Experience Rating and Moral Hazard in Insurance Markets

Honglin Li and Hao Zheng*

November 7, 2024

(Click Here for the Latest Version)

Abstract: Experience rating, in which insurance premiums are adjusted based on claim histories, is a common practice in many markets. In this paper, we empirically and theoretically examine how changes in experience rating impact behavioral responses and social welfare. First, we empirically estimate the effects of increasing experience rating on risky driving and ex-post claiming behaviors in the Chinese auto insurance market, by exploiting the staggered rollout of a policy change that increased the penalties for experiencing claims across the nation and leveraging administrative data for a 5% sample of automobile insurance policies in China over four years (more than 7 million policies). We find a decrease in monthly claim rate for liability coverage by 28% (2 claims per 1,000 policies), which can plausibly be attributed to reductions in risky driving and accidents. At the same time, we estimate a sizable reduction in reporting of small-sized losses using collision damage coverage claims, over which insureds have more discretion about claiming. We develop a model of insurance with experience rating to understand the welfare implications. Our model highlights the key tradeoff that experience rating penalties worsen risk protection but can provide incentives to reduce the moral hazard effects of insurance. Reducing accidents has a clear welfare benefit, while the welfare benefits of reducing insurance claims ex-post for losses that occurred are ambiguous. Using sufficient statistics derived from the model, we estimate that the policy change of increasing experience rating improves social welfare by about $33 per insured driver per year.

*Li: Wisconsin School of Business, University of Wisconsin-Madison, hli434@wisc.edu. Job Market Paper. Zheng: School of Insurance and Economics, University of International Business and Economics, zhenghao@uibe.edu.cn. Honglin thanks Justin Sydnor, Dan Sacks, Ty Leverty, and Naoki Aizawa for their invaluable mentorship and advice. We also thank Panle Jia Barwick, Dani Bauer, Stuart V. Craig, Yijia Lin, Xiaolu Liu, Anita Mukherjee, Philip Mulder, Richard Phillips, Casey Rothschild, Joan Schmit, Peng Shi, Ashley Swanson and the participants at the PhD Student Research Meeting and R&I Workshop at the UW-Madison, 2024 PhD Student Research Symposium at the University of Georgia and 2024 ARIA for helpful comments and suggestions. All remaining errors are our own.

1 Introduction

Experience rating, in which insurance premiums are adjusted based on claim histories, is a very common practice in many markets, such as auto insurance (Abbring et al., 2003b; Dionne et al., 2013; Abbring et al., 2008; Jeziorski et al., 2017), health insurance (Daley et al., 2012; Craig, 2022), workers' compensation (Ruser, 1985), and unemployment insurance (Johnston, 2021; Spaziani, 2023). A key rationale for experience rating is to help with risk classification. Additionally, experience rating, similar to partial insurance, can provide incentives for risk prevention and mitigate the moral hazard effects associated with insurance coverage (Pinquet, 2013). Despite its widespread use, there has been relatively less economic research that comprehensively evaluates behavioral and welfare implications of experience rating as a mechanism to reduce moral hazard. We bridge this gap by providing empirical evidence that distinguishes between different types of moral hazard behaviors, quantifying their relative importance, highlighting the welfare trade-off between risk protection and moral hazard reduction under experience rating, and quantifying the overall welfare impact of these changes.

This paper examines behavioral and welfare impacts of increasing experience rating incentives in insurance markets. We empirically estimate the effects of increasing experience rating penalties using a large-scale natural experiment that increased the penalties for experiencing claims in the Chinese auto insurance market. Our estimates indicate a large reduction in risky driving and a sizable reduction in reporting of small-sized claims. We then develop a general model of insurance with experience rating to shed light on the fundamental tradeoffs associated with the increase of experience rating. Using a sufficient statistic formula derived from the model, combined with our empirical estimates, we show that the policy change of increasing experience rating improves social welfare by $33 per insured driver per year.

Experience rating in auto insurance markets, also known as the Bonus-Malus system, rewards drivers with premium discounts for accident-free years and imposes premium penalties on those who file claims. It is widely used by insurance companies and is also commonly adopted as government policy in many countries.[1] In 2015, the Chinese insurance regulator initiated a reform that significantly increased the impact

[1] In some countries, such as China, France, Portugal, and the Netherlands, the Bonus-Malus system is fully regulated by the government and standardized across insurance companies. In countries where it is not government-regulated, such as the US, insurance companies typically set their own experience rating schemes.

of past claims on the premiums for voluntary coverage.[2] For example, before the reform, the premium for consumers with one claim in the previous year was 10% higher than for those with no claims; after the reform, this difference rose to 15%. The policy change was rolled out across the country in a staggered fashion.

To provide causal estimates of the reform's effects, we exploit the staggered roll out of the reform and employ difference-in-differences (DID) approaches, which use regions that had not increased experience rating as a comparison group for those that had. We combine the natural experiment with comprehensive administrative data. Our primary dataset includes a 5% sample of auto insurance policies in China, spanning a 4-year period and covering over 7 million policies. The dataset provides detailed information at the policy and claim level. The policy data include coverage choice, premiums, demographics of the insured, risk class (i.e. the insured's past claim history), and vehicle characteristics. The claim data, matched with policy level data, include information on the types of coverage under which the claim was filed, the time of filing, and the size of the claim.

We begin by presenting evidence that increasing experience rating penalties leads to a reduction in risky driving behavior. A challenge to examining risky driving behavior in insurance is that we typically observe claims, rather than actual accidents. To overcome this challenge, we focus on liability claims for damages to third parties, over which the insured has limited discretion about claiming. We find a decrease in claim rate for liability coverage by about 28% (2 claims per 1,000 policies per month). This suggests that the policy led to approximately 0.8 million fewer multi-car accidents per year in China once fully implemented nationwide. These results provide the first evidence from a large-scale natural experiment that increasing experience rating penalties meaningfully reduced accident risk.

In addition to reducing accidents, we further show that increased experience rating reduces claim filing in cases where accidents do occur. Voluntary collision coverage provides an ideal context for examining ex-post claiming behavior because it insures against losses to the insured's own property when the driver is at fault, giving the insureds considerable discretion in deciding whether to file a claim. However, increases in experience rating in voluntary insurance markets may lead insureds to skip filing claims for smaller

[2] Auto insurance in China includes two types of coverage: compulsory and voluntary. Compulsory coverage protects against losses to third parties, while voluntary coverage provides additional protection on top of compulsory coverage. The policy change only impacted voluntary coverage, leaving compulsory coverage unchanged.

losses, but also change incentives for purchasing voluntary coverage, leading to changes in composition of risks among buyers. To obtain clean estimates of incentive effects on claim rates and sizes, we restrict our analysis to individuals who always purchased collision coverage throughout the sample period. The monthly claim rate for collision damages decreases by about 15 claims per 1,000 policies following the increase in experience rating (52% of the pre-reform mean). In addition, the average claim size increases by about 690 yuan. Examining the distribution of claim sizes, we observe a clear reduction in claims for small-dollar collision losses after the reform. These findings are consistent with the presence of ex-post moral hazard effects. Moreover, there is a smaller reduction in claiming for much larger losses that is consistent with the reduction in accident rates estimated from the liability coverage discussed above. This implies that both reductions in risky driving and in reporting small-sized losses contribute to reductions in claim rate for voluntary collision coverage. Assuming a constant semi-elasticity of risky driving with respect to experience rating, reductions in ex-post claiming account for approximately 46% of the decrease in collision claim rates.

While reducing moral hazard is generally beneficial, increasing experience rating can have potential downsides. To better understand the welfare implications of our empirical findings, we develop a model of insurance with experience rating, accounting for different types of behavioral responses. One key downside of increased experience rating penalties is that it reduces risk protection, resulting in reduced consumption smoothing across states of the world. The loss of risk protection, however, can be offset by the benefits of reduced moral hazard. Our model highlights that it is important to separate risk mitigation behavior (ex-ante moral hazard)[3] and the reporting and claiming of realized losses (ex-post moral hazard)[4]. Reductions in ex-ante moral hazard have positive welfare effects: lower accident rates result in lower average costs, which

[3]Insurance leads to ex-ante moral hazard when it reduces the level of (unobserved and costly) prevention effort the agent chooses (Abbring et al., 2008; Jeziorski et al., 2017; Gong, 2017; Einav and Finkelstein, 2018). This aligns with the notion of hidden action in the literature (Bolton and Dewatripont, 2005), and is similar to ex-ante moral hazard discussed in health insurance literature (Ehrlich and Becker, 1972; Spenkuch, 2012).

[4]Ex-post moral hazard broadly involves the level of insured costs conditional on a loss event or harm occurring. In our context of property insurance, the ex-post moral hazard behavior is whether or not to file an insurance claim after experiencing an accident because premium penalties are triggered by claiming or not (Abbring et al., 2008; Jeziorski et al., 2017; Gong, 2017). This is consistent with the notion of a "pseudodeductible": a latent, unobserved threshold that determines whether or not an insured loss is large enough to trigger the policyholder to file a claim (Braun et al., 2006). Ex-post moral hazard has been studied extensively in health insurance settings, where the idea is that insurance by lowering the cost of medical services can cause individuals to inflate their use of services (and hence the size of the insurance claim) relative to what they would use if they were not insured (Pauly, 1968; Cutler and Zeckhauser, 2000; Einav and Finkelstein, 2018).The difference of ex-post moral hazard between health insurance and property insurance is natural because experience rating in property creates incentives for claiming or not claiming but not differential incentives for the size of loss/spending, as seen in cost sharing in health insurance.

in turn reduced premiums for insureds and ultimately improves consumer welfare. Therefore, if increased experience rating penalties reduce accidents by a sufficient amount to offset the loss of risk protection, it can improve overall welfare. Typically, premiums will capture both expected losses to the insured's property and damages to others via liability coverage.[5] In contrast, the welfare implications of binary claim-decision version of ex-post moral hazard are ambiguous. If there are no additional frictions or costs associated with filing claims, increased experience rating penalties are welfare-reducing regardless of the impact of ex-post moral hazard. This is because any premium reductions from not filing claims for some losses are fully offset by the individuals' need to bear those losses themselves. As a result, marginal increases in experience rating penalties in the presence of ex-post moral hazard only reduced consumption smoothing. Ex-post moral hazard effects can be welfare-improving, though, if there are frictions or additional costs associated with filing or processing claims. In this case, ex-post moral hazard responses have impacts similar to ex-ante moral hazard, but only through the channel of reduced "friction costs" associated with filing claims, not from the forgone claims themselves.

We derive a sufficient statistics formula from our model to quantify the net welfare benefit of the experience rating reform and combine it with our empirical estimates. We calculate a net welfare gain of approximately \$33 per insured driver per year from the policy change. About \$20 of this net welfare gain comes from the voluntary coverage for which the experience rating program and policy change are in place. Another \$12 of the welfare gain comes from reductions in accidents that damage third parties and are internalized by the compulsory liability system. The estimated welfare gain of \$33 per insured driver per year presented above include only the costs that come through insurance (property and liability) but not additional uninternalized externalities. Therefore, these estimates should be considered as a lower bound on the reform's total welfare impact.

This paper contributes to three strands of literature. First, we contribute to the literature on moral hazard and risk protection. A large body of literature has examined how moral hazard responses affect insurance design. A central result from this literature suggests that full insurance is suboptimal; instead, partial coverage, such as deductibles, is warranted to mitigate moral hazard (Pauly, 1968; Holmström, 1979;

[5] Reductions in costs through liability coverage will capture some of the externality costs of risky driving, though there may be additional externalities from risky driving that are not fully incorporated into the liability system. For example, the full value of lost life from mortality events may not be covered in the liability system and would constitute an un-internalized externality.

Winter, 2013; Einav and Finkelstein, 2018).An extensive empirical literature in health insurance estimates the magnitude of moral hazard responses to changes in coverage levels to inform evaluations of policies and assess optimal coverage levels. (Einav et al., 2015, 2017; Marone and Sabety, 2022). Prior literature on moral hazard and insurance design has rarely considered experience rating except for Jeziorski et al. (2017), who developed a model of coverage choice and effort choice-efforts that alter risk levels-to estimate key parameters such as the cost of effort and risk aversion. They found that while the Portuguese experience rating system was ineffective at sorting consumers by risk type, it successfully incentivized safer driving. Our paper, focusing on moral hazard, first highlights that the fundamental tradeoffs between risk protection and moral hazard remain similar under the design of experience rating as they are under partial insurance. Moreover, experience rating introduces additional potential for individuals not to file claims, which differs from incentives created by coverage levels. This highlights the importance of disentangling the ex-ante from ex-post response when evaluating the design of experience rating penalties. Failing to distinguish between ex-ante and ex-post moral hazard can lead to setting an experience rating penalty that is higher than optimal. While the two channels have been previously mentioned and discussed in the literature, our paper is the first to combine these forces and provide the welfare analysis of experience rating. In addition, we contribute to the literature by providing a sufficient statistic formula to quantify the welfare implications of modifications to experience rating (Einav et al., 2017; Chetty, 2009).

Second, this paper contributes to the empirical literature examining the impact of experience rating on driver behavior in the context of auto insurance. Several previous papers are especially relevant: they employ dynamic insurance data to test for moral hazard by examining whether claim probabilities decrease after insureds experience more claims (Abbring et al., 2003a; Dionne et al., 2013; Abbring et al., 2008, 2003b). The findings on the existence of moral hazard are mixed in previous literature. However, testing for negative correlation between claim probabilities and claim frequencies may capture factors other than moral hazard. For example, Shum and Xin (2022) find that drivers may adopt more conservative driving habits after "near-miss" accidents due to time-varying risk aversion. We contribute to the literature by exploiting quasi-experimental variation to causally identify moral hazard, providing a more direct test of the impact of changes in experience rating on moral hazard.[6] Furthermore, we quantify the magnitude of moral

[6]In addition to studies using experience rating schemes to identify moral hazard effects, there is also a related empirical literature testing for moral hazard reactions to coverage levels in auto insurance (Weisburd, 2015; Cohen and Dehejia, 2004).

hazard behaviors rather than merely testing for their existence, allowing for empirical estimates relevant to welfare analysis. To the best of our knowledge, we are the first to quantify the magnitude of moral hazard effects under experience rating. More importantly, we are the first to separately quantify the magnitude and relative importance of the ex-ante and ex-post moral hazard channels. Abbring et al. (2008) shows existence of ex-post moral hazard by comparing the distributions of first and second claim sizes. In contrast, our reduced-form methods allow us to directly test both channels and quantify their relative importance under modest assumptions, which enables a deeper understanding of the welfare effects. A prior study by Dionne and Liu (2021) used a similar pilot experience rating reform in a Chinese city to test for moral hazard effects and found the reform decreased claim frequency. However, the pilot reform coincided with a new requirement for insurers to share past claims data across all companies.[7] Compared to Dionne and Liu (2021), we avoid such confounders by exploiting a cleaner natural experiment. We also use a larger and richer data to separately identify ex-ante and ex-post moral hazard and couple it with the model to quantify welfare effects.

Finally, we contribute to the literature on policies designed to deter risky driving behavior, a critical global public health issue causing significant loss of life and substantial property damage. An estimated 1.3 million people die each year as a result of road traffic crashes (WHO, 2022a,b). In China alone, over 250,000 lives were estimated to have been lost in 2019 (WHO, 2022a). Even in developed countries, risky driving and roadway fatalities are still significant issues. For example, motor vehicle crashes in the United States are estimated to cause over 40,000 death annually and cost more than $340 billion, roughly 1.6% of the country's GDP (WHO, 2022a; NHTSA, 2023; Blincoe et al., 2022). Recognizing the need for effective policies to improve road safety, previous studies have examined a range of policies, including graduated driver licensing (Karaca-Mandic and Ridgeway, 2010), the minimum legal driving age (Huh and Reif, 2021), stricter drink driving laws (Francesconi and James, 2021), mandatory seatbelt laws (Carpenter and Stehr, 2008), and penalty-point system (Abay, 2018; Rebollo-Sanz et al., 2021), among others. We provide evidence that experience rating in auto insurance can be a potential effective tool to deter risky behavior, albeit at the expense of reduced risk protection.

[7] When insures do not share data with each other, there exists the possibility for drivers to circumvent penalties by switching insurers. The simultaneous implementation of data sharing requirements eliminates this possibility. Therefore, the cessation of avoiding penalties due to the data sharing requirements might result in an overestimate of the effects of changes in experience rating schemes.

The paper proceeds as follows. Section 2 introduces Chinese auto insurance markets, explains the policy changes related to experience rating, describes data used in our analyses and presents summary statistics and premium changes. Section 3 presents our empirical results. Section 4 provides a simple model of insurance with experience rating and discusses the welfare implications of ex-ante and ex-post moral hazard. Section 5 discusses the welfare impacts using sufficient statistic approach. Section 6 concludes.

2 The Chinese Auto Insurance Market

Our empirical application focuses on the Chinese auto insurance market. In recent decades, vehicle ownership by households has risen sharply, from fewer than 50 million vehicles in 2009 to over 250 million in 2021. Along with this, there was a sharp increase with auto insurance purchases. The total gross written premiums of the auto insurance market in 2021 were $108.4 billion (Fang and Xu, 2023).

2.1 Auto Insurance Coverage in China

Auto insurance coverage in China is similar to that in developed countries such as the US. In China, there are two types of automobile insurance coverage: the compulsory traffic accident liability insurance policy (compulsory liability coverage hereafter) mandated by the government and the voluntary personal automobile insurance policy (voluntary coverage hereafter) (Fang and Xu, 2023).

Compulsory coverage. The compulsory policy covers financial losses incurred by a third party (including both bodily injury and property damage) for which the policyholder is liable while operating the vehicle. The coverage amounts for third party's bodily injury and property damage are standardized by the China Insurance Regulatory Commission (CIRC), the regulator of insurance industry in China. During our study period, the compulsory policy offered coverage up to 110,000 yuan (approximately 16,000 USD) for death and disability compensation, 2,000 yuan (approximately 300 USD) for property damage, and 10,000 yuan (approximately 1,500 USD) to cover medical expenses. The premium of compulsory policies is uniformly set at 950 yuan (about 135 USD) nationwide and is provided through commercial insurers.[8] Drivers are required to display a sticker on the vehicle's window to show compliance with the compulsory policy purchase,

[8] Note that the final premium for a driver's compulsory policy is adjusted by a factor based on the driver's past claim record, i.e. experience rating factor. But experience rating scheme for compulsory coverage is different from that for voluntary coverage. During our study period, the experience rating scheme for the compulsory coverage did not change.

making noncompliance easily detectable and resulting in a nearly 95% take-up of compulsory coverage.

Voluntary coverage. On top of the compulsory coverage, drivers have options to purchase different types of voluntary coverage. A voluntary policy is a standardized contract that provides comprehensive coverage by integrating multiple insurance agreements. There are two major voluntary policies purchased by consumers: additional third-party liability coverage (voluntary liability coverage hereafter) and collision damage coverage (voluntary collision damage coverage hereafter). Additional third-party liability policies cover the third party's financial losses that exceed the coverage limits of the compulsory policy, when the driver is at fault. When purchasing an additional third-party liability policy, drivers select the coverage amount, which covers losses up to a specified coverage limit. Typically, the available coverage amounts are available in fixed increments, ranging from 100,000 to 1,000,000 yuan (approximately \$14,000 to \$140,000), in steps of 50,000 or 100,000 yuan. A collision damage policy covers the driver's own property losses when the driver is at fault. [9] The insured value provided by a collision damage policy is determined by the covered vehicle's market value. The market value of a vehicle depends on the vehicle's age, manufacturer, model and other characteristics. By default, both voluntary liability coverage and voluntary collision coverage only cover 85% of the losses. The insured can further pay an extra 15% of the premiums to avoid the coinsurance, which is a more prevalent practice.

Premium setting for voluntary coverage. The premiums for voluntary policies are highly regulated by the government and are typically determined by three multiplicative factors: the base premium, the experience rating factor, and the insurer factor. We discuss the premiums for voluntary liability policy and collision damage policy separately. The premium of an additional voluntary liability policy for individual i, vehicle v, coverage level j from insurer f in region r, denoted as p^{tl}_{ivjkr}, is calculated as follows:

$$p^{\text{tl}}_{ivjfr} = \underbrace{b^{\text{tl}}_{jvr}}_{\text{base premium}} \times \underbrace{e_i}_{\text{experience rating factor}} \times \underbrace{l_{fr}}_{\text{insurer factor}}, \tag{1}$$

where b^{tl}_{jvr} represents the base premium, which is a function of coverage limit j, region (province or city) r, and vehicle v's seating capacity, regardless of insurer and individual characteristics. The base premium

[9] Note that the collision damage policy doesn't include thefts, break-ins, or vandalism (non-collision damage/loss).

is determined by the regulator. e_i represents the individual i's experience rating factor. Each driver is in a risk class, denoted by k_i, from 5 or more claims in the past policy year, to 0 claims in the last three policy years. Each driver's risk class k_i corresponds to a specific factor. For example, before the policy change, an individual in the 5 or more claims in the past year category would have an experience rating factor e_i of 1.3, whereas $e_i = 0.7$ for those with with no claims in the last three years. The experience rating factor, also referred to as the No Claims Discount (NCD) factor, is determined by the regulator. We provide a detailed discussion on the regulation and its changes in Section 2.2. Last, l_{fr} represents the factors of insurer f in region r. The regulator regulates the range of l_{fr}, within which insurers have discretion to set their own factors. Although the Chinese auto insurance market is highly regulated, it is pretty competitive, with 5 to 30 insurers serving in each province. Based on Zheng et al. (2021), the HHI at the province level ranges from 1343 to 5075, with 25 out of 32 provinces below 2500, indicating the market is from competitive to moderately concentrated. Given this competitive environment, most insurers adopt the lowest limit of the government-regulated range for their insurer factor.

Similar to premiums of voluntary liability policies, the premium, denoted as p^{cd}_{ivkr}, for voluntary collision coverage for individual i with vehicle v from insurer f in region r is as follows:

$$p^{\text{cd}}_{ivkr} = \underbrace{b_r(V_v)}_{\text{base premium}} \times \underbrace{e_i}_{\text{experience rating factor}} \times \underbrace{l_{fr}}_{\text{insurer factor}}, \tag{2}$$

where $b_r(V_v)$ is the regulated region-specific base premium, depending on V_v, the value of the covered vehicle v in region r. Experience rating factor e_i and insurer factor l_{fr} are the same as those for voluntary liability coverage and subject to the same regulation.

Enrollment and claim process. A standard automobile insurance policy, regardless of the type of coverage , offers coverage for a one-year period. At the start or renewal of the policy, drivers are required to purchase compulsory insurance coverage. Additionally, drivers can decide whether to purchase voluntary coverage and, if purchasing, what type of voluntary coverage. The total premium is the sum of compulsory coverage premium and the premiums for any selected voluntary coverage. [10]

[10] Suppose a consumer decides to purchase third-party liability coverage of 100,000 yuan and collision damage protection coverage of 100,000 yuan. The base premium for the 100,000 yuan third-party liability coverage is 2,124 yuan, and for the collision damage coverage, it is 1,761 yuan. The total base premium is 3,885 yuan. This total is then multiplied by the experience rating

During the enrollment year, a loss might occur. When a loss occurs, the insured has two options: filing a claim with the insurer to get covered under the relevant policies or incurring the loss without filing. To compensate third parties' losses, the at-fault insured can file a claim under their compulsory coverage. If the loss size exceeds the compulsory policy's limit, the insured may also file a claim under their voluntary liability coverage. Should the insured opt not to file for either compulsory or voluntary coverage, the insured must negotiate a mutually agreeable compensation amount with the third party, which can be time-consuming and risky. To cover the losses to their own vehicle when the driver is at fault, either involved with third parties or not, the driver can either claim under their collision damage policy or pay the repair costs themselves. This decision typically doesn't involve third parties and is solely at the driver's discretion. Overall, the at-fault party has more discretion regarding filing claims for collision damage coverage compared to compulsory and voluntary liability coverage.

2.2 Policy Changes of Experience Rating

The regulator initiated a series of reforms on premium setting in the automobile insurance industry beginning in 2015. These reforms started with changes in regulations on premium setting for voluntary coverage. We will use this policy change in the premium setting as a natural experiment to investigate its impact on consumer selection and moral hazard behavior.

Foremost, the regulator updated the experience rating factors. Figure 1 shows experience rating factors for various claim history scenarios before and after reform. It introduced more stringent penalties for drivers with a history of multiple claims, while concurrently offered more generous discounts for those who have maintained a claim-free record for one or more consecutive years. The policy change only applied to experience rating factors of voluntary coverage and did not impact those of compulsory coverage. It is important to note that the opportunity cost of an additional claim increased for drivers of all risk classes after the policy change, as shown in Figure 1b. For example, before the reform, drivers of lowest risk class received 30% discount off the base premium when having no claim , but had to pay the full base premium when having a claim. In contrast, the discount of having no accident/claim increased to 40%, suggesting

factor of the driver. For example, if the driver has one claim in the past year, the experience rating factor would be 1. Additionally, the insurance company might provide an additional discount or surcharge. In this instance, assuming the insurer offers a 20% discount, the total premium faced by the consumer would then be 4,058 yuan (calculated as $3,885 \times 0.8+$ compulsory coverage premium of 950 yuan).

larger incentives for consumers even in the lowest risk classes to avoid accidents/claims.

In addition to changes in experience rating factors, the regulator revised base premium and insurer factor. First, the regulator updated base premiums based on the latest data collected by the regulator. Overall the base premium for all types of coverage slightly decreased after the reform. Second, the regulator granted insurers more discretion in premium adjustment. Before the reform, the insurer factor was typically set at one, unless special approval was granted by the regulator, which was very rare. In contrast, the insurers have greater flexibility to tailor premiums based on underwriting conditions and sales channels, with the insurer factor ranging from 0.7 to 1.4. However, due to the potentially competitive environment, most insurers adopt lowest insurer factors. As estimated in our data, the mean insurer factor is 0.73 with a very small standard deviation at 0.05.[11]

To summarize, the policy change introduces premium adjustments through two main channels. First, it changes dynamic incentives for drivers by altering the cost associated with claims through increased penalties and discounts, i.e. changes in experience rating factors. Second, it adjusts the static components of the premiums, i.e. base premium and insurer factor. The overall shifts in realized premiums after the reform are predominantly attributed to changes in experience rating factors. We elaborate more on this in Section 2.5.

The policy change rolled out across the country in a staggered fashion. The regulator grouped the 36 provinces/cities in the country into three reform groups[12], with the first group of 6 provinces/cities starting on June 1st, 2015, the second group of 12 on January 1st, 2016, and the remaining 18 on July 1st, 2016. These dates are the declared in the official announcements, which were typically released three months before the formal start dates. In practice, the reform started one or two or three months earlier than the formal dates specified in the announcements in several provinces. For example, one province in the third group started the reform three months ahead of the schedule, three provinces began two months earlier, and a few more started only one month in advance. The early start time is verified in the media coverage and in our data, where we observe provinces adopting new experience rating factors before the publicly announced start dates. Given that different provinces had slightly different actual start dates from the announced start dates and the

[11] There are additional changes in the reform, for example, giving the insurer the freedom to design their own insurance contract. These additional changes are not binding and no insurers have implemented other changes by the end of our sample period.

[12] The 36 provinces/cities include all provinces and several important cities, including Beijing, Shanghai, Tianjin, Chongqing, Shenzhen, Ningbo, Dalian, and Qingdao.

advance notice period of three months, our empirical analysis sets the start dates for each reform group to three months before the officially announced start dates. The first three months after announcement month are interpreted as partial reform effects of early implementations as well as anticipatory effects. Section F provides details of the timeline of the reform in each reform group.

2.3 Data and Sample

Our primary dataset is collected and obtained from the regulator CIRC. Per regulation, all auto insurance companies in China are mandated to report their underwriting and claim data to the regulator for rate-making purposes, allowing us to observe individual-level policy and claim data. The sample comprises a panel dataset including 5% of all auto insurance policies issued between June 1st, 2013 (over two years before the reform) and June 30th, 2017 (one to two years after the reform).[13] We exclude two provinces that used different experience rating code systems and another three provinces/cities that had missing or incomplete experience rating codes prior to the reform. Our final analytical sample consists of approximately 1.7 million vehicles and more than 7 million policies from June 1st, 2013 to June 30th, 2017.

The dataset includes a comprehensive set of variables used in insurer ratemaking, with policy data on a policy(-year) level and claim data at the individual claim level. First, we observe each consumer's selection of voluntary coverage, including the type and amount of coverage, along with the corresponding premiums paid. Second, the dataset includes the driver's annual experience rating factor, as well as individual and vehicle characteristics. Third, the claim data include detailed information of all claims under compulsory coverage and different types of voluntary coverage, including claim filing time, claim amount, and the driver's percentage of liability in an accident.

2.4 Summary Statistics.

Table 1 reports descriptive statistics separately for each type of coverage. The observation level is at policy/claim level, with more than 7.6 million compulsory policies. There are about 7.6 million voluntary liability policies and 6 million voluntary collision damage coverage, implying about 100% and 80% of

[13]The sample, provided by the regulator, is restricted to vehicles that had at least one voluntary policy both before and after the reform. However, in practice, the rate of purchasing any type of voluntary coverage, esp additional voluntary liability coverage, is about 97%. Therefore, the concern about this sample selection is not a big issue. Note that this sample selection might limit our ability to understand the entry and exit of drivers over time, but we will have a cleaner interpretation that our observed changes are mostly from behavioral changes over time instead of composition changes.

consumers purchased voluntary liability and collision policies, respectively.

The claim rate for compulsory liability coverage is 0.08, implying about 8 claims per 100 compulsory policies annually. In comparison, the claim rate for voluntary liability coverage is 4 claims per 100 voluntary liability policies and for voluntary collision coverage is about 34 claims per 100 policies. The reason for lower claim rate for voluntary liability coverage is that it covers additional third parties' losses that exceed the compulsory insurance's coverage limit, which are usually large sized losses and less frequent. The claim rate for voluntary collision coverage is much higher than the other two types of coverage, because collision damage policies cover the driver's own losses. This is also higher than the claim rate in other studies; for example, the annual claim rate of "comprehensive" coverage, similar to our collision damage coverage, was 24.5% in Israel from 1994 to 1999 (Cohen and Einav, 2007). The insurers' average cost, including payouts and zeros, is lowest for compulsory liability coverage at 275 yuan, and highest for voluntary collision coverage at 927 yuan. Conditional on claiming, the average claim size is much larger among compulsory and voluntary liability coverage, suggesting losses of 3,500 yuan and 8,000 per claim. The claim size for voluntary collision coverage is smaller, at about 2,700 yuan per claim. All statistics have quite large standard deviations, implying potential heterogeneity across consumers.

2.5 Effects on Premiums of Voluntary Coverage

Figure 2 displays the evolution of premiums of voluntary liability coverage with coverage limit at 300,000 yuan, as an example, and premiums of voluntary collision coverage, separately for drivers with different rating classes. First, premiums for different risk classes change differently after the reform. Premiums increased sharply for consumers in higher risk classes, i.e. those with 3 or 2 or 1 claims in the last year. In contrast, consumers in lower risk classes, i.e. those with no claims in last two or three years, experience a drop in their actual premiums. Additionally, actual premiums for consumers with no claims in the previous year are stable after the reform. Further, this group of consumers had a very small decrease in the experience rating factor from 0.9 to 0.85, implying that the premiums before multiplying the experience rating factor, that is, the base premium and the insurer factor, had a very small change. Overall, it suggests that adjustments in the experience rating factor introduced by the reform were effectively implemented, and changes in the experience rating factors primarily drive the changes in the total premiums.

3 Empirical Analysis

In this section, we start with introducing empirical strategy. Then we investigate the incentive effects of increasing experience rating on risky driving behaviors using liability coverage and on ex-post claiming behavior using collision coverage. We end with summary of the results.

3.1 Empirical Strategy

Dynamic Effects. To provide causal effects of increasing experience rating factors, we employ a event-study style two-way fixed-effect difference-in-differences (TWFE DID) approach as our primary empirical strategy. The estimating equation of TWFE DID regression is

$$y^c_{irt} = \sum_{e \neq -1} \gamma^c_e D^e_{irt} + \mu_r + \tau_t + \varepsilon_{irt}, \tag{3}$$

where the outcome variable is the whether there is a claim or the claim size of driver i with coverage type c in the province/city r in year-month t, D^e_{irt} is an event-time dummy equal to 1 if consumer i in region r has been treated for exactly e periods in period t and 0 otherwise. For example, D^0_{rt} equals 1 for provinces that just become treated in period t and 0 otherwise. μ_r denotes province/city fixed effects and τ_t denotes calendar year-month fixed effects. Coefficient α^c_e measures the difference in the outcome variable between treatment and control group e months after the announcement of the reform relative to the baseline period, which is set to the period just before the reform announcement. Further, we interpret the coefficients α^c_0, α^c_1, and α^c_2 as the anticipatory and/or partial reform effect as 1) drivers might react to policy announcement by changing behaviors in advance; and 2) some provinces starts earlier after the announcement. Standard errors are clustered at the province/city level. The dynamic TWFE specification yields sensible causal estimates, under the assumptions of parallel trends and homogeneous treatment effects across groups (Sun and Abraham, 2021). We will discuss the assumptions when presenting the results.

Average Treatment Effects. To provide an overall estimate of the reform effects, we turn to stacked DID as the baseline specification. Our dynamic TWFE regressions yield sensible estimates under the assumptions of parallel trend and homogeneous treatment effects across groups,. This allows the dynamic effects to be heterogeneous across time. However, the static TWFE regressions might produce biased estimates of the average treatment effects under the heterogeneous across time (Callaway, 2022). The biased estimates arise from "bad comparison", i.e. a comparison between a late-treated group to an already-treated group. Therefore, for overall average treatment effects results, we rely more on the stacked DID estimates. We construct two stacks for regressions: one with the first reform group as the treatment group and another with the second reform group as the treatment group, both with 20 and 12 months before and after the announcement.[14] We pool the two stacks for estimation. The estimating equation of stacked DID is

$$y^{c}_{irt} = \beta^{c}_{1}\text{anticipation}_{irst} + \beta^{c}_{2}\text{post}_{irt} + \sum_{e\neq -1} d^{e}_{rst} + \mu_{rs} + \tau_{ts} + \varepsilon_{irst}, \tag{4}$$

where the anticipation $_{irst}$ equals to one when the event dummies 0, 1, or 2 equal to 1, that is, treated within three months of the announcement month, otherwise 0. post$_{irst}$ equals to 1 for driver i in provinces/cities r stack s that are treated in period t and 0 otherwise. β^{c}_{1} is interpreted as the average anticipatory/partial reform effects, and β^{c}_{2} is interpreted as the average treatment effects of the reform. Standard errors are clustered at the province/city level.

Robustness. The first robustness check we conduct is to estimate the simple event study model, without controlling calendar year-month fixed effects compared to the TWFE model. The simple event study model provide estimates of average changes in outcome variables over time relative to baseline period by re-centering the event period as well as controlling for different levels across reform groups.

Recent literature on DID suggests that dynamic TWFE estimates yields a sensible causal estimand when heterogeneity is only related to the time since treatment and may not be robust in the presence of heterogeneous effects across treatment groups (Goodman-Bacon, 2021; Callaway, 2022; Roth et al., 2023). To mitigate any concerns about potential biases in our TWFE estimates, we use the stacked DID approach (Butters et al., 2022) and the approach proposed by Callaway and Sant'Anna (2021) to provide alternative dynamic treatment effect estimates. However, using the methods suggested by recent literature comes with

[14]To be specific, for the first stack, provinces in the first reform group serve as the treatment group, and the periods cover from July 2013 to Mar 2016, which is the latest month before the third group is treated. To construct the control groups for the first reform group in the first stack, we use provinces in the third reform group from July 2013 to Mar 2016 (20 months from July 2013 to Feb 2015, the month just before the announcement of reform in the first group, and 13 months including and after the announcement month of the reform, from Mar 2015 to Mar 2016). In addition, we can also include provinces from the second reform group from July 2013 till Sep 2015, the month before the announcement of the reform in the second group. Similarly, the second stack consists of data from the second and third groups from Feb 2014 (the 20th month before the reform announcement for the second group) till Mar 2016 (5 months after the reform announcement).

costs. These approaches emphasize constructing better controls for the treatment group and avoiding unfavorable comparisons between late-treated and already-treated groups. Typically, never-treated groups are used to construct suitable control groups. In our specific context, we do not have a never-treated group; instead, we can only use not-yet-treated groups. Consequently, our ability to estimate dynamic effects is limited due to short intervals between reform start dates in the three reform groups. Furthermore, we lose some statistical power, as the third reform group cannot be matched with a suitable control group and is only used as a control group for estimating the reform effects. We defer the discussion of implementation and results in Appendix B.

3.2 Effects on Liability Coverage Claim Rate and Size

First, we use compulsory liability coverage to understand whether increasing (voluntary) experience rating factors leads to any reductions in risky driving. Compulsory liability coverage is ideal for testing changes in risky driving, induced by larger penalties of voluntary coverage, for several reasons. First, compulsory liability coverage covers third parties' losses when the driver is liable, and the insureds have limited discretion regarding filing claims for liability coverage. Additionally, compulsory liability coverage was not affected by the changes in experience rating factors, so the claiming incentives remain constant before and after the reform. As a result, if drivers are driving more carefully, the claim rate for compulsory coverage will decreases. Further, any such changes in claim rate for compulsory coverage would come from changes in driving behavior, i.e. changes in ex-ante moral hazard effects.

Figure 3 displays time series of claim rate and size of compulsory coverage. Prior to the reform, the claim rates of compulsory coverage were relatively stable, fluctuating around 1.2 claims per 100 policies, with the level slightly higher in the first reform group. Upon the announcement of the reform in a region, we observed a decline in the claim rate in the region relative to regions where the reform had not yet been introduced. Ultimately, the claim rate across all regions converged to a level below 1 claim per 100 policies. In contrast, there were no significant changes in the claim size around the time of the reform. We interpret this pattern as strong evidence of a reduction in risky driving. Figure 4 display TWFE results Claim rate of compulsory coverage decreases by about 1-2 claim per 1000 policies monthly (13%-34% of prior reform average rate) and there are no significant changes in claim sizes over time. These findings collectively suggest that consumers may exercise greater caution in their driving behavior, indicative of reductions in ex-ante moral hazard.

Drivers can purchase voluntary liability coverage on the top of compulsory liability coverage. Though voluntary, the purchase rate is as high as about 100%. Similar to compulsory coverage, the insureds have limited discretion regarding filing claims. We further use the claim rates and sizes for voluntary liability coverage to examine the changes in risky driving behaviors. Appendix Figure 9 displays the trends in claim rate and claim size and Appendix Figure 10 display event study and TWFE results for voluntary liability coverage. The results for voluntary coverage imply a decrease in claim rate and stability of claim size.

Column (1) in Table 2 report the average treatment effect results using stacked DID. The monthly claim rate dropped by 2 per 1,000 claims after reform (28% of pre-reform mean), and the claim size slightly increase but is very small (0.25% of pre-reform mean) and insignificant. Overall the reductions in claim rate and stability in claim size for liability coverage show significant reductions in risky driving behaviors.

3.3 Effects on Voluntary Collision Coverage

We turn to voluntary collision coverage to understand claiming behaviors. Voluntary collision coverage insures against losses to the insured's own property when the driver is liable regardless of whether there is a third party involved. The insureds have considerable discretion in deciding whether to file a claim or personally absorb the loss, making this type of coverage more susceptible to ex-post moral hazard. As a result, when experience rating penalties increase, both reductions in risky driving behaviors and in reporting small-sized losses could contribute to changes in claim rate, which would also leads to changes in claim sizes. At the same time, changes in premiums due to experience rating changes might affect the demand for voluntary collision coverage. This would lead to changes in compositions of risks of drivers purchasing voluntary collision coverage, complicating how we estimate claim rate changes. So in the section, we start by showing demand responses to premium changes. As a result of concerns of changes in compositions of risks, we use drivers who always buy voluntary collision coverage to estimate clean incentive effects on claim rate and size, and we further show both ex-ante and ex-post moral hazard contribute to changes in claim rate.

3.3.1 Demand Responses and Composition Changes

As premiums of voluntary collision coverage increased for higher risk class drivers and decreased for lower risk class ones, consumers might respond to the premium changes and switch into or out of the coverage. If drivers in different risk classes have different risk profiles, the compositions of drivers buying voluntary collision coverage might change.

We examine the demand response of voluntary collision coverage by plotting time series of share of drivers buying voluntary collision coverage and estimating Equation 3, with outcome variable as whether drivers purchase voluntary collision coverage. Appendix Figure 11 and Appendix Figure 12 displays the time series plots and TWFE estimates by risk classes. Drivers of higher risk classes switched out of coverage and those of lower risks switched in, despite of smaller effects. The selection pattern implies that, among drivers purchasing collision coverage, there are fewer drivers from higher risk classes and more from lower risk classes. So when we examine the claim rate trends using all voluntary collision policies, it would include the incentive effects of experience rating factors as well as the composition changes, i.e. fewer consumers from higher risk classes and more from lower risk classes.[15]

3.3.2 Changes in Claim Rate and Claim Size

We examine changes in claim rate and size of voluntary collision coverage. As we discussed the potential concerns of composition changes above, to isolate incentive effects of experience rating changes, we restrict to our sample to drivers who always purchase voluntary collision coverage (always buyers) during our sample period. We start with presenting the time series plots of claim rates and sizes in Figure 5. The claim rates in three reform groups are parallel and slightly trend down, with the level in the third reform group is slightly higher. After the announcement, claim rates drop quickly by about 2 claims per 100 policies in all three regions, which is much larger than the decrease in claim rate for liability coverage. At the same time, claim size increase from 2,200 yuan to 3,300 yuan after reform. The drop in claim size and increase in size reflect that drivers might be less likely to report the small-dollar losses after reform. This is also verified

[15] Another interesting point arising from selection is whether the drivers who choose to select into or out of voluntary coverage, within each risk class, are of higher or lower risk, i.e. advantageous/adverse selection. To identify risk selection effects, we need additional exogenous variations in premiums (Finkelstein et al., 2010). However, we lack of such variations in the paper and abstract the selection within risk class from current paper.

in Figure 7a, which plots the histogram of claim size of voluntary collision coverage before and after the reform. In contrast, histogram of claim size of voluntary liability coverage, as shown in Figure 7b, doesn't change much. The increase in claim rates and the decrease in claim sizes together indicate potential ex-post claiming behavior.

We estimate Equation 3 for voluntary collision coverage among always buyers. Figure 6 displays results. The estimates of collision claim rates remained stable around zero before the reform announcement. The claim rate drop modestly during the initial three months after the reform announcement, followed by a more substantial decrease of 1-2 claims per 100 policies. Furthermore, alongside the claim rate changes, the reform increased claim size by approximately 1,000 yuan. Column (2) in Table 2 reports the average effects for voluntary collision coverage. On average, the claim rate decrease by 1.5 per 100 policies (52% of pre-reform mean), and the size increased by about 600 yuan. Overall, the combined evidence of the decline in claim rate and the rise in claim size suggest that drivers might be less likely to report the small-dollar losses, that is, reductions in ex-post moral hazard. In addition, we estimate the model use all buyers. Figure 13 shows that the results are almost the same as above.

However, in addition to reductions in reporting small-sized losses, reductions in risky driving could also contribute to reductions in the claim rate for collision damage coverage. To show the role of reductions in risky driving, we categorize claims for voluntary collision coverage into four groups by their claim sizes. The four groups are divided by the quartiles of claim sizes: the first group ranges from 0 to 790 yuan, the second group from 791 to 1,275 yuan, the third group from 1,256 to 2,550 yuan, and the fourth group above 2,250 yuan. For each group, we calculate the monthly claim rate as number of claims per 100 voluntary collision policies. Figure 8 displays the results by claim sizes. The effects are smaller for larger size claims, with the claim rate of the smallest size dropping by 0.5 to 1.5 p.p., while the claim rate for the largest dropping by 0.2 to 0.4 p.p., which is consistent with drops in the compulsory coverage. The results suggest that reductions in large-sized claims are primarily from reductions in ex-ante moral hazard behavior, while reductions in smaller-sized claims combines both sources from ex-ante and ex-post moral hazard. Assuming a constant semi-elasticity of risky driving with respect to experience rating under liability coverage and collision coverage, reductions in ex-post claiming account for approximately 46% of the decrease in collision claim rates.[16]

3.4 Summary of Empirical Findings

Overall we find that the reform reduced claim rates. We estimate a size reduction in risky driving from the increase in claims-related insurance penalties, implying reductions in ex-ante moral hazard. Simultaneously, we observe a much larger decreases in collision-related claim rate. The marked reductions in claim rate, along with increases in claim sizes for collision coverage, imply the reduction in ex-post claiming moral hazard, as the largest effect seems to be a reduction in claims for smaller dollar amount collision losses. However, it is important to recognize that this also incorporates a reduction in ex-ante moral hazard, as illustrated by the reduction in frequency of large-sized claims. Therefore, to fully understand the effects of increased in experience rating, we need to consider the different implications of the two separate sources of reductions.

4 A Model of Insurance With Experience Rating

We develop a conceptual model to help illuminate the key tradeoffs and welfare implications involved in the use of experience rating. We follow the model developed by Abbring et al. (2008) and take account into ex-ante and ex-post moral hazard.

We start with a static model with a single period for simplicity. An agent, defined as the insured driver and denoted by i, faces a probability of an accident/loss $\overline{p}$ but can lower this probability by exerting some costly effort. We model this as a choice of the probability of loss p from an interval $[\underline{p}, \overline{p}]$ with convex costs $\Gamma(p)$, with $\Gamma'(p) < 0$ and $\Gamma''(p) > 0$, and $\Gamma(\overline{p}) = 0$.

In the event of a loss with probability p, the loss size l is random and follows a CDF $F(l)$ and a corresponding PDF $f(l)$. We assume that the loss size is independent of the loss probability for simplicity.[17]

[16] The claim rate for collision coverage is given by $\Pr(\text{claim}) = \Pr(\text{claim}|\text{accident})\Pr(\text{accident})$. The semi-elasticity of the collision coverage claim rate with respect to changes in experience rating is defined as $\frac{d\Pr(\text{claim})/d\theta}{\Pr(\text{claim})} = \frac{d\Pr(\text{claim}|\text{accident})/d\theta}{\Pr(\text{claim}|\text{accident})} + \frac{d\Pr(\text{accident}/d\theta)}{\Pr(\text{accident})}$. Assuming a constant semi-elasticity of risky driving with respect to experience rating under both liability and collision coverage, i.e., $\frac{d\Pr(\text{accident})/d\theta}{\Pr(\text{accident})}$ at 28% as estimated under compulsory liability coverage, and given an estimated $\frac{d\Pr(\text{claim})/d\theta}{\Pr(\text{claim})}$ of 52%, we calculate the semi-elasticity of ex-post claiming at 24%, accounting for a 46% reduction in the total claim rate of collision coverage.

[17] This is motivated by the empirical evidence that sizes of compulsory liability coverage remained unchanged when risky driving decreased.

Conditional on loss size l, the agent decides whether to file the claim with the insurer.[18] If the agent chooses not to file the loss, she bears the loss herself. If the agent files the claim, the insurer covers the whole claimed losses (full insurance).[19] Following Abbring et al. (2008), we model an agent's decision about whether to file a claim as a choice of claim threshold x: when $l < x$, the agent doesn't claim; otherwise, the agent files the claim. We start with the simple assumption that there are no additional frictions or costs associated with filing claims and/or claiming process. The only cost to the agent for not filing is the realized loss.

The dynamics of consumption are as follows. We assume that each period the agent receives a constant amount of income, denoted as w. With a claim filed, the premium increases by θq, resulting in a total premium of $(1+\theta)q$, where θ is defined as the experience rating factor or experience rating penalty intensity. Should there be no claim regardless of no accident or no reporting, the agent pays base premium q.

The model is a single period model with two stages. The timing of the model is as follows. In the first stage, a risk-averse expected utility-maximizing agent anticipates the distribution of realized losses and chooses the optimal probability of accident, which depends on the plan coverage as well as experience rating. In the second stage, conditional on the realization of losses, the agent chooses to claim or not. Payoffs are realized at the end of the period with realized premiums and any incurred unclaimed loss. We model that the agent chooses the probability of accident and the claim threshold simultaneously because the agent has no incentives to deviate from the simultaneous choice when the loss is realized.

4.1 The Static Model

The agent's problem can be written as:

$$V(\theta) = \max_{p,\, x} \underbrace{p \left(\overbrace{\int_0^x u(w-l-q)dF(l)}^{\text{not claim}} + \overbrace{\int_x^\infty u(w-(1+\theta)q)dF(l)}^{\text{claim}} \right)}_{\text{utility w/ an accident}} + \underbrace{(1-p)u(w-q)}_{\text{utility w/o an accident}} - \underbrace{\Gamma(p)}_{\text{cost}}. \quad (5)$$

[18] In our model, the size of losses is assumed to be exogeneously determined. Consequently, ex-post moral hazard is limited to the decision of whether to file a claim following a loss. This excludes the potential issues related to the intensive margin, that is, the insured would increase the expenditures beyond the optimal level as commonly discussed in health insurance. This seems to be less of an issue in our model because we always assume full insurance and changes in experience rating do not have direct effects on intensive margin.

[19] In practice, an agent might pay coinsurance. But we abstract it away for simplicity.

Effects of increasing experience rating penalty on claim filing and the probability of loss. By the first order condition w.r.t. x and p,[20] the optimal claim threshold is $x^* = \theta q$ and the optimal level of probability of losses equates the marginal cost and the marginal benefit:

$$-\Gamma'(p^*) = u(w-q) - \left[\int_0^{x^*} u(w-l-q)dF(l) + \int_{x^*}^{\infty} u(w-(1+\theta)q)dF(l)\right]. \tag{6}$$

Obviously, a higher penalty intensity θ leads to a higher threshold. This is straightforward: if loss size l is smaller than penalty θq, the insured benefits by saving $\theta q - l$ through not filing the loss. Conversely, if the loss size l is larger than penalty θq, the insured benefits from claiming. Also, the optimal p^* decreases with the penalty intensity θ because $\frac{dp^*(\theta)}{d\theta} = \frac{-u'(w-(1+\theta)q)q(1-F(x^*))}{\Gamma''(p^*(\theta))} < 0$.

4.1.1 Welfare Effects of Increasing Experience Rating Penalty

To understand how consumer welfare changes when experience rating penalties increases, we take the derivative of $V(\theta)$ w.r.t. θ:

$$\frac{dV(\theta)}{d\theta} = \underbrace{-u'(w-(1+\theta)q)p^*(\theta)(1-F(x^*))q}_{<0;\uparrow\text{penalty with a claim}} + \tag{7}$$

$$\underbrace{\left(p^*(\theta)\int_0^x u'(w-l-q)dF(l) + p^*(\theta)(1-F(x^*))u'(w-(1+\theta)q)(1+\theta) + (1-p^*(\theta))u'(w-q)\right)(-\frac{dq}{d\theta})}_{>0;\uparrow\text{benefit from base premium}\downarrow}. \tag{8}$$

This directly demonstrates that, as θ increases, changes in consumer welfare depend on the tradeoffs between cost of additional penalties and potential reduction in equilibrium premiums. We replace $u'(w-l-q)$ and $u'(w-(1+\theta)q)$ with their first order Taylor approximations at consumption level, $w-q$, in no accident

[20] See Appendix Section D for detailed mathematical derivation.

state into Equation 7, rearrange the equation, and normalize by $u'(w-q(\theta))$:

$$\frac{dV(\theta)}{d\theta}/u'(w-q) = \tag{9}$$

$$\underbrace{\left(-\overbrace{\frac{dq(\theta)}{d\theta}}^{\text{premium}\downarrow}\left(1+p^*(\theta)(1-F(x^*)\theta)\right)-\overbrace{p^*(\theta)(1-F(x^*))q(\theta)}^{\text{expected penalty}\uparrow}\right)}_{\text{expected consumption level}\uparrow} \tag{10}$$

$$-\underbrace{\left(-\frac{u''(\cdot)}{u'(\cdot)}\right)}_{\text{abs. risk aversion}}\underbrace{\left(\theta q(\theta)\overbrace{p^*(\theta)(1-F(x^*))q(\theta)}^{\text{expected penalty}\uparrow}+\overbrace{\frac{dq(\theta)}{d\theta}}^{\text{premium}\downarrow}\left(p^*(\theta)\int_0^{x^*} l dF(l)+\theta q(\theta)p^*(\theta)(1-F(x^*))(1+\theta)\right)\right)}_{\text{consumption smoothing}\downarrow}, \tag{11}$$

which highlights the key tradeoff of increasing experience rating penalties. The first line shows potential benefits from increases in expected consumption level, resulting from reduced premiums and increased penalties, whereas the second line highlights the reduced consumption smoothing resulting from higher penalties, offset by the reduced equilibrium base premiums in the state of a loss. The marginal utility in the loss state is higher as consumption is lower at $w-q-l$ or $w-(1+\theta)q$, compared to $w-q$ in the state of no losses. With an increase in θ, the penalty in the loss state increases, resulting in further consumption reductions in this adverse state and reduced consumption smoothing. However, the decrease in premiums, $\frac{dq(\theta)}{d\theta}$, would soften the increase in penalty in the loss state.

Externality. Our model highlights the tradeoff faced by the insureds. One significant missing component is the externalities to third parties from traffic accidents (Edlin and Karaca-Mandic, 2006). Externalities arise in the context of ex-ante moral hazard: More cautious driving leads to reduced accidents impacting third parties. Therefore, to account for externality cost, we could think of an extra term in Equation 9 as reductions in risky driving, multiplied by average cost of such losses.

4.1.2 Premium and Welfare Implications Under the Competitive Market Assumption

As noted that the welfare effect depends on the extent of premium changes $\frac{dq(\theta)}{d\theta}$. To understand how increasing experience rating penalties affects base premium levels, we assume a competitive insurance market.

Then we will come back to the welfare implications.

Effects of increasing experience rating penalty on the base premium. In a competitive insurance market, the equilibrium base premium q is determined by equating expected premium revenue and expected cost:

$$\underbrace{p^* \int_{x^*}^{\infty} l dF(l)}_{\text{expected cost}} = \underbrace{\overbrace{p^*(1-F(x^*))(1+\theta)q}^{\text{claim in an accident}} + \overbrace{p^* F(x^*) q}^{\text{not claim in an accident}} + \overbrace{(1-p^*)q}^{\text{no accident}}}_{\text{expected premium}}, \tag{12}$$

and thus we have $q = \frac{p^* \int_{x^*}^{\infty} l dF(l)}{1+p^*(1-F(x^*)\theta)}$. We take a total derivative of Equation 12 w.r.t. θ and rearrange all terms except for $\frac{dq(\theta)}{d\theta}$ on the RHS:

$$\frac{dq(\theta)}{d\theta}\left(1+p^*(1-F(x^*)\theta)\right) = \tag{13}$$

$$\underbrace{\frac{dp^*(\theta)}{d\theta}\int_{x^*}^{\infty} l dF(l)}_{<0;\downarrow \text{ accident rate}} - \underbrace{\frac{dp^*(\theta)}{d\theta}(1-F(x^*))\theta q(\theta)}_{<0;\downarrow \text{ penalty revenues}} \tag{14}$$

$$+ \underbrace{p^*(\theta)\frac{\partial}{\partial\theta}\int_{x^*}^{\infty} l dF(l)}_{<0;\downarrow \text{ claim filings}} \quad - \underbrace{p^*(\theta)\frac{\partial}{\partial\theta}\left(1-F(x^*)\right)\theta q(\theta)}_{<0;\downarrow \text{ penalty revenues}} \tag{15}$$

$$\underbrace{-p^*(1-F(x^*))q}_{<0;\text{ direct effect of }\uparrow\theta}. \tag{16}$$

The changes in base premiums can be decomposed into three channels: changes in ex-ante moral hazard; changes in ex-post moral hazard; and mechanical changes. The first line show the effects through changes in ex-ante moral hazard, $\frac{dp^*(\theta)}{d\theta}$. The base premium decreases due to a lower accident rate and subsequently a lower average cost, which is somewhat offset by lower penalty revenues. Typically the average claim size $\int_{x^*}^{\infty} l dF(l)$ is larger than the penalty size $\theta q(\theta)$. So reductions in ex-ante moral hazard generally lead to lower premiums. Second, changes through x^* or $F(x^*)$, i.e. ex-post moral hazard, are shown on the second line. Fewer claiming, due to a higher claim threshold, decreases the base premiums by decreasing expected cost. However, as fewer drivers claim, the revenues from penalties decreases, and thus the base premium increase. With the optimal claim threshold set at $x^* = \theta q$ under the assumption of no additional claim cost, the two terms on the second line cancel out. Intuitively, with a marginal change in θ, reductions in cost θq due to fewer claims are exactly from individuals who avoid claims and incur the penaltyθq on the margin. Therefore, changes in ex-post moral hazard have no effects on the base premiums. The last term represents the direct effects of higher experience rating factor: higher penalties directly lower the base premium. To summarize, the effects of increasing experience rating on base premiums are only through reductions in ex-ante moral hazard and direct changes, while changes in ex-post moral hazard have no effects on the premium changes, under the assumption of no costs associated with filing agents.

Implications on welfare and optimal experience rating penalty. Under the assumption of competitive markets and no additional cost associated with claiming, the expected consumption level only includes the ex-ante moral hazard part: $\frac{dp^*(\theta)}{d\theta}\left(\int_{x^*}^{\infty} l dF(l) - (1-F(x^*))\theta q(\theta)\right)$. As we show before, the channel through ex-post moral hazard cancel out, as any reductions in the cost from the decision not to file claims for some losses are fully offset by the reductions in penalty/premium revenues. Additionally, the direct effects of increasing experience rating factor on premiums cancel out the direct effects of increasing experience rating factor on expected consumption, since any premium reductions from larger penalties are fully offset by the individuals' need to absorb those losses themselves in the case of an accident. Thus, the welfare effects only depend on the extent of behavioral responses of driving: a larger reduction in the accident probability due to increased penalty intensity leads to larger decreases in base premiums and higher consumer welfare. In an extreme case where there is no behavioral response of ex-ante moral hazard, consumer welfare decreases unambiguously when the penalty intensity θ increases. Intuitively, the consumer welfare decreases because experience rating further decreases consumption in the state where marginal utility is highest (due to the penalty) and "equally" divided into all states through base premiums. Therefore, the optimal θ would be zero in such case. In the other extreme case where the behavioral response of decreasing the probability of loss is large enough, the benefits from lower premiums on the expected consumption level are very large. Moreover, the large premium reductions would decrease or even overturn the reduced consumption smoothing caused by the penalty. Thus, increasing experience rating penalties is unambiguously welfare improving, leading to a corner solution of maximum θ. With moderate responses to θ, there should be an optimal θ that balances the marginal benefit of premium saving and consumption risks. In summary, the impact of increasing experience rating penalties on consumer welfare depends on the extent of reduction in ex-ante moral hazard.

4.2 Discussions of the Model

4.2.1 Costs Associated With Claiming

So far we assume no additional cost associated with claiming. We relax the assumption to allow for additional costs associated with claiming or processing insurance claims. For simplicity, we assume a (fixed) positive cost associated with filing a claim for drivers. As a result, the optimal threshold is $x^* = \theta q + c$, which is higher than the previous threshold by the amount of c. With this assumption, an increase in θ leads to the following changes in the equilibrium premiums:

$$\frac{dq(\theta)}{d\theta} = \frac{\frac{dp^*(\theta)}{d\theta}\left(\int_{x^*}^{\infty} l dF(l) - (1-F(x^*))\theta q\right) - p^* c f(x^*)\frac{\partial x^*}{\partial \theta} - p^*(1-F(x^*))q}{1+\theta p^*(1-F(x^*))}. \tag{17}$$

Changes in ex-post moral hazard on premiums are reflected through the term $p^* c f(x^*)\frac{\partial x^*}{\partial \theta}$, where a larger response of ex-post moral hazard, i.e. larger $f(x^*)\frac{\partial x^*}{\partial \theta} = \frac{\partial F(x^*)}{\partial \theta}$, result in greater reductions in equilibrium base premiums. There are multiple sources of additional costs involving with claiming, such as administrative costs incurred by insurers to deal with claims. In this case fewer claims lead to less administrative costs to deal with claims and further to lower premiums.

With positive costs associated with claiming, the expected consumption level increases by $p^* c f(x^*)\frac{\partial x^*}{\partial \theta}$. The size of increased consumption level depends on the claim cost c as well as the response of ex-post moral hazard $f(x^*)\frac{\partial x^*}{\partial \theta}$. However, different sources of claiming costs have different implications on consumption smoothing. If additional costs are from driver's hassle or time cost of dealing with the claims, they need to incur such cost in the state of loss and claiming, which further reduced consumption smoothing. However, if the additional costs are from administrative cost, it will not show up in the consumption directly.

What causes the difference in welfare implications with or without costs associated claiming? The answer is whether there are any "real" reductions in cost. If there are no costs associated with claiming, the increase in θ only leads to transfer from consumers to the insurers (which further transfer to consumers via decreased base premiums but spread in all states in competitive markets). In contrast, when there are costs associated claiming, the reductions in premiums are further from external sources of such additional cost.

4.2.2 Experience Rating vs Deductible

So far we assume the model is static with a single period. Although it works well to capture the tradeoff of experience rating, the mechanism works the same way as deductible, another instrument commonly used in insurance. To show how experience rating works differently from deductible, we extend the model to multi-period. For simplicity and tractability, we assume that the agent chooses the probability of accident and applies it to all periods. We present the details of the model in Appendix Section C. We obtain two primary conclusions from the multi-period model. First, experience rating and deductible are complements to each other. We show that consumers are always better off with small positive experience rating, regardless of the level of deductible. The intuition is that the two instruments extract penalties in different period-state, which reduced the adverse effects of reduced consumption smoothing. An increase in deductible leads to lower consumption in the loss states as of the accident period. A further increase in deductible continues to lower consumption in the same period-state. However, increasing experience rating penalties lowers the consumption in the next period, in which includes an accident state and a no-accident state with relatively lower marginal utility. Therefore, experience rating is a complement to deductible. The value added by experience rating is because it allows penalties or lower consumption from the lower marginal state, thus leading to less reduced consumption smoothing. Second, our multi-period model indicates that the static model is a good approximation for per period utility when the number of periods is large. In a multi-period model, we always start with a good state in which no accident occurs and consumers only pay base premiums, and followed by periods with possibility of good and bad states. As the number of periods increases, the effects of first period with no losses or penalties diminish.

4.3 Summary of Theoretical Model

To summarize, the model hights the tradeoff between benefits from decreased equilibrium premiums and losses of risk protection from reduced consumption smoothing. The extent of welfare changes due to increasing penalties depends on the degree of changes in risky driving behavior, which determines the changes in the base premiums . When the cost of modifying driving behaviors is low, the accident rate becomes more responsive to penalties, leading to a larger decrease in base premium decreases, and higher consumer welfare. We find that, under the assumption that claiming incurs no additional costs or benefits except for the loss itself, increasing experience rating penalties have zero impacts on the base premiums through claiming behavior because the decreased cost due to higher claim threshold is fully offset by lower claim rate on the margin. When considering the costs associated with filing a claim, lower claiming rate also reduced premiums through reducing claiming cost.

5 Quantifying Welfare Effects of Increasing Experience Rating Factor - A Sufficient Statistic Approach

We use the sufficient statistic approach to evaluate the welfare effects of experience rating factor changes. Equation 9 shows how consumer surplus (CS) changes when θ changes. Considering that consumers in different risk classes face different θ, we rewrite Equation 9 with subscript r. To obtain the CS of the whole population, we first quantify the CS for each risk class, and then compute weighted average CS for the whole population. The sufficient statistic formula of risk class r is:

$$CS_r = \frac{dV(\theta_r)}{d\theta_r} / u'(w - q_r) \tag{18}$$

$$= \left(- \overbrace{\left(\frac{dq_r(\theta_r)}{d\theta_r}\right)}^{\text{prm changes}} \left(1 + \underbrace{p_r{}^*(1 - F(x_r^*))}_{\text{claim rate}} \overbrace{\theta_r}^{\text{ER factor}}\right) - \underbrace{p_r^*(1 - F(x_r^*))}_{\text{claim rate}} \overbrace{q_r}^{\text{premium}} \right) \tag{19}$$

$$- \underbrace{\left(-\frac{u''(\cdot)}{u'(\cdot)}\right)}_{\text{coeff. of abs. risk aversion}} \left(\underbrace{p_r^*(1 - F(x_r^*))}_{\text{claim rate}} \theta_r q_r^2 + \overbrace{\left(\frac{dq_r(\theta_r)}{d\theta_r}\right)}^{\text{prm changes}} \left(\underbrace{p_r^* \int_0^{x_r^*} l dF(l)}_{\text{unclaimed loss}} + \underbrace{p_r^*(1 - F(x_r^*))}_{\text{claim rate}} (1 + \theta_r)\theta_r q_r \right) \right). \tag{20}$$

There are three types of terms in Equation 18 to estimate CS_r.

Observed terms. Several terms in Equation 18 can be directly observed in the data. First, $p_r(1 - F(x_r))$ is the average claim rate of risk class r, which is measured before reform. Second, the base premium and experience rating factor can also be observed in our data.[21]

[21] The base premium in the data is defined as the premium with one claim in the past year, and premium with zero claims in the past year is a percentage of base premium. We re-normalize the base premium in the data to define the base premium as the premium with no claims in the past year, denoted as q_r, to match our model. Similarly we transform the discount factor in the data to experience rating factor θ_r. See Appendix E for detailed steps.

Unobserved terms. Two terms cannot observed in our data: average unclaimed loss size and constant absolute risk aversion (CARA) coefficient. First, $p_r^*(\theta)\int_0^{x_r^*} l dF(l)$ measures the average unclaimed cost, which is the total unclaimed losses over the number of policies. We use the one fourth of average (claimed) cost as a baseline measure to approximate the unobserved value. Moreover, we provides a bound for this value: the lower bound of zero and the upper bound corresponding the average cost of claims (including zeros for no claims). This bound is very conservative, but can still provide us informative estimates of welfare changes. Second, $\left(-\frac{u''(c)}{u'(c)}\right)$ represents CARA coefficient. Currently, we calibrate it from existing literature. Our benchmark calibration use a value of 0.0067 from Cohen and Einav (2007), divided by the average exchange rate, 6.4 CNY/USD, during the sample period to convert to inverse US dollar unit. In addition, we conduct robustness checks of how our welfare estimates change with a wide range of CARA coefficients.

Estimable term. We quantify $\frac{dq(\theta)}{d\theta}$ under the assumption of competitive market. Equation 13 shows how premiums change under the assumption of competitive market, which requires to separate out changes in ex-ante and ex-post moral hazard. It is worthwhile noting that we can rewrite the Equation 13 by combining derivatives of expected cost and of penalty income as

$$\frac{dq(\theta)}{d\theta} = \frac{1}{1+p^*(1-F(x^*))\theta}\left(\underbrace{\frac{d}{d\theta}\left[p^*\int_{x^*}^{\infty} l dF(l)\right]}_{\text{changes in avg cost}} \underbrace{-\frac{d}{d\theta}\left[p^*(1-F(x^*))\right]}_{\text{changes in claim rate}}\theta q(\theta)\underbrace{-p^*(1-F(x^*))q}_{\text{direct changes}}\right). \tag{21}$$

This formula simplifies Equation 13 and only consists of changes in the average cost (including claims with zero values) and changes in the claim rate, as well as direct changes, where the first two terms can be estimated from the data. Though ex-ante and ex-post moral hazard do not directly show up in the equation, relative changes in the first two terms reflect the relative contribution of changes in ex-ante and ex-post moral hazard. If all the changes in the claim rate are from ex-post moral hazard, the firm two terms will cancel out. However, if reductions in ex-ante moral hazard contribute to more to the reductions in claim rate, the base premium will experience a larger decrease.

To obtain changes in base premiums, we first estimate changes in average costs and claim rates and obtain how premium changes for each risk class separately.[22] Second, we pool claims of all risk classes, as insurers do in practice, to determine how the base premium changes. The premium changes due to pooling are a weighted average of changes across each risk class, with the weight being the claim rate's proportion attributable to each risk class.

5.1 Welfare Results.

Table 3 presents the welfare changes along with the intermediary steps required to calculate these changes. First, we present some summary statistics information for each risk class. About 57% of insureds have no claim in the past one to three years. Experience rating factors increase by 7 p.p. for highest risk class, but increase by 24 p.p. for lowest risk class due to the reform. Second, we report the pre-reform claim rate and average costs and effects of the policy change on claim rate and average cost. Annual claim rate for voluntary coverage is as high as 469 claims per 1,000 policies for the highest risk class, and is 122 claims per 1,000 policies for the lowest risk class. Additionally, the average cost is from more than 1,700 yuan for the highest risk class to about 500 yuan for the lowest risk class. The overall effects of the reform on claim rate and average cost decrease by risk class, with annual claim rate decreasing by more than 100 claims per 1,000 policies p.p. for higher risk classes and by 40-80 claims per 1,000 policies for lower risk classes. The effects on average cost ranges from 550 yuan to 80 yuan for higher and lower risk classes.

The bottom two panels report the baseline and bounded welfare changes. Implied premiums changes under the competitive insurance markets range from -78 yuan for higher risk classes. to more than 226 yuan for lower risk classes. Welfare gains from decreased moral hazard and increased expected consumption are approximately 7-183 yuan across risk classes, with larger amount in higher risk classes. The value of reduction in risk protection amounts from negative around 20 yuan, implying the premium decreases dominate the penalty effects and leading to an increase in risk protection. Overall, welfare increases by approximately 30 yuan annually per insured driver for higher risk categories and by about 200 yuan for lower risk categories, leading to an increase in welfare by 126 yuan ($20) per insured driver per year. We additionally bound the welfare changes arising from reductions in risk protection, where the lower bound shows smaller effects while the upper bound presents larger effects.

[22] As mentioned before, base premium in the data is defined slightly different from that in our model. To obtain how base premium q_r changes with θ_r, we first estimate how premiums with one claim change with decrease in discount factor. Then we re-normalize to obtain $\frac{dq_r}{d\theta_r}$. See Appendix E for detailed steps.

Externality. The welfare analysis has so far focused on voluntary coverage as a self-contained system, accounting for adjustments in the experience rating, claims, and premiums. As demonstrated by Edlin and Karaca-Mandic (2006), the addition of a vehicle or an extra mile driven imposes additional costs on third parties by increasing the likelihood of involving other drivers in accidents. Such externality costs would be larger when the additional mileage involves greater risk. Therefore, reductions in risky driving behaviors decrease the likelihood of causing damage to third parties. However, the external costs of risky driving to third parties are typically reflected in liability insurance premiums. While the analysis of voluntary coverage incorporates some liability costs, the bulk of the costs from compulsory liability coverage remains unaccounted for. To quantify the externality costs associated with compulsory coverage, we rely on the reduction in the compulsory liability claim rate, multiplied by the average claim size. Our estimates show a reduction of 25 claims per 1,000 policies in the annual compulsory coverage claim rate, with an average claim cost of approximately 3,482 yuan, yielding an annual reduction in externality costs of 88 yuan ($14) per policy.

Moreover, the liability insurance costs in our setting likely underestimate the full extent of externality costs, in extreme scenarios such as fatalities, which are often insufficiently covered. To address this extreme case, we apply the best available estimates of reductions in risky driving, i.e. reductions in the compulsory coverage claim rate, combined with the per-accident fatality rate (0.15 from Wang et al. (2019)) and the statistical value of life in China (4 million yuan). The reduction in fatality-related costs could reach as much as 1,500 yuan, calculated as .025 times .015 times 4 million yuan.

CARA coefficient sensitivity. In our analysis, we have utilized a CARA coefficient of .0067/6.4, which is considered large relative to existing literature. To provide robust welfare change estimates, we explore a wide range of CARA coefficients, from 0.0001 to .001 (divided by 6.4), including various CARA estimates from the literature. Appendix Figure 14 presents the welfare estimates with different CARA coefficients. Our results show that welfare changes are robust across the wide range of CARA coefficients.

Summary. This section evaluates the welfare effects of the policy changes that increase experience rating using a sufficient statistic framework. We find that welfare improves by 214 yuan per insured driver per year per year, where 59% is from the increase in expected consumption level captured by voluntary coverage, and

41% is from the reduced externality cost captured by compulsory system. Our results are generally robust to CARA coefficients. Additionally, we explores the externalities associated with compulsory coverage and risky driving, noting how these factors influence overall welfare.

6 Discussion

In this paper, we empirically and theoretically evaluate the behavioral and welfare implications of increasing experience rating. We empirically estimate the effects of experience rating in auto insurance markets. Leveraging administrative data for a 5% sample of automobile insurance policies in China over a 4-year period, totaling more than 7 million policies, we exploit the staggered rollout of this policy change across the nation and employ a combination of approaches, including recent developments in DID approaches. To estimate the impact on risky driving behavior, we focus on claims for liability for damages to third parties, over which the insured have little discretion about claiming. We find a decrease in monthly claim rate by 28% (2 claims per 1,000 policies). These results show that increasing experience rating penalties had a meaningful impact on reducing accident risk. We also estimate the impact on the claim rate for collision damage coverage, which likely reflects a combination of ex-ante and ex-post moral hazard since insureds have more discretion about making these claims. The monthly claim rate for collision damages decreases by 52% after the increase in experience rating (15 claims per 1000 policies). The average claim size increases by about 690 yuan and we observe a large reduction in claims for small-dollar collision losses, consistent with ex-post moral hazard effects.

Next, we develop a model of insurance with experience rating. We show the downside of increased experience rating penalties is that it reduces risk protection. The loss of risk protection can be offset with the benefits of reduced moral hazard. Our model, though, highlights that it is important to separate ex-ante moral hazard) versus ex-post moral hazard. Reductions in ex-ante moral hazard have positive welfare effects. So if an increase in experience rating penalties reduces accidents by a sufficient amount to offset the losses in risk protection, it can be welfare improving. In contrast, the welfare implications of ex-post moral hazard are ambiguous. If there are no additional frictions or costs associated with filing claims, experience rating penalties are welfare reducing regardless of the impact of ex-post moral hazard. This is because the marginal increase in experience rating penalties in the presence of ex-post moral hazard only increases consumption risk, and any reductions in premiums that arise from the decision not to file claims for some losses are fully offset by the individuals' need to absorb those losses themselves. Ex-post moral hazard effects can be welfare improving, though, if there are frictions or additional costs associated with filing or processing insurance claims. In this case, ex-post moral hazard responses have similar impacts as ex-ante moral hazard. We further quantify the welfare implications using a sufficient statistic formula derived from the model, combined with empirical estimates. We find the welfare of the policy change increases by $30 per insured driver per year.

This paper has focused on the incentive effects of experience rating, specifically examining how increasing experience rating impacts both risky driving behaviors and ex-post claiming, and their corresponding welfare effects. However, our paper abstracts from the role of experience rating in risk selection and classification, which could be explored in future research.

References

Kibrom A Abay. How effective are non-monetary instruments for safe driving? panel data evidence on the effect of the demerit point system in denmark. *The Scandinavian Journal of Economics*, 120(3):894–924, 2018.

Jaap H Abbring, Pierre-André Chiappori, and Jean Pinquet. Moral hazard and dynamic insurance data. *Journal of the European Economic Association*, 1(4):767–820, 2003a.

Jaap H Abbring, James J Heckman, Pierre-André Chiappori, and Jean Pinquet. Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association*, 1(2-3):512–521, 2003b.

Jaap H Abbring, Pierre-André Chiappori, and Tibor Zavadil. Better safe than sorry? ex ante and ex post moral hazard in dynamic insurance data. 2008.

Lawrence Blincoe, Ted R Miller, Jing-Shiarn Wang, David Swedler, Tom Coughlin, Bruce Lawrence, Feng Guo, Sheila Klauer, and Thomas Dingus. The economic and societal impact of motor vehicle crashes, 2019. Technical report, 2022.

P Bolton and M Dewatripont. , contract theory (cambridge: Mit press). 2005.

Michael Braun, Peter S Fader, Eric T Bradlow, and Howard Kunreuther. Modeling the "pseudodeductible" in insurance claims decisions. *Management Science*, 52(8):1258–1272, 2006.

R Andrew Butters, Daniel W Sacks, and Boyoung Seo. How do national firms respond to local cost shocks? *American Economic Review*, 112(5):1737–1772, 2022.

Brantly Callaway. Difference-in-differences for policy evaluation. *arXiv preprint arXiv:2203.15646*, 2022.

Brantly Callaway and Pedro HC Sant'Anna. Difference-in-differences with multiple time periods. *Journal of econometrics*, 225 (2):200–230, 2021.

Christopher S Carpenter and Mark Stehr. The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and crash-related injuries among youths. *Journal of health economics*, 27(3):642–662, 2008.

Raj Chetty. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(Volume 1, 2009):451–488, 2009. ISSN 1941-1391.

Alma Cohen and Rajeev Dehejia. The effect of automobile insurance and accident liability laws on traffic fatalities. *The journal of law and economics*, 47(2):357–393, 2004.

Alma Cohen and Liran Einav. Estimating risk preferences from deductible choice. *American economic review*, 97(3):745–788, 2007.

Stuart V Craig. Competition in employer sponsored health insurance: Implications for a public option. Technical report, Working paper, 2022.

David M Cutler and Richard J Zeckhauser. The anatomy of health insurance. In *Handbook of health economics*, volume 1, pages 563–643. Elsevier, 2000.

Claire Daley, James Gubb, E Clarke, and E Bidgood. Healthcare systems: Switzerland, 2012.

Manasi Deshpande and Yue Li. Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–248, 2019.

Georges Dionne and Ying Liu. Effects of insurance incentives on road safety: Evidence from a natural experiment in china. *The Scandinavian Journal of Economics*, 123(2):453–477, 2021.

Georges Dionne, 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*, 11(4):897–917, 2013.

Aaron S Edlin and Pinar Karaca-Mandic. The accident externality from driving. *Journal of Political Economy*, 114(5):931–955, 2006.

Isaac Ehrlich and Gary S Becker. Market insurance, self-insurance, and self-protection. *Journal of political Economy*, 80(4): 623–648, 1972.

Liran Einav and Amy Finkelstein. Moral hazard in health insurance: what we know and how we know it. *Journal of the European Economic Association*, 16(4):957–982, 2018.

Liran Einav, Amy Finkelstein, and Paul Schrimpf. The response of drug expenditure to nonlinear contract design: Evidence from medicare part d. *The quarterly journal of economics*, 130(2):841–899, 2015.

Liran Einav, Amy Finkelstein, and Paul Schrimpf. Bunching at the kink: implications for spending responses to health insurance contracts. *Journal of Public Economics*, 146:27–40, 2017.

Hanming Fang and Xian Xu. Chinese insurance markets: Developments and prospects. Technical report, National Bureau of Economic Research, 2023.

Amy Finkelstein, Liran Einav, and Mark R Cullen. Estimating welfare in insurance markets using variation in prices. 2010.

Marco Francesconi and Jonathan James. None for the road? stricter drink driving laws and road accidents. *Journal of health economics*, 79:102487, 2021.

Qing Gong. Ex post moral hazard in automobile insurance markets with experience rating. 2017.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

Benjamin R Handel. Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–2682, 2013.

Bengt Holmström. Moral hazard and observability. *The Bell journal of economics*, pages 74–91, 1979.

Jason Huh and Julian Reif. Teenage driving, mortality, and risky behaviors. *American Economic Review: Insights*, 3(4):523–539, 2021.

Przemyslaw Jeziorski, Elena Krasnokutskaya, and Olivia Ceccarini. Adverse selection and moral hazard in the dynamic model of auto insurance. *UC Berkeley, Haas School of Business working paper*, 2017.

Andrew C Johnston. Unemployment insurance taxes and labor demand: Quasi-experimental evidence from administrative data. *American Economic Journal: Economic Policy*, 13(1):266–293, 2021.

Pinar Karaca-Mandic and Greg Ridgeway. Behavioral impact of graduated driver licensing on teenage driving risk and exposure. *Journal of health economics*, 29(1):48–61, 2010.

Victoria R Marone and Adrienne Sabety. When should there be vertical choice in health insurance markets? *American Economic Review*, 112(1):304–342, 2022.

NHTSA. Early estimate of motor vehicle traffic fatalities in 2022. Technical report, National Highway Traffic Safety Administration, 2023. URL https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813428.

Mark V Pauly. The economics of moral hazard: comment. *The american economic review*, 58(3):531–537, 1968.

Jean Pinquet. *Experience Rating in Nonlife Insurance*, pages 471–485. Springer New York, New York, NY, 2013. ISBN 978-1-4614-0155-1.

Yolanda Rebollo-Sanz, Jesús Rodríguez-López, and Núria Rodríguez-Planas. Penalty-point system, deterrence and road safety: A quasi-experimental analysis. *Journal of Economic Behavior & Organization*, 190:408–433, 2021.

Jonathan Roth, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 2023.

John W Ruser. Workers' compensation insurance, experience-rating, and occupational injuries. *The RAND Journal of Economics*, pages 487–503, 1985.

Matthew Shum and Yi Xin. Time-varying risk aversion? evidence from near-miss accidents. *Review of Economics and Statistics*, 104(6):1317–1328, 2022.

Sara Spaziani. Optimal unemployment insurance financing: Theory and evidence from two us states. Working paper, 2023.

Jörg L Spenkuch. Moral hazard and selection among the poor: Evidence from a randomized experiment. *Journal of health economics*, 31(1):72–85, 2012.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

Justin Sydnor. (over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199, 2010.

Deyu Wang, Qinyi Liu, Liang Ma, Yijing Zhang, and Haozhe Cong. Road traffic accident severity analysis: A census-based study in china. *Journal of Safety Research*, 70:135–147, 2019. ISSN 0022-4375.

Sarit Weisburd. Identifying moral hazard in car insurance contracts. *Review of Economics and Statistics*, 97(2):301–313, 2015.

WHO. Road traffic deaths, data by country. Technical report, World Health Organization, 2022a. URL https://apps.who.int/gho/data/node.main.A997.

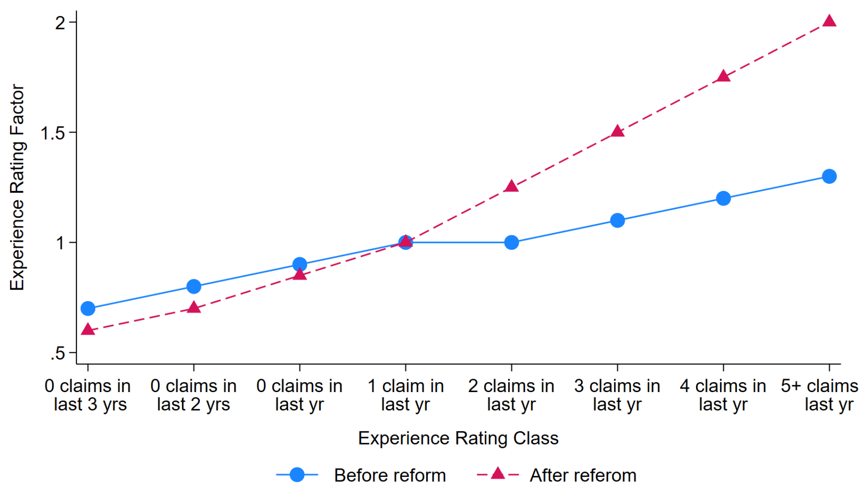
WHO. Road traffic injuries. Technical report, World Health Organization, 2022b. URL https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Ralph A Winter. Optimal insurance contracts under moral hazard. *Handbook of insurance*, pages 205–230, 2013.

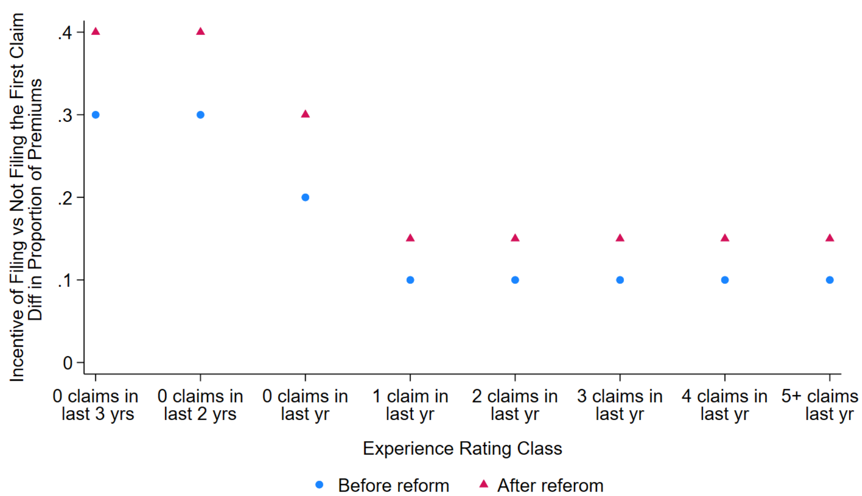
Wei Zheng, Yi Yao, Peng Shi, Yinglu Deng, and Hao Zheng. Deregulation, competition, and consumer choice of insurer: Evidence from liberalization reform in china's automobile insurance market. *The Geneva Risk and Insurance Review*, pages 1–43, 2021.

Figures and Tables

Figure 1: Experience Rating Factor Changes Before and After Reform



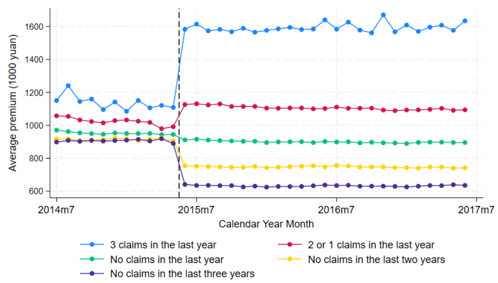
(a) Experience Rating Factor



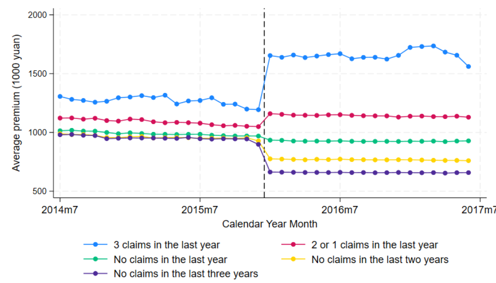
(b) Implied incentives to File v.s. Not Filing the First Claim in a Coverage Year

Note: The figure plots the experience rating factor before and after the reform. The x-axis shows the experience rating class in the past year. The y axis in Panel (a) shows the experience rating factor corresponding to the experience rating class. The y axis in Panel (b) shows the incentives of filing vs not filing the first claim in a coverage year for the insured in different experience rating class.

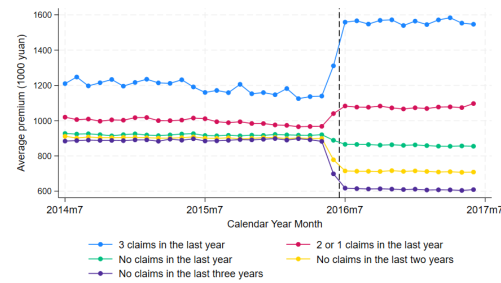
Figure 2: Trends of Premiums of Voluntary Liability and Collision Coverage by Risk Class and Reform Group



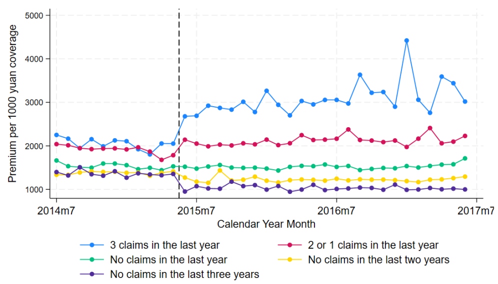
(a) Voluntary Liability Coverage - Reform Group 1



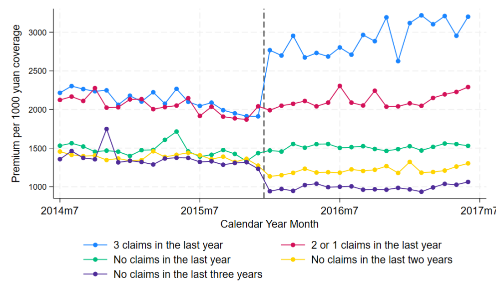
(b) Voluntary Liability Coverage - Reform Group 2



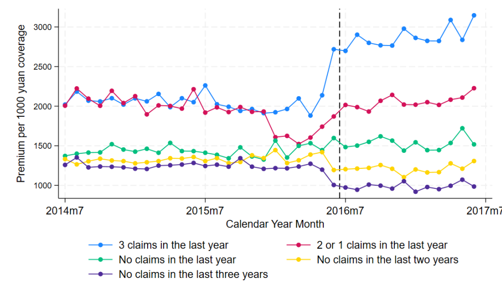
(c) Voluntary Liability Coverage - Reform Group 3



(d) Voluntary Collision Coverage - Reform Group 1



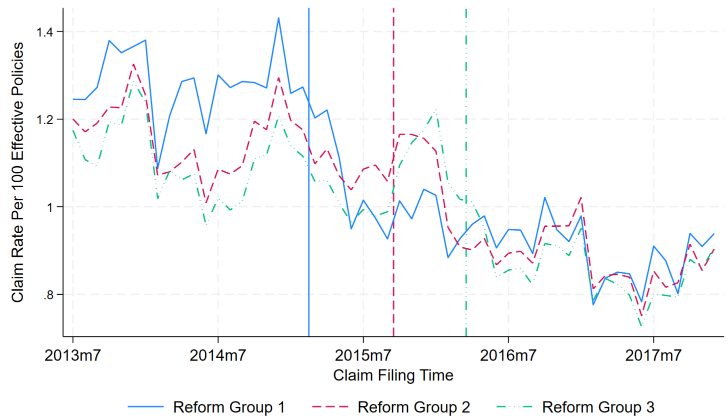
(e) Voluntary Collision Coverage - Reform Group 2



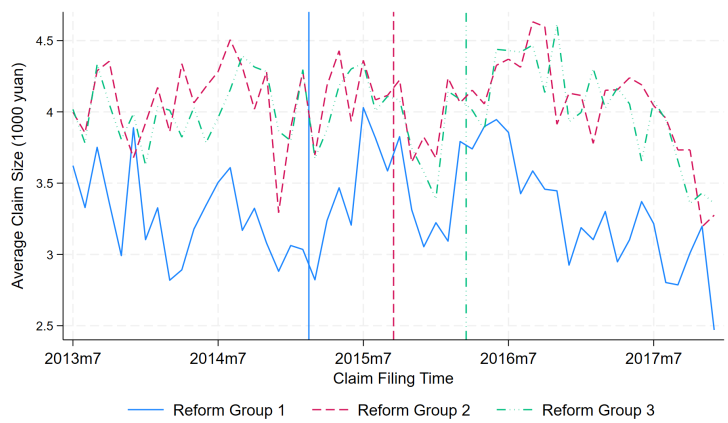
(f) Voluntary Collision Coverage - Reform Group 3

Note: Panel (a) - (c) plots average premiums of voluntary liability coverage with 300,000 yuan coverage limit filed in each month for each risk class in each reform group. Panel (d) - (f) plots average premiums of voluntary collision coverage for each risk class in each reform group. Panel (g) - (i) plots average premiums of voluntary liability coverage normalized by the vehicle market value. The vertical line indicates the announced start date of the reform in each reform group.

Figure 3: Raw Data Plots of Claim Rate and Size of Compulsory Liability Coverage



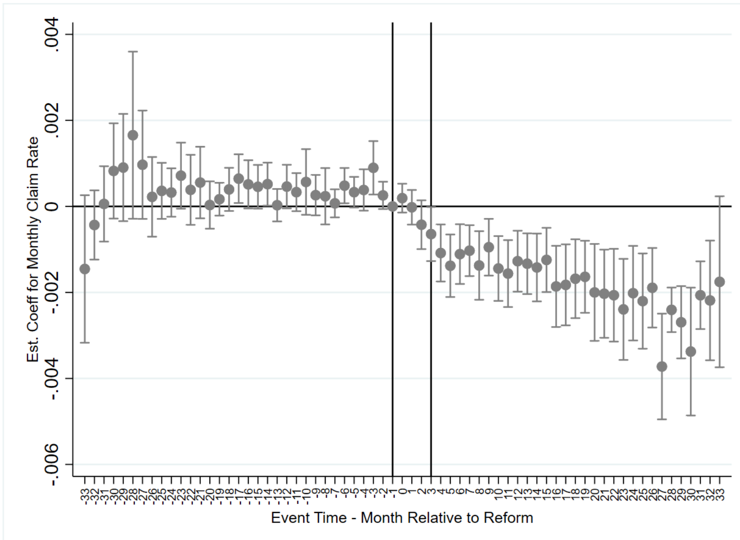
(a) Claim Rate of Compulsory Coverage



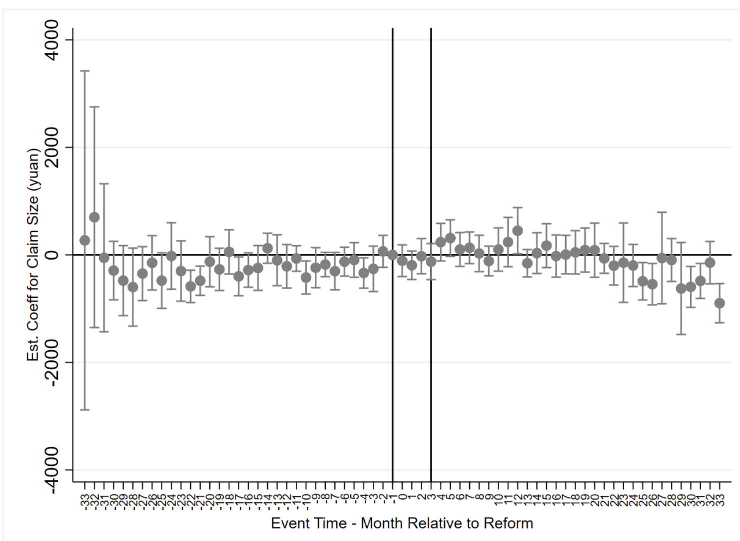
(b) Claim Size of Compulsory Coverage

Note: The figures plot average claim rate or average claim size of compulsory liability coverage over time in each region. The vertical line indicates the date when the reform is announced. The actual start time is usually within two or three months of the announcement.

Figure 4: Dynamic TWFE Estimates of Effects on Claims for Compulsory Liability Coverage



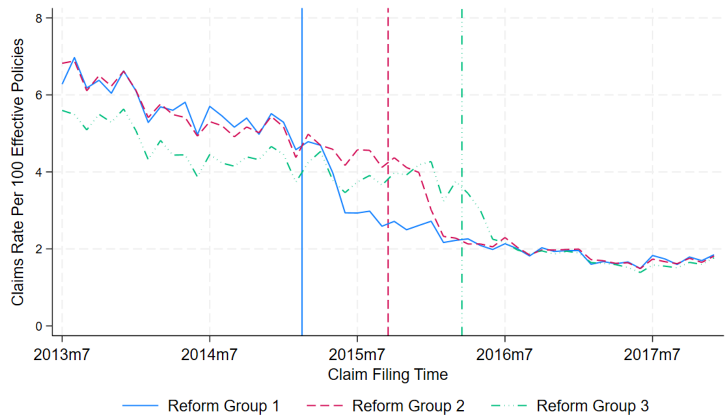
(a) Claim Rate



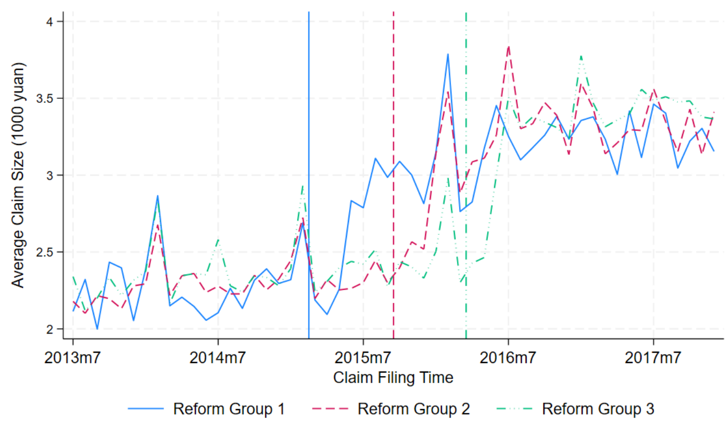
(b) Claim Size

Note: The figures plot estimated coefficients γ_e^c from Equation 3 separately for claim rate and claim size of compulsory liability coverage. Observation is at the **claim-year-month level**. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 5: Raw Data Plot of Claim Rate and Size of Voluntary Collision Coverage



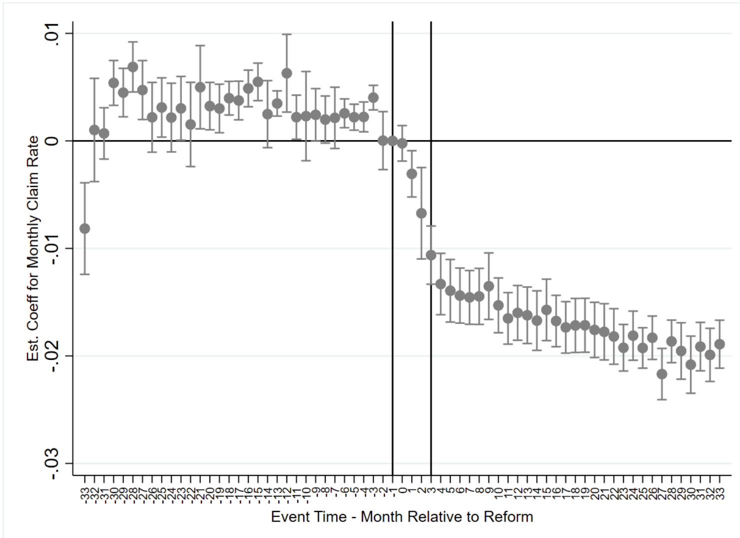
(a) Claim Rate



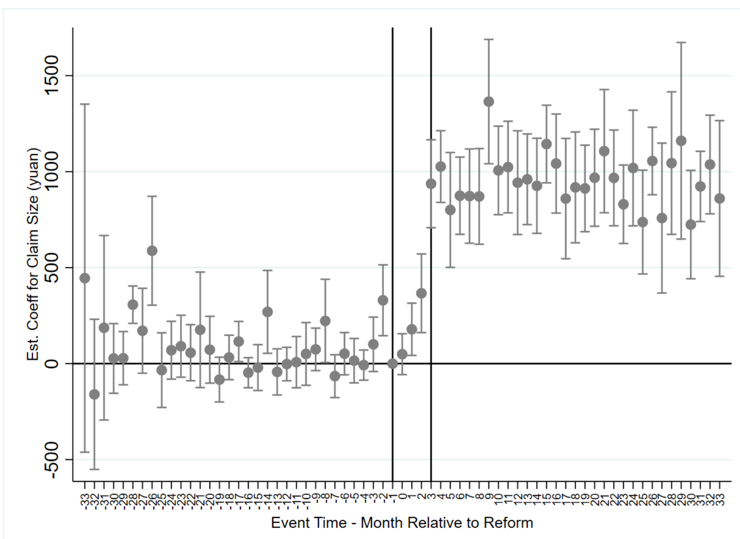
(b) Claim Size

Note: The figures plot average claim rate or average claim size of collision damage coverage among always buyers over time in each region. The vertical line indicates the date when the reform is announced. The actual start time is usually within two or three months of the announcement.

Figure 6: Dynamic TWFE Estimates of Effects on Claims for Voluntary Collision Coverage Among Always Buyers



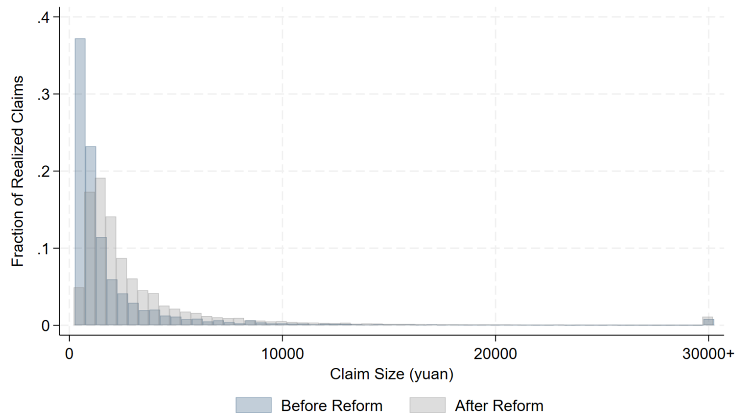
(a) Claim Rate



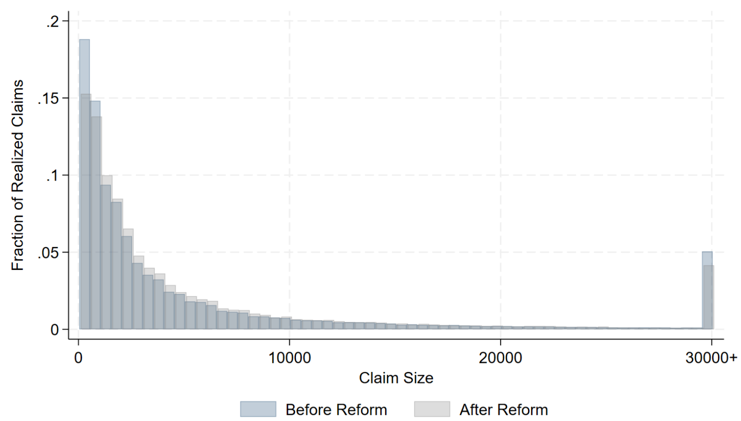
(b) Claim Size

Note: The figures plot estimated coefficients γ_e^c from Equation 3 separately for claim rate and claim size of collision damage coverage. Observation is at the claim-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 7: Histogram of Claim Size for Voluntary Collision Coverage and Voluntary Liability Coverage Before and After Reform



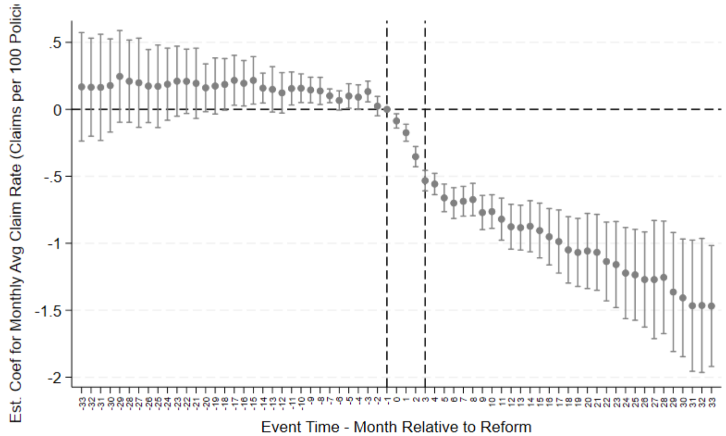
(a) Collision Damage Coverage



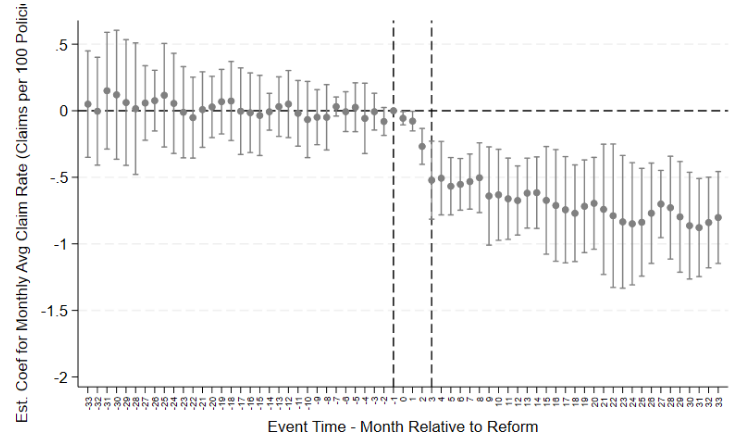
(b) Third Party Liability Coverage

Note: The figures plot histogram of realized claim size (> 0) for collision coverage and voluntary liability coverage. In panel a-b, the bin size is 500 yuan (70 usd). The last bar indicates proportion of claims equal or greater than 30,000 yuan.

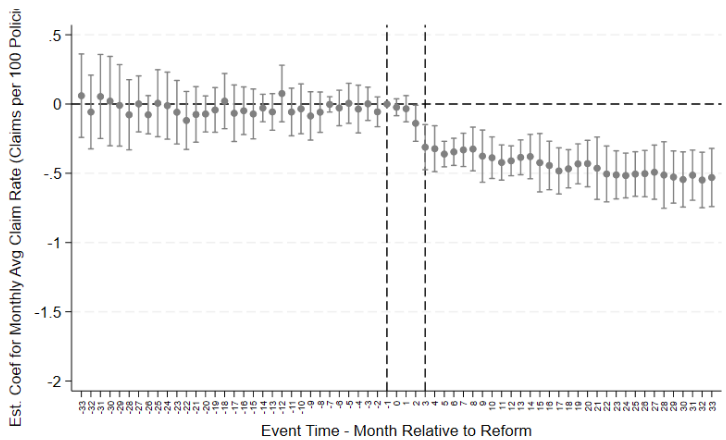
Figure 8: TWFE Estimates of Claim Rate For Voluntary Collision Coverage by Size



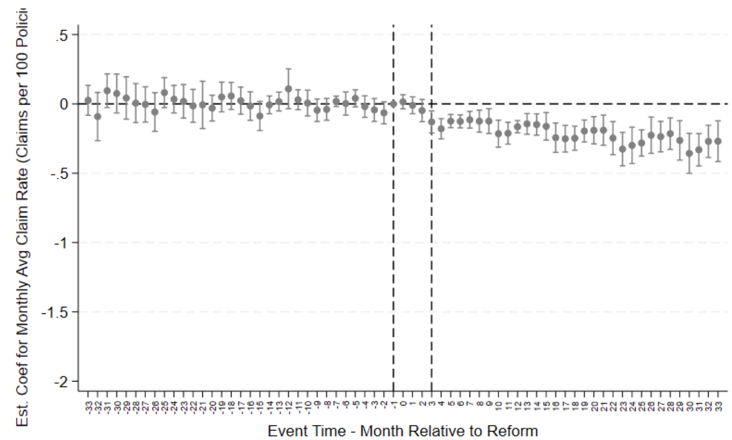
(a) Claim Size Below 1st Quartile



(b) Claim Size Between 1st Quartile - 2nd Quartile



(c) Claim Size Between 2nd Quartile - 3rd Quartile



(d) Claim Size Above 3rd Quartile

Note: Figure 8 plots estimated coefficients from Equation 3, where the outcome variable is the claim rate for collision damage coverage claims, the size of which is within a specific range. Observation is at the jurisdiction-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Table 1: Summary Statistics

	Full Sample				
	Obs	Mean	Std. dev.	Min	Max
Compulsory Liability Coverage					
Number of claims	7,604,293	0.0791	0.3035	0	7
Claim cost (yuan)	7,604,293	274.9	3,609.1	0	209,735
Claim size (conditional on claiming) (yuan)	605,737	3,474.1	12,302.5	0.01	134,100
Voluntary Liability Coverage					
Number of claims	7,595,315	0.0404	0.2042	0	5
Claim cost (yuan)	7,595,315	325.5	6,060.1	0	1,000,000
Claim size (conditional on claiming) (yuan)	307,133	8,050.3	29,037.0	0.01	1,000,000
Voluntary Collision Coverage					
Number of claims	6,047,334	0.3449	0.6467	0	14
Claim cost (yuan)	6,047,334	927.1	4,027.4	0	727,980
Claim size (conditional on claiming) (yuan)	2,085,836	2,687.9	6,111.5	0.01	724,554

Note: This table reports the summary statistics for different types of coverage. The observation level is at policy year level. Always buyers are those vehicles who always purchase voluntary collision coverage during our sample period.

Table 2: Effects of the Reform on Claim Rate and Size

	Claim Rate			Claim Size		
	Compulsory Liability Full Sample	Voluntary Liability Full Sample	Voluntary Collision Always buyers	Compulsory Liability Full Sample	Voluntary Liability Full Sample	Voluntary Collision Always buyers
Anticipation	-0.00032	-0.00027**	-0.004***	24	-159	106**
	(0.0002)	(0.00012)	(0.00063)	(62)	(382)	(48)
Post	-0.0021***	-0.00062***	-0.015***	88	-16	687***
	(0.00039)	(0.00014)	(0.0014)	(103)	(432)	(56)
N	65,661,884	65,564,466	52,741,477	481,134	157,658	1,445,000
Pre-Reform Mean Y	0.01	0.00	0.03	3,366.31	7,458.46	2,612.86
Anticipation % of Pre-Reform Mean	-4.20	-7.50	-13.83	0.71	-2.13	4.06
Post % of Pre-Reform Mean	-27.65	-17.36	-51.61	1.61	-0.21	26.29

Note: The table reports estimates from Equation 4. The observation level is at stack-policy/claim-month level. We controls for stack by calendar-year-month fixed effects and stack by province fixed effects. Full sample indicates all policies/claims, and always buyers indicates policies/claims of individuals who always purchased voluntary collision coverage during our sample period. Standard errors are clustered at province/city level. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table 3: Welfare Effects of Experience Rating Factor Changes

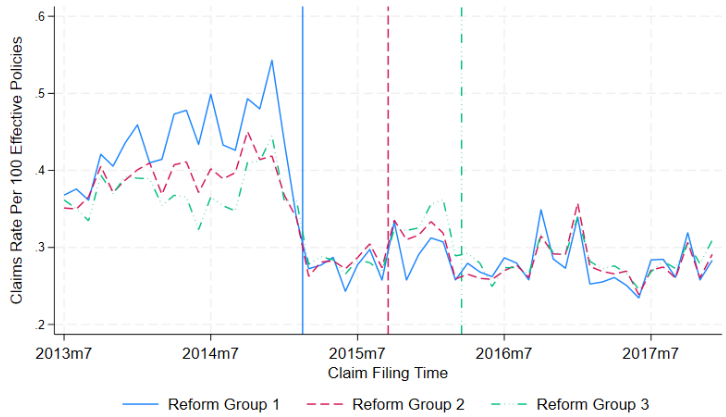
	Risk Class					
	4+ claims last year	**3 claims last year**	**1-2 claims last year**	**0 claims last year**	**0 claims last 2 years**	**0 claims last 3 years**
Proportion of insureds	0.20%	1%	41%	25%	15%	17%
Experience rating factor θ_r						
Before reform	0.11	0.11	0.11	0.25	0.43	0.43
After reform	0.18	0.18	0.18	0.43	0.67	0.67
Change	0.07	0.07	0.07	0.18	0.24	0.24
Base premium, average claim rate and average cost						
Base premium (yuan)	2542.5	2542.5	2542.5	2260	1977.5	1977.5
Pre-reform annual claim rate of any first voluntary coverage claim	0.469	0.466	0.328	0.224	0.170	0.122
Pre-reform average cost (yuan) of first claim	1741	1545	1131	832	665	493
Changes in claim rate and average cost						
Overall change in annual claim rate	-.0984	-.144	-.132	-.078	-.0624	-.0432
Overall change in avg cost (yuan)	-552	-22.8	-204	-144	-99.6	-80.4
Baseline Welfare Changes						
Implied Premium Change Under Competitive Eqm (yuan)	-77.6	-77.6	-77.6	-220.0	-226.4	-226.4
Increase in expected consumption (yuan)	6.6	7.0	28.0	145.2	165.9	183.0
Decrease in risk protection - Baseline (yuan)	-25.5	-21.6	-16.1	-33.4	-20.6	-15.7
Overall welfare - Baseline (yuan)	32.2	28.6	44.1	178.7	186.6	198.7
Bounded Welfare Changes						
Decrease in risk protection - LB (yuan)	9.8	9.8	6.9	14.5	18.8	13.5
Decrease in risk protection - UB (yuan)	-131.6	-115.7	-85.0	-177.2	-138.9	-103.4
Overall welfare - LB (yuan)	-3.2	-2.8	21.1	130.8	147.1	169.5
Overall welfare - UB (yuan)	138.3	122.7	113.0	322.4	304.8	286.4

Note: This table presents the welfare changes along with summary statistics and intermediary steps required to calculate these changes. It begins with the proportion of insureds per risk class and experience rating factors before and after the reform. It then displays the average base premium, claim rate and cost as well as the changes due to the reform. Last, it presents the welfare changes under varying assumptions, along with breakdowns into increase in expected consumption level as well as reductions in moral hazard. Baseline indicates the unclaimed cost is set to one fourth of average cost, lower bound (LB) indicates the unclaimed cost is zero, and upper bound (UB) indicates the unclaimed cost is the average cost.

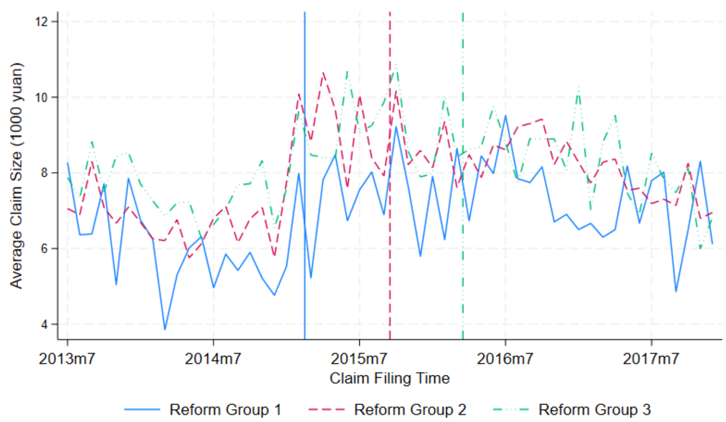
A Appendix Figures

A.1 Raw Data Plot and TWFE Estimates of Voluntary Collision Coverage

Figure 9: Raw Data Plot of Claim Rate and Size of Voluntary Liability Coverage



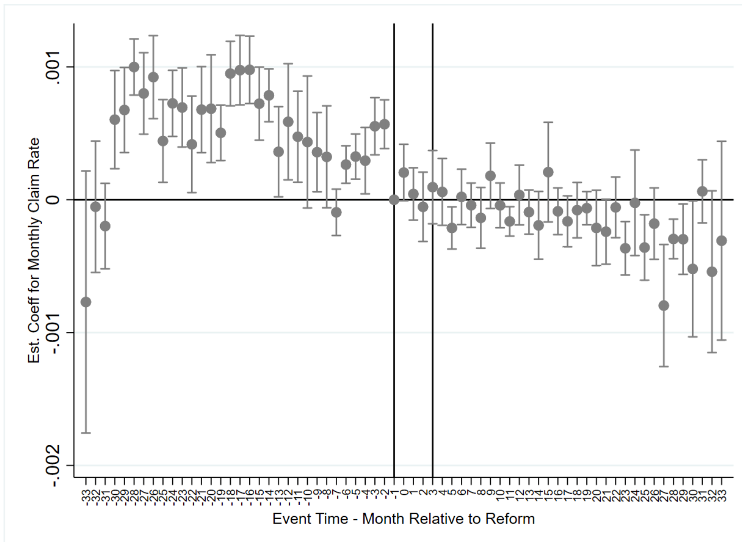
(a) Claim Rate



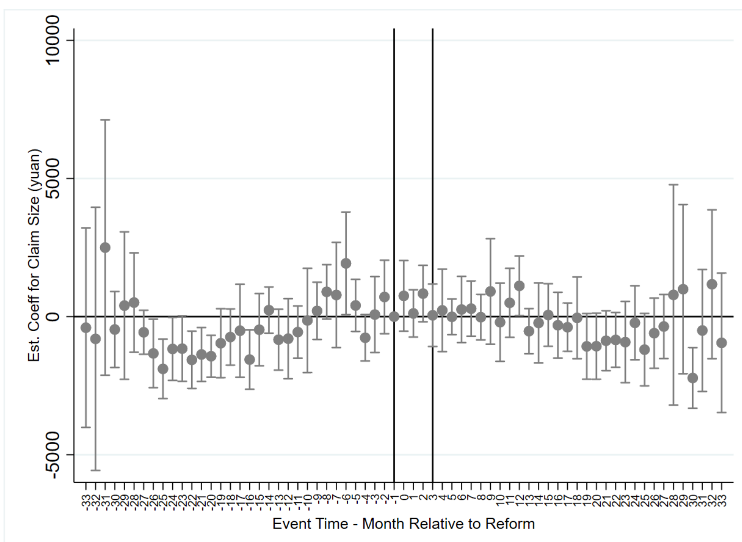
(b) Claim Size

Note: The figures plot average claim rate or average claim size of voluntary liability coverage over time in each region. The vertical line indicates the date when the reform is announced. The actual start time is usually within two or three months of the announcement.

Figure 10: Dynamic TWFE Estimates of Effects on Claims for Voluntary Liability Coverage - Claim data



(a) Claim Rate

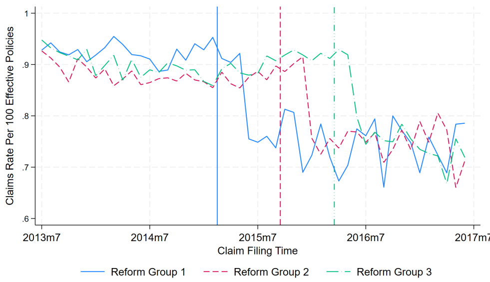


(b) Claim Size

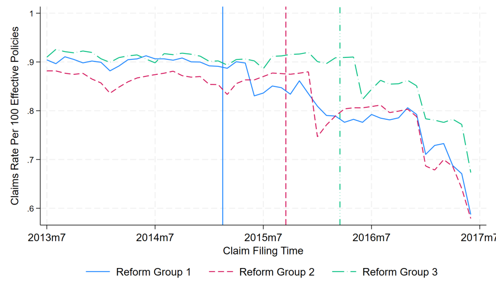
Note: The figures plot estimated coefficients γ_e^c from Equation 3 separately for claim rate and claim size of voluntary liability coverage. Observation is at the claim-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

A.2 Premium Changes and Demand Responses

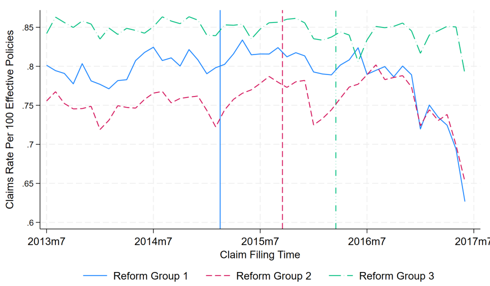
Figure 11: Times Series of Share of Consumers Purchasing Collision Damage Coverage



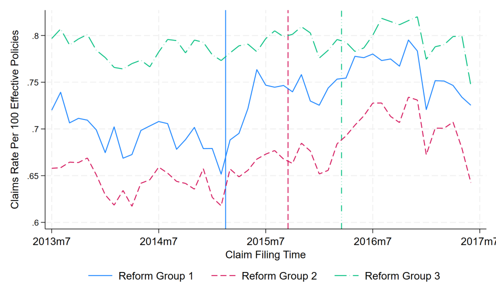
(a) 3 Claims Last Year



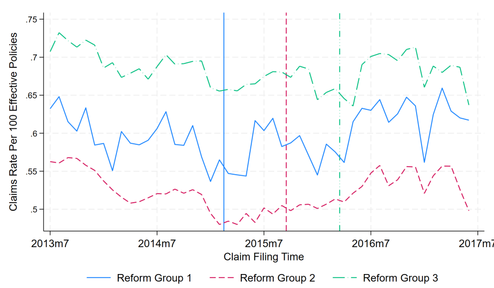
(b) 1 or 2 Claims Last Year



(c) No Claims Last Year



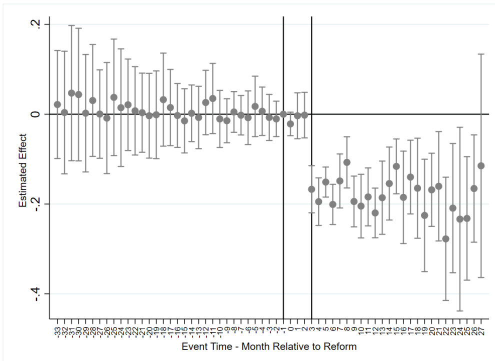
(d) No Claims Last 2 Year



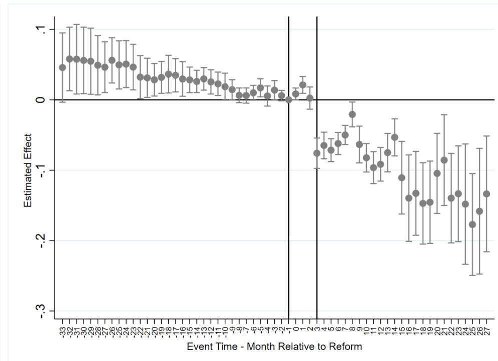
(e) No Claims Last 3 Year

Note: Figure 11 plots the time series of enrollment share of voluntary collision coverage. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

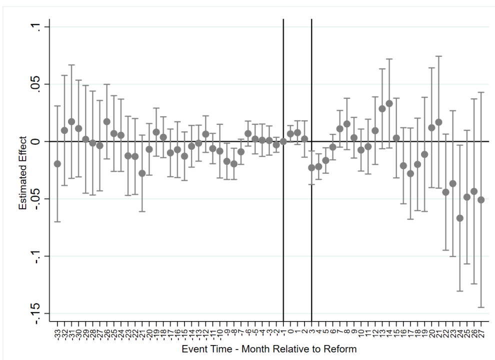
Figure 12: TWFE Estimates of Changes in Share of Consumers Purchasing Collision Damage Coverage



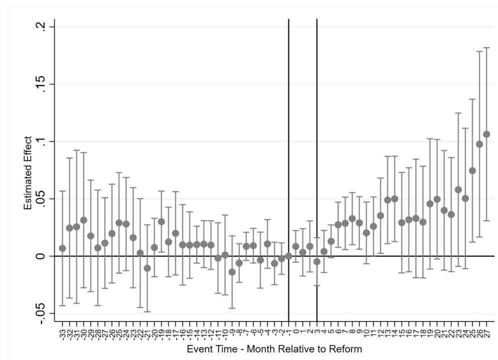
(a) 3 Claims Last Year



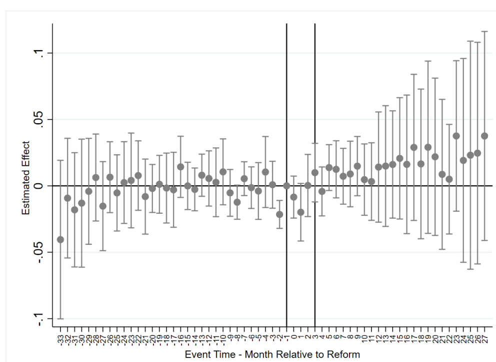
(b) 1 or 2 Claims Last Year



(c) No Claims Last Year



(d) No Claims Last 2 Year

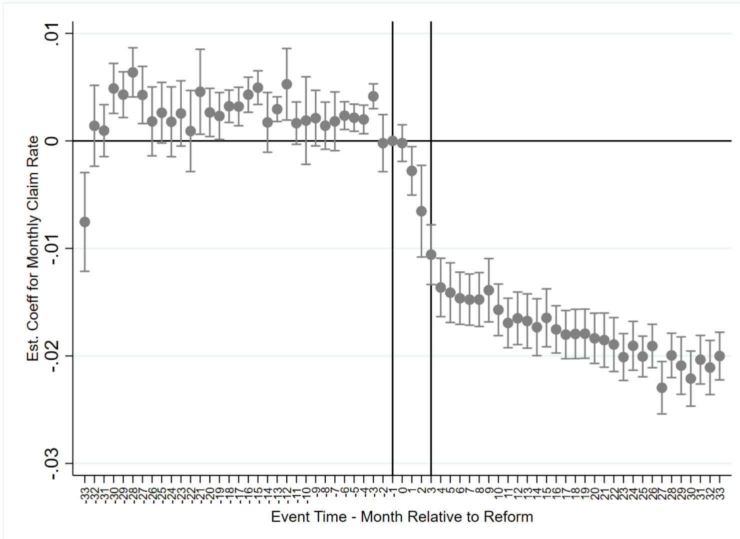


(e) No Claims Last 3 Year

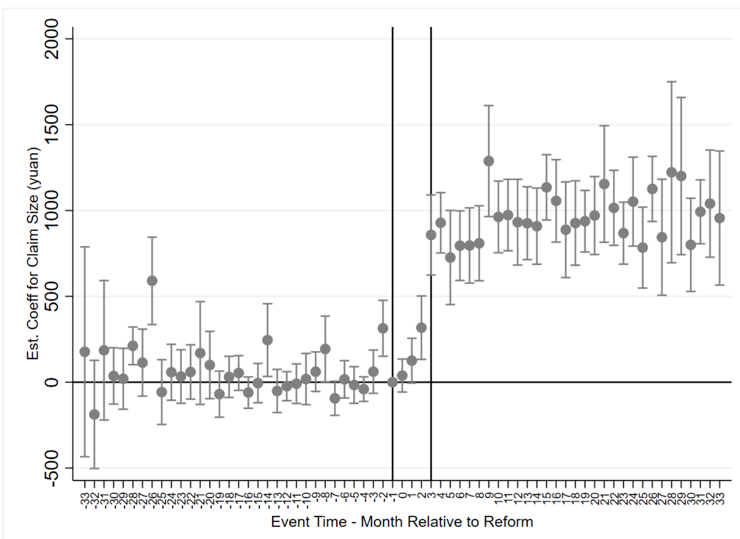
Note: Figure 12 plots estimated coefficients from Equation 3, where the outcome variable is whether consumers purchase voluntary collision coverage. Observation is at the policy level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

A.3 Results Using All Voluntary Collision Coverage

Figure 13: Dynamic TWFE Estimates of Effects on Claims for Voluntary Collision Coverage Using All Policies



(a) Claim Rate



(b) Claim Size

Note: The figures plot estimated coefficients γ_e^c from Equation 3 separately for claim rate and claim size of collision damage coverage. Observation is at the claim-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

A.4 Welfare Changes with Different CARA Coefficients γ

Figure 14: Welfare Changes with Different CARA Coefficients γ

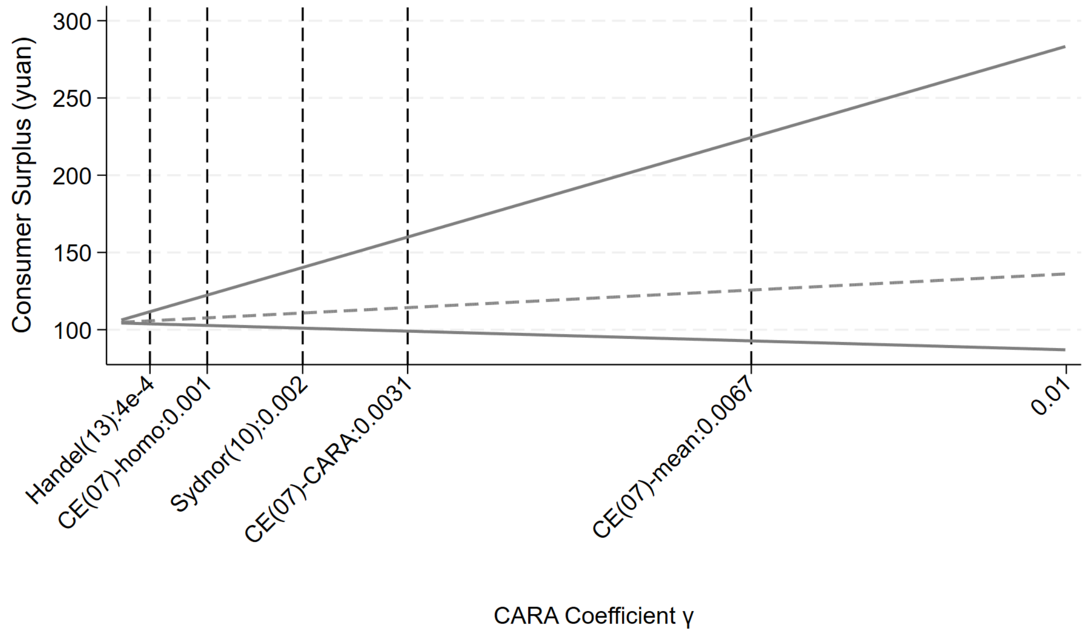


Figure 15: Overall Welfare Changes

Note: The figure how estimates of aggregated welfare changes with different magnitudes of CARA coefficients. The x-axis is the CARA measured in the inverse USD unit. The actual CARA used in the welfare calculation is divided by the average exchange rate 6.4 during the sample period. Handel(13) refers to the CARA estimates by Handel (2013). CE(07) refers to Cohen and Einav (2007), where homo indicates the estimates under the assumption of homogeneity, CARA indicates the estimates using CARA utility functions, and mean indicates the mean of CARA estimates using the baseline model. Sydnor(10) refers to estimates by Sydnor (2010).

B Robustness Check

B.1 Robustness Checks of Dynamic TWFE Estimates

We have presented the raw data plots of different types of coverage in the three reform groups. To summarize and quantify the time series results presented in the raw data plots, we estimate an event study model that only include the event-time dummies and province/city fixed effects. Our event study results provide reasonable estimates of the policy's effects. Unlike difference-in-differences estimates, our event study regressions do not control calendar year-month fixed effects. As a result, it might capture effects of events coincided with the reform, which is not a significant issue in our setting, as our raw data plots show that outcomes of interest experiences similar trends across all groups until the announcement of the reform and started to change only after the reform and only within the reform groups. This provides sufficient evidence of the relative changes over time. Additionally, the event study regressions avoid the recent concerns in DID literature regarding potential bias in the TWFE DID estimates in a staggered adoption style (Callaway, 2022; Roth et al., 2023). The event study estimating equation is

$$y^c_{irt} = \sum_{e \neq -1} \alpha^c_e D^e_{irt} + \mu_r + \varepsilon_{rt}. \tag{22}$$

Our dynamic TWFE regressions provide sensible causal effects of the reform. However, recent literature on DID suggests that dynamic TWFE estimates may not be robust in the presence of heterogeneous effects across treatment groups. The dynamic specification yields a sensible causal estimand when heterogeneity is only related to the time since treatment (Goodman-Bacon, 2021; Callaway, 2022; Roth et al., 2023). If treatment effects are heterogeneous across adoption groups, the "treatment lead" coefficients are not guaranteed to be zero even if parallel trends hold in all periods (and vice versa). Consequently, using these coefficients to evaluate pre-trends could lead to misleading conclusions. In contrast, if the treatment effects are homogeneous across groups, dynamic TWFE DID regression provides sensible estimates of the dynamic effects, that is, the coefficient γ_e capture the homogeneous treatment effects in the e-th period after treatment under reasonable parallel trend assumption (Sun and Abraham, 2021). Based on patterns observed in our raw data plots, heterogeneity across groups does not appear to be a significant issue. Therefore, the dynamic TWFE regression provides us sensible estimates of dynamic effects of the experience rating reform.

However, to mitigate any concerns about potential biases in our TWFE estimates and establish the robustness of our estimates, we use both the stacked DID approach (Butters et al., 2022) and the approach proposed by Callaway and Sant'Anna (2021) to provide alternative treatment effect estimates. However, using the methods suggested by recent literature comes with costs. These approaches emphasize constructing better controls for the treatment group and avoiding unfavorable comparisons between late-treated and already-treated groups. Typically, never-treated groups are used to construct suitable control groups. In our specific context, we do not have a never-treated group; instead, we can only use not-yet-treated groups. Consequently, our ability to estimate dynamic effects is limited due to short intervals between reform start dates in the three reform groups. Furthermore, we lose some statistical power, as the third reform group cannot be matched with a suitable control group and is only used as a control group for estimating the reform effects.

The estimating equation of stacked DID for dynamic effects is:

$$y^c_{irst} = \sum_{e \neq -1} \beta^c_e D^e_{irst} + \sum_{e \neq -1} d^e_{irst} + \mu_{rs} + \tau_{ts} + \varepsilon_{irst}, \tag{23}$$

where the outcome variable is whether consumer i or size of claim i of coverage type c in the province/city r stack s in year-month t, D^e_{irst} is an event-time dummy, which equals to 1 if consumer i in region r stack s has been treated for exactly e periods in period t from the reform announcement date and 0 otherwise. d^e_{irst} is a relative time dummy, which equals to 1 if consumer i in region r stack s is exactly e periods in period t from the reform announcement date of treated group in the stack and 0 otherwise. μ_{rs} denotes province/city-stack specific fixed effects, and τ_{ts} denotes the calendar year-month-stack fixed effects. β^c_k are interpreted as the difference in outcome of interest in event period e relative to the baseline period between treatment and control provinces. Standard errors are clustered at the province/city level.

To estimate the average treatment effects, we also estimate event study model and TWFE model. The estimating equation for event study model is

$$y^c_{rt} = \beta_1 \text{ anticipation}_{rt} + \beta_2 \text{ post}_{rt} + \mu_r + \varepsilon_{rt},$$

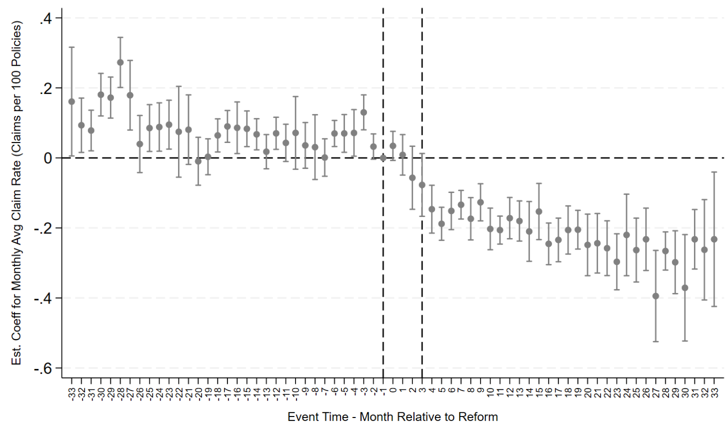
and the estimating equation for TWFE model is

$$y_{rt}^{c} = \beta_1 \text{ anticipation}_{rt} + \beta_2 \text{post}_{rt} + \mu_r + \tau_t + \varepsilon_{rt},$$

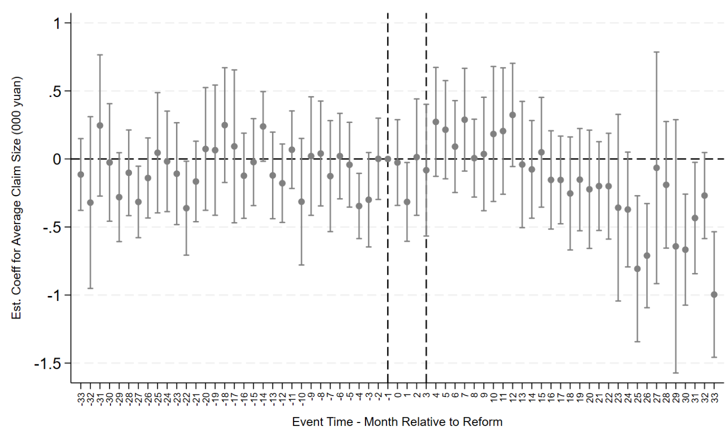
where where we additionally control calendar year-month fixed effects, denoted by τ_t.

B.1.1 Event Study Estimates

Figure 16: Event Study Estimates of Effects on Claims for Compulsory Liability Coverage



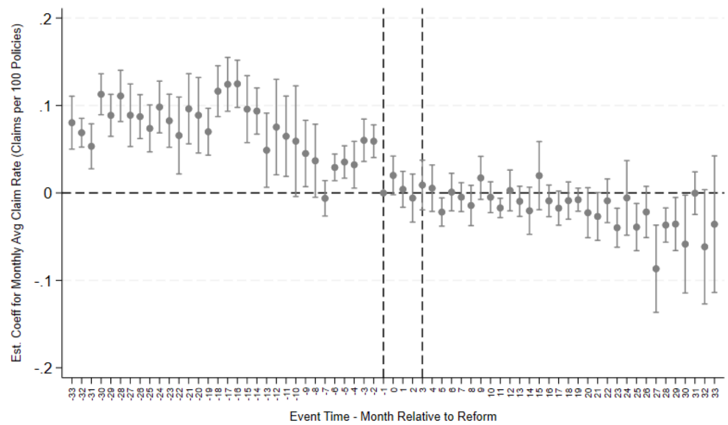
(a) Claim Rate



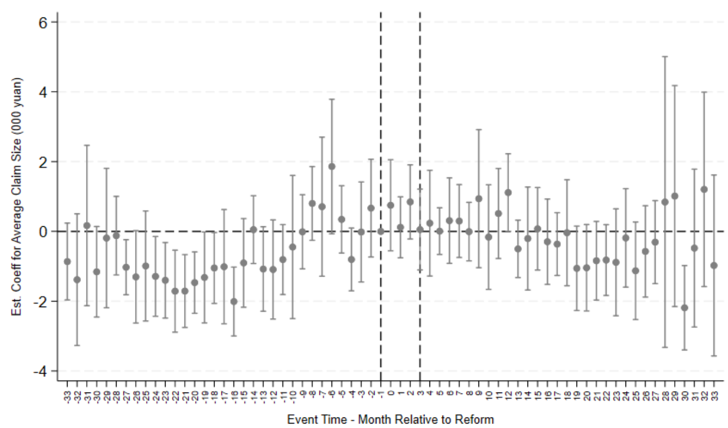
(b) Claim Size

Note: The figures plot estimated coefficients α_e^c from Equation 22 separately for claim rate and claim size of compulsory liability coverage. Observation is at the jurisdiction-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 17: Event Study Estimates of Effects on Claims for Voluntary Liability Coverage



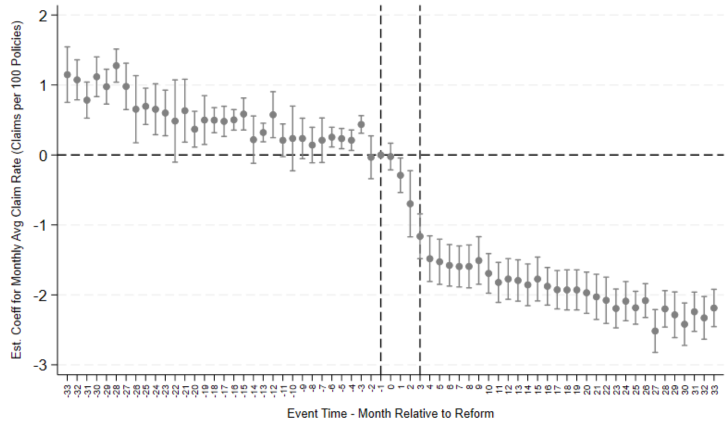
(a) Claim Rate



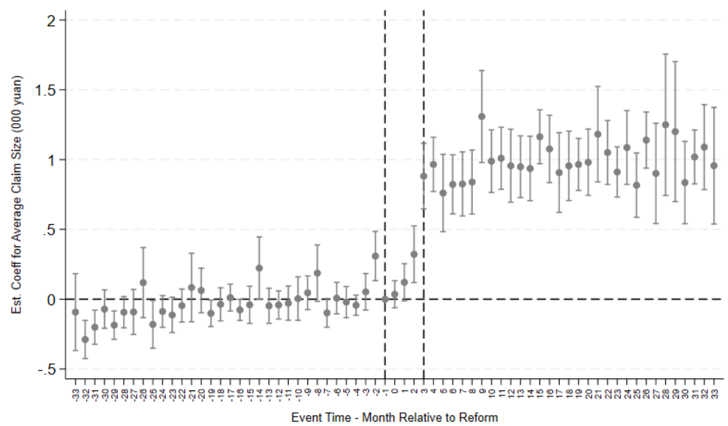
(b) Claim Size

Note: The figures plot estimated coefficients α_e^c from Equation 22 separately for claim rate and claim size of voluntary liability coverage. Observation is at the jurisdiction-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 18: Event Study Estimates of Effects on Claims for Voluntary Collision Coverage



(a) Claim Rate

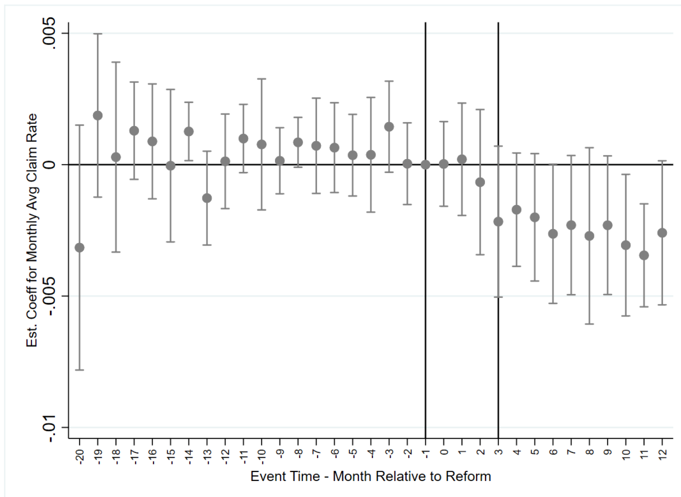


(b) Claim Size

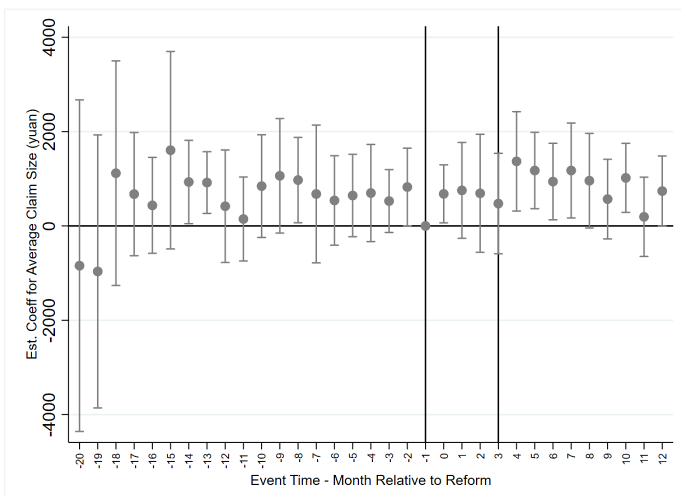
Note: The figures plot estimated coefficients α_e^c from Equation 22 separately for claim rate and claim size of collision damage coverage. Observation is at the jurisdiction-year-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

B.1.2 Results Using Stacked DID - Reform Group 1 vs Reform Group 3

Figure 19: Dynamic Stacked DID Estimates of Effects on Claims for Compulsory Liability Coverage



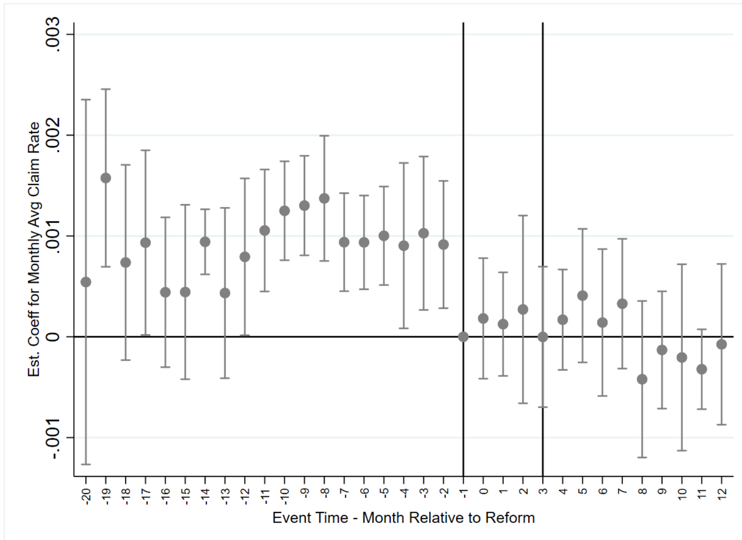
(a) Claim Rate



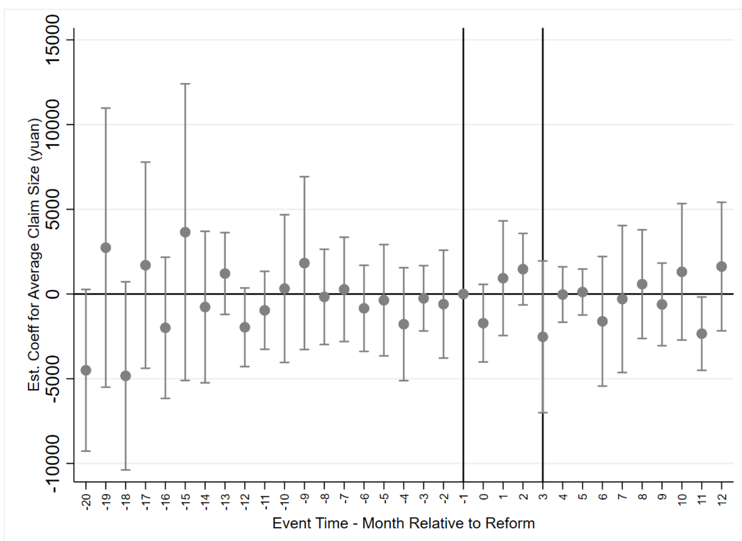
(b) Claim Size

Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of compulsory liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 20: Dynamic Stacked DID Estimates of Effects on Claims for Voluntary Liability Coverage



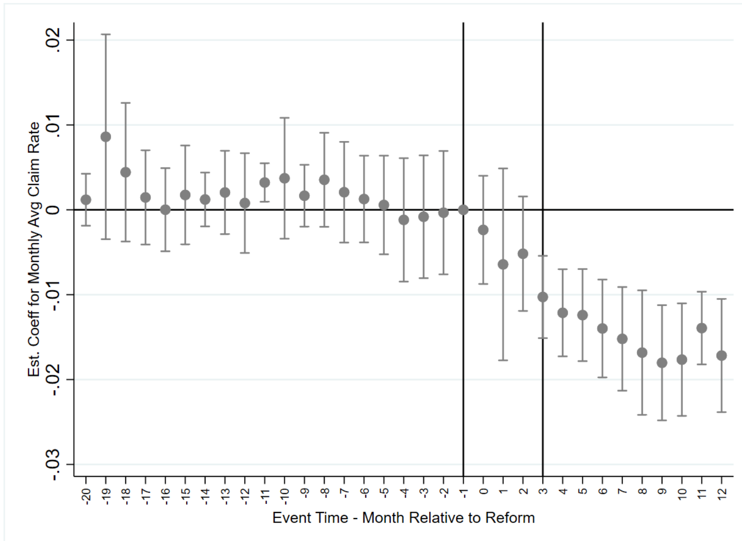
(a) Claim Rate



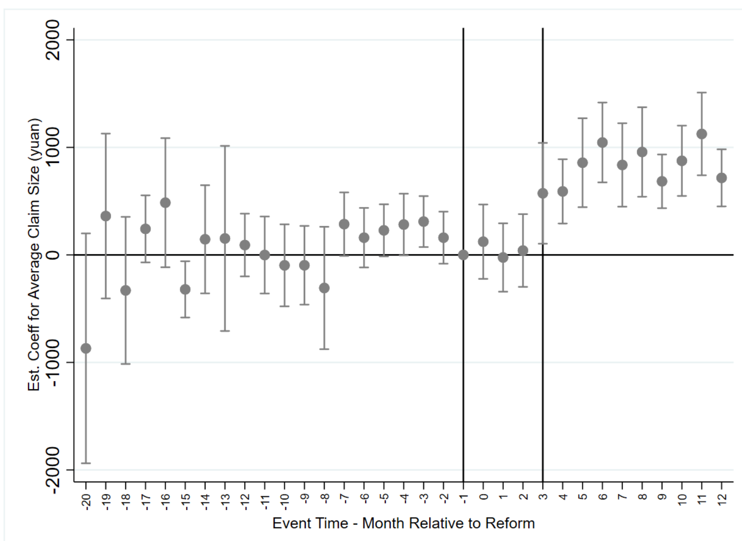
(b) Claim Size

Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of voluntary liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 21: Dynamic Stacked DID Estimates of Effects on Claims for Voluntary Collision Coverage



(a) Claim Rate

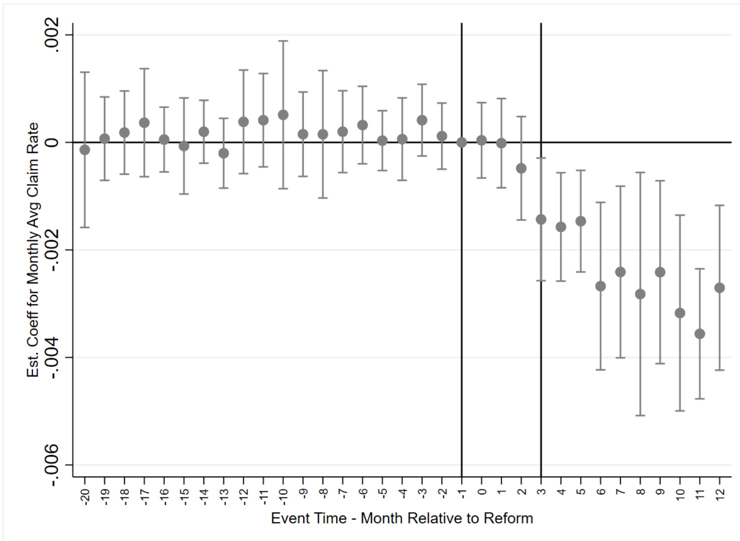


(b) Claim Size

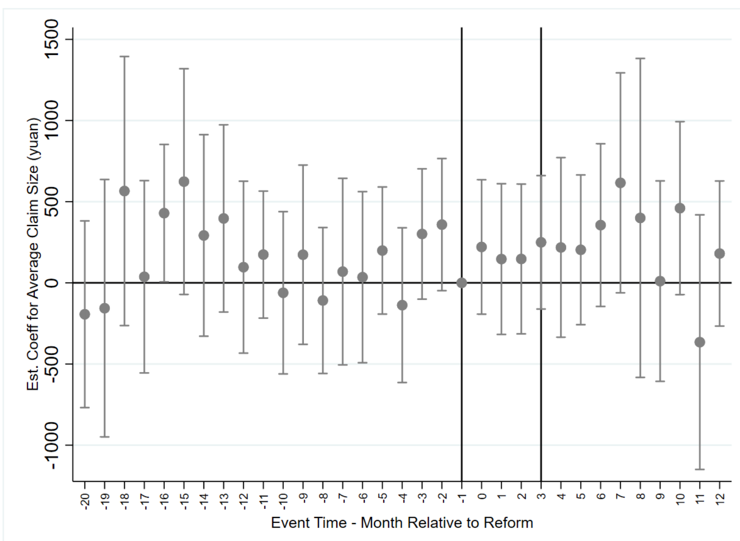
Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of compulsory liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

B.1.3 Results Using Stacked DID - Two Stacks

Figure 22: Dynamic Stacked DID Estimates of Effects on Claims for Compulsory Liability Coverage



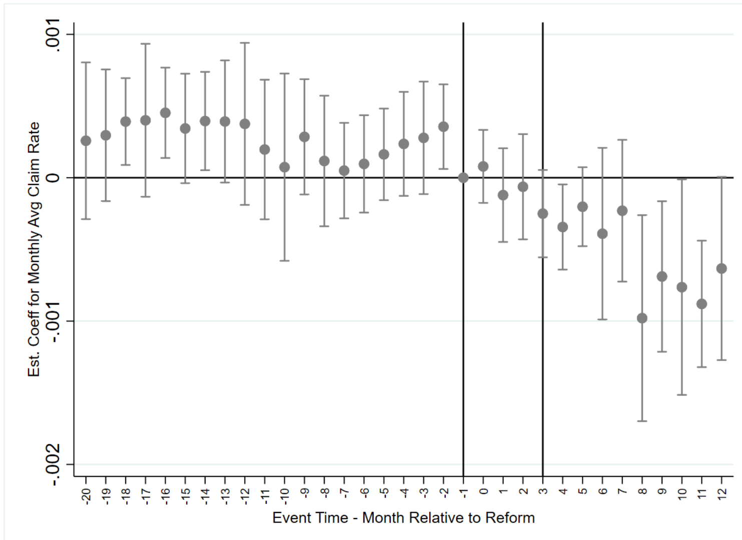
(a) Claim Rate



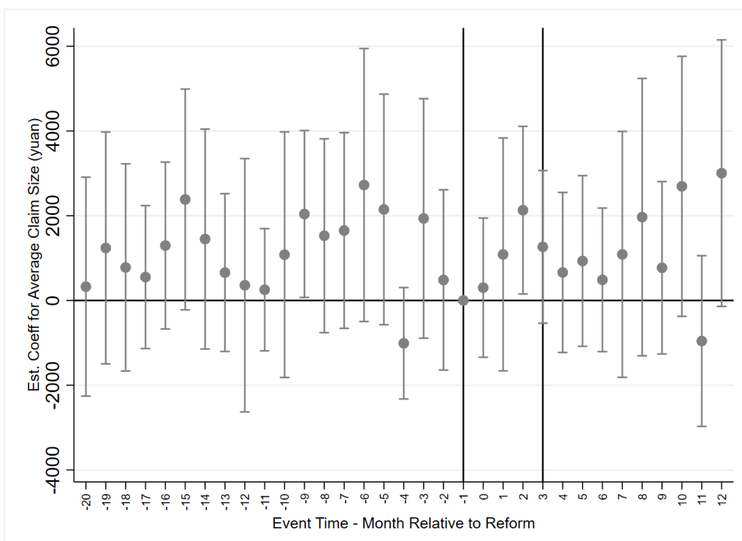
(b) Claim Size

Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of compulsory liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 23: Dynamic Stacked DID Estimates of Effects on Claims for Voluntary Liability Coverage



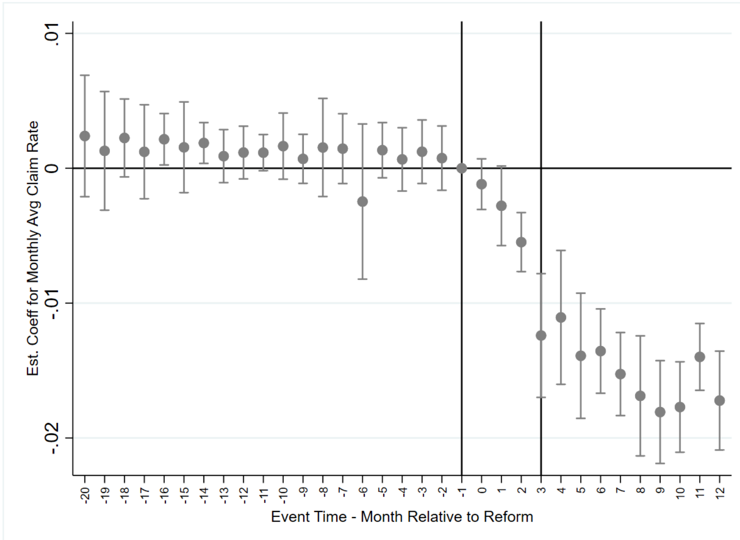
(a) Claim Rate



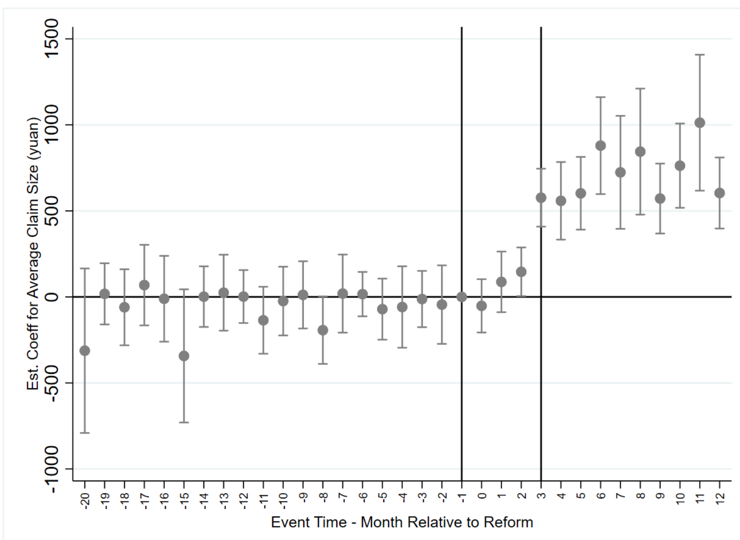
(b) Claim Size

Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of voluntary liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

Figure 24: Dynamic Stacked DID Estimates of Effects on Claims for Voluntary Collision Coverage



(a) Claim Rate



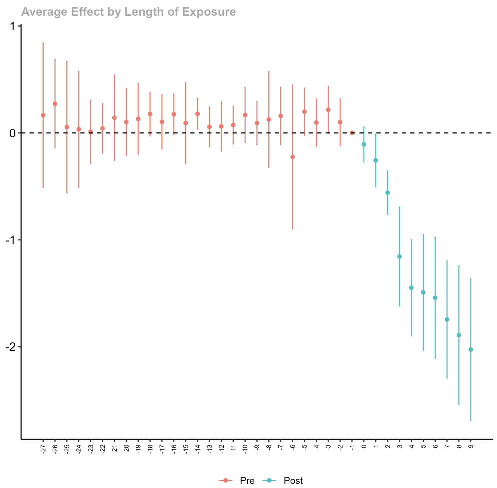
(b) Claim Size

Note: The figures plot estimated coefficients β_e^c from Equation 23 separately for claim rate and claim size of compulsory liability coverage. The treatment group is reform group 1 and the control group is reform group 3. Observation is at the policy/claim-month level. The point denotes the point estimate, with -1 period is the omitted baseline. The bar indicates the 95% confidence interval. The first vertical line indicates the announcement of the reform, and the second vertical line indicates the first three months after initial announcement, that is, the start month in the announcement date.

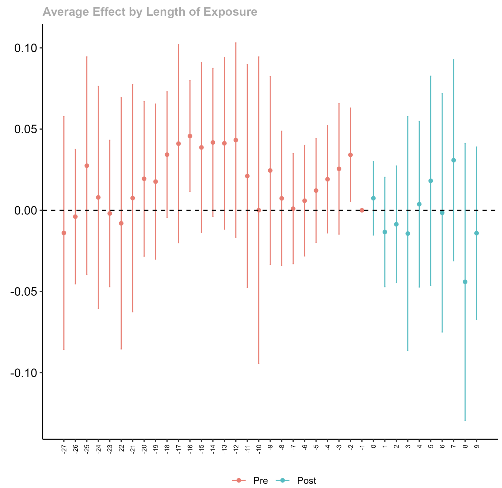
B.2 Estimates Using Method Proposed by Callaway and Sant'Anna (2021)

We estimate the dynamic effects using the approach proposed by Callaway and Sant'Anna (2021) (hereafter referred to as CS estimates). To achieve CS estimates, we first estimate the group-time average treatment effects, $ATT(g,t)$, for reform group 1, $g = 2015m3$, or group 2, $g = 2015m10$, and for each month t. This gives us the average treatment effect at time t for the cohort g. We use the not-yet-treated units as the comparison groups. Thus we cannot estimate $ATT(3,t)$ for the third reform group as there is no control group for it. With parallel trends, $ATT(g,t)$ is identified by comparing the expected changes in outcome for cohort g between the baseline and t to that to that for a control group not-yet-treated at period t. Next, we aggregate the building blocks $ATT(g,t)$ to event study parameter e period after the reform by choosing appropriate weight w_g, that is, $ATT(e) = \sum_g w_g ATT(g, g+e)$. In addition, we can construct the overall average treatment effects using $ATT(g,t)$ (Callaway and Sant'Anna, 2021). The results are presented in Figure 25.

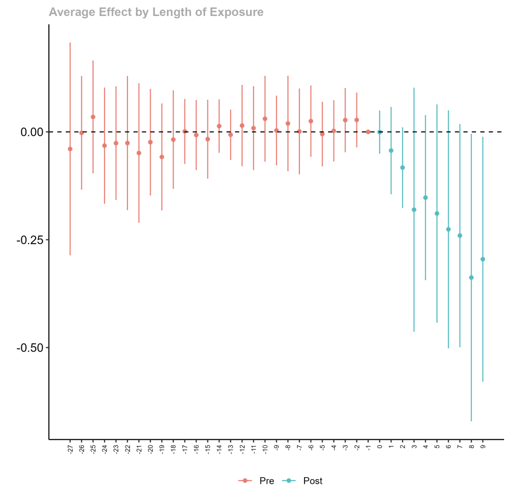
Figure 25: CS DID estimates of claim rate and size



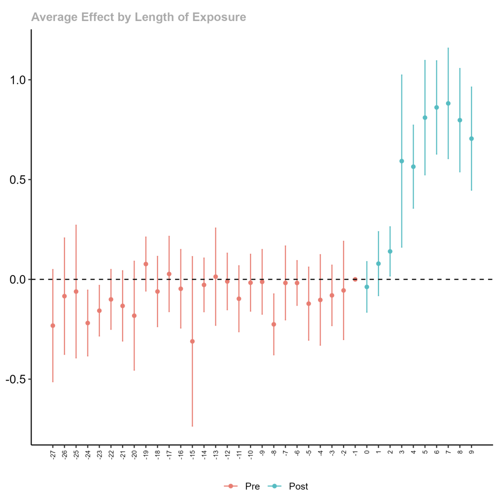
(a) Claim rate of collision damage insurance



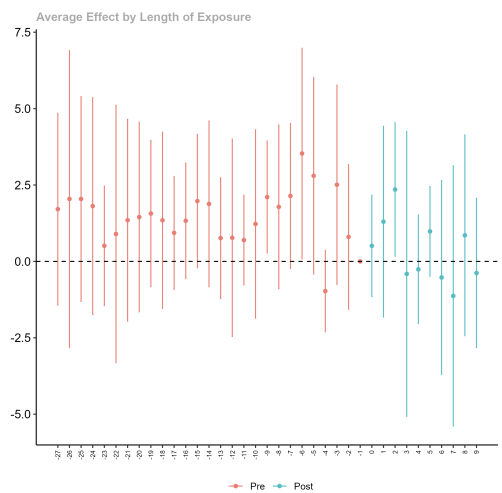
(b) Claim rate of 3rd party liability insurance



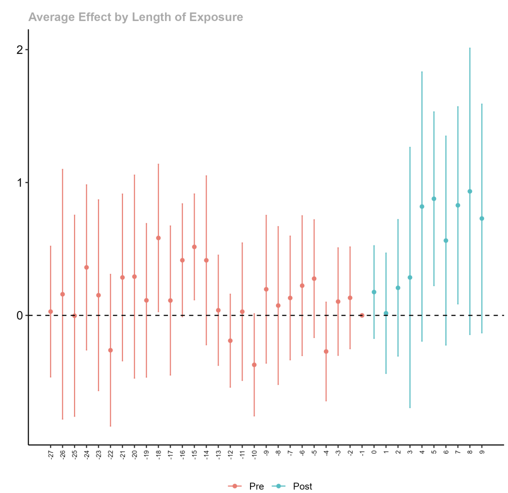
(c) Claim rate of compulsory insurance



(d) Claim size of collision damage insurance



(e) Claim size of 3rd party liability insurance



(f) Claim compulsory insurance

Note: The figures plot aggregated CS estimates.

B.3 Average Treatment Effects using Event Study and TWFE Model

Table 4: Average Treatment Effects using Event Study and TWFE Model

	Claim Rate			Claim Size		
	Compulsory Full Sample	Voluntary Liability Full Sample	Voluntary Collision Always buyers	Compulsory Compulsory	Voluntary Liability Always buyers	Voluntary Collision Always buyers
Panel A: Event Study						
anticipation	-.00045**	-.00042***	-.0061***	89	815***	120***
	(.00021)	(.000076)	(.00084)	(71)	(243)	(30)
post	-.0018***	-.00055***	-.019***	228***	154	902***
	(.00045)	(.000077)	(.00098)	(45)	(196)	(78)
N	73,569,764	73,492,238	49,588,880	501,800	224,341	1,220,846
Pre-Reform Mean Y	0.01	0.00	0.04	3,366.31	7,458.46	2,612.86
Anticipation % of Pre-Reform Mean	-5.73	-11.60	-17.30	2.66	10.92	4.58
Post % of Pre-Reform Mean	-23.33	-15.32	-52.22	6.76	2.07	34.53
Panel B: TWFE						
anticipation	-.000051	-.00013**	-.0029***	44	7.6	127**
	(.0004)	(.000058)	(.00081)	(66)	(254)	(51)
post	-.0011***	-.00037***	-.012***	144	-261	712***
	(.00033)	(.000061)	(.0011)	(105)	(301)	(57)
N	73,569,764	73,492,238	49,588,880	501,800	224,341	1,220,846
Pre-Reform Mean Y	0.01	0.00	0.04	3,366.31	7,458.46	2,612.86
Anticipation % of Pre-Reform Mean	-0.64	-3.61	-8.14	1.32	0.10	4.87
Post % of Pre-Reform Mean	-13.98	-10.30	-34.63	4.28	-3.50	27.24

Note: Panel A reports estimates from Equation B.1, and Panel B reports estimates from Equation B.1, where outcome variable is monthly average claim rate (claims per 100 policies). The observation level is at year-month-province level. Standard errors are clustered at jurisdiction level. $*p < 0.10, **p < 0.05, ***p < 0.01$.

C T period model

Now we extend to T-period model. The consumer chooses p to maximize the utility of T periods:

$$V = \max_p u(w-q) + (T-1)\left[(1-p)u(w-q) + pu(w-q-\delta)\right] - T\Gamma(p) \tag{24}$$

where the equilibrium premium is determined by

$$TpL = Tq + (T-1)p\delta. \tag{25}$$

Thus the base premium is $q = pL - \frac{(T-1)}{T}p\delta$, and change in δ leads to

$$T\frac{qr}{d\delta} = T\frac{dp}{d\delta}L - (T-1)\frac{dp}{d\delta}\delta - (T-1)p. \tag{26}$$

The change in value function with small change in δ is

$$\frac{dV}{d\delta} = -(T-1)pu'(w-q-\delta) + \left[u'(w-q) + (T-1)\left((1-p)u'(w-q) + pu'(w-q-\delta)\right)\right]\left(-\frac{dq}{d\delta}\right) \tag{27}$$

$$= -(T-1)p\left[u' + u''(-\delta)\right] + \left[u' + (T-1)\left[u' + pu''(-\delta)\right]\right]\left(-\frac{dq}{d\delta}\right) \tag{28}$$

$$= u'\left[-(T-1)p - T\frac{dq}{d\delta}\right] + u''\left[(T-1)p\delta + (T-1)p\delta\frac{dq}{d\delta}\right] \tag{29}$$

$$= -u'\left[T\frac{dp}{d\delta}L - (T-1)\frac{dp}{d\delta}\delta\right] + u''(T-1)p\delta\left[1 + \frac{dp}{d\delta}L - \frac{(T-1)}{T}\frac{dp}{d\delta}\delta - \frac{(T-1)}{T}p\right], \tag{30}$$

where the second line is obtained by first order Taylor approximation of u'. Dividing the equation by marginal utility, we have the sufficient formula as

$$\frac{dV}{d\delta}/u'(w-q) = -\left[T\frac{dp}{d\delta}L - (T-1)\frac{dp}{d\delta}\delta\right] - \gamma(T-1)p\delta\left[1 + \frac{dp}{d\delta}L - \frac{(T-1)}{T}\frac{dp}{d\delta}\delta - \frac{(T-1)}{T}p\right]. \tag{31}$$

Remark Note that we obtain per-period value change by dividing $\frac{dV}{d\delta}/u'(w-q)$ by T: $\frac{dV}{d\delta}/u'(w-q)T$. When T is large enough, the per-period value change is the close to that in the static model.

D Model Derivation

D.1 Optimal Driving Behavior in Ex-Ante Model

The optimal p_0^* satisfies the first order condition:

$$\underbrace{-\Gamma'(p^*)}_{\text{MC}} = \underbrace{u(w-q) - u(w-(1+\theta)q)}_{\text{MB}}, \tag{32}$$

and we get $p^* = \Gamma'^{-1}\left(u(w-(1+\theta)q) - u(w-q)\right)$ Using equation 32, we have

$$\frac{dp^*(\theta)}{d\theta} = \frac{-u'(w-(1+\theta)q)q}{\Gamma''(p^*)} < 0, \tag{33}$$

which implies larger penalty decrease p^*, i.e. improving driving behavior. Clearly, the larger $\Gamma''(p^*)$, the more costly it is for drivers to change their behavior, the smaller the change in behavior is.

D.2 Welfare Implications in Ex-Ante Model

Now we consider how consumer welfare changes with an increase in θ. We take derivative of $V(\theta)$ w.r.t. θ:

$$\frac{dV(\theta)}{d\theta} = \underbrace{-p^*u'(w-(1+\theta)q)q}_{\text{cost when claiming}} + \underbrace{\left(u'(w_0-q) + p_0u'(w-(1+\theta)q)(1+\theta) + (1-p_0)u'(w-q)\right)(-q')}_{\text{benefit from decrease in base premium}}. \tag{34}$$

Next, we approximate $u'(w-(1+\theta)q)$ using the first order of Taylor approximation that $u'(w-(1+$

$\theta)q) = u'(w-q) - u''(w-q)\theta q$ and rearrange, we have

$$\frac{dV(q)}{d\theta} = -pq(u'(w-q) - u''(w-q)\theta q) + \tag{35}$$

$$(u'(w_0 - q) + p_0(1+\theta)(u'(w-q) - u''(w-q)\theta q) + (1-p_0)u'(w-q))\,(-q') \tag{36}$$

$$= -pqu'(w-q) + pqu''(w-q)\theta q - p_0(1+\theta)u''(w-q)\theta q(-q') \tag{37}$$

$$+ (u'(w_0-q) + p_0(1+\theta)u'(w-q) + (1-p_0)u'(w-q))\,(-q') \tag{38}$$

$$= \underbrace{u'(w-q)\,(-pq + (2+p_0\theta)\,(-q'))}_{\text{benefits from prm decrease}} + \underbrace{p\theta qu''(w-q)\,(q - (1+\theta)(-q'))}_{\text{risk from higher penalty}} \tag{39}$$

$$= \underbrace{u'(w-q)\left(-\frac{dp_0^*(\theta)}{d\theta}\right)(l-\theta q)}_{\downarrow\text{ risky driving}} + \underbrace{p\theta qu''(w-q)\left(\overbrace{q}^{\uparrow\text{ penalty}} - \overbrace{(1+\theta)(-q')}^{\text{offset by }\downarrow\text{ in prm in the state}}\right)}_{\text{consumption risk from higher penalty}}. \tag{40}$$

D.3 Welfare Implications in Ex-Post Model When No Additional Filing Cost

Now we focus on consumer welfare, and take derivative of $V(\theta)$ w.r.t. θ.

$$\frac{dV(\theta)}{d\theta} = \underbrace{-u'(w-(1+\theta)q)\overline{p}(1-F(x))q}_{\text{cost when claiming}} + \tag{41}$$

$$\underbrace{\left(u'(w-q) + \overline{p}\int_0^x u'(w-l-q)f(l)dl + \overline{p}u'(w-(1+\theta)q)(1-F(x))(1+\theta) + (1-\overline{p})u'(w-q)\right)(-\frac{\partial q}{\partial\theta})}_{\text{benefit from decrease in base premium}}. \tag{42}$$

We approximate $u'(w-c(l)-q)$ and $u'(w-(1+\theta)q)$ using Taylor series. That is,

$$u'(w-c(l)-q) = u'(w-q) - u''(w-q)l, \tag{43}$$

and

$$u'(w-(1+\theta)q) = u'(w-q) - u''(w-q)\theta q. \tag{44}$$

Thus, replacing the Taylor approximations in Equation 41 as well as the premium changes, we have

$$\frac{dV(\theta)}{d\theta} = -\left(u'(w-q) - u''(w-q)\theta q\right)\overline{p}(1-F(x))q + \left(-\frac{\partial q}{\partial \theta}\right) \times \tag{45}$$

$$\Bigg(u'(w-q) + \overline{p}\int_0^x \left(u'(w-q) - u''(w-q)c(l)\right) f(l)dl + \tag{46}$$

$$\overline{p}(u'(w-q) - u''(w-q)\theta q)(1-F(x))(1+\theta) + (1-\overline{p})u'(w-q)\Bigg) \tag{47}$$

$$= -u'(w-q)\overline{p}(1-F(x))q + u''(w-q)\theta q\overline{p}(1-F(x))q + \frac{\overline{p}(1-F(x))q}{2+\overline{p}\theta(1-F(x))}\times \tag{48}$$

$$\Bigg(u'(w-q) + \overline{p}\int_0^x u'(w-q)f(l)dl + \overline{p}u'(w-q)(1-F(x))(1+\theta) + (1-\overline{p})u'(w-q) \tag{49}$$

$$-\overline{p}\int_0^x u''(w-q)c(l)f(l)dl - \overline{p}u''(w-q)\theta q(1-F(x))(1+\theta)\Bigg) \tag{50}$$

$$= -u'(w-q)\overline{p}(1-F(x))q + u''(w-q)\theta q\overline{p}(1-F(x))q + \frac{\overline{p}(1-F(x))q}{2+\overline{p}\theta(1-F(x))}\times \tag{51}$$

$$\Bigg(u'(w-q) + \overline{p}u'(w-q)F(x) + \overline{p}u'(w-q)(1-F(x))(1+\theta) + (1-\overline{p})u'(w-q) \tag{52}$$

$$-\overline{p}u''(w-q)\int_0^x lf(l)dl - \overline{p}u''(w-q)\theta q(1-F(x))(1+\theta)\Bigg) \tag{53}$$

$$= -u'(w-q)\overline{p}(1-F(x))q + u'(w-q)\times \tag{54}$$

$$\frac{\overline{p}(1-F(x))q}{2+\overline{p}\theta(1-F(x))}\times\Bigg(\underbrace{1+\overline{p}F(x)+\overline{p}(1-F(x))(1+\theta)+(1-\overline{p})}_{=2+\overline{p}\theta(1-F(x))}\Bigg)+ \tag{55}$$

$$u''(w-q)\times\frac{\overline{p}(1-F(x))q}{2+\overline{p}\theta(1-F(x))}\times \tag{56}$$

$$\Bigg(\theta q\left(2+\overline{p}\theta(1-F(x))\right) - \overline{p}\underbrace{\int_0^x lf(l)dl}_{=xF(x)-\int_0^x F(l)dl} - \overline{p}\theta q(1-F(x))(1+\theta)\Bigg) \tag{57}$$

$$= \underbrace{u''(w-q)}_{\leq 0}\times\frac{\overline{p}(1-F(x))q}{2+\overline{p}\theta(1-F(x))}\times\underbrace{\left(\theta q(2-\overline{p}) + \int_0^{\theta q} F(l)dl\right)}_{>0} \leq 0. \tag{58}$$

D.4 Welfare Implications in Ex-Post Model With Additional Filing Cost

With assumption that the insured have additional cost or benefit to not file a claim, that is, $c(l) \neq l$, we have $x^* \neq \theta q$. For simplicity, we can assume $c(l) = l - c$. If $c > 0$, it implies that the insured has additional benefit from not filing, say, from no hassle cost need to filing the claim.

To express the changes in premiums w.r.t. θ in terms of observed term, we have

$$\overline{p}\left[\underbrace{\frac{\partial}{\partial\theta}\int_{x^*}^{\infty}c(l)f(l)dl}_{<0;\downarrow\text{expected cost}}+\underbrace{\theta q\frac{\partial F(x^*)}{\partial\theta}}_{>0,\downarrow\text{ penalty income}}\right]\underbrace{-\overline{p}(1-F(x^*))q}_{<0;\downarrow\text{ as higher exp. rated}}=\frac{dq(\theta)}{d\theta}(2+\overline{p}\theta(1-F(x^*))), \tag{59}$$

when there is no cost to file, as the assumption in the above section, we know all the decrease in the expected cost is from the decrease in penalty when the insured choose not to file the claim. Thus we have the first two term equal to zero. If the insured benefits from not filing, we have the first two term as negative, implying the decrease in the premiums is more negative.

Thus the welfare changes is

$$\frac{dV(\theta)}{d\theta}=\underbrace{-u'(w-(1+\theta)q)\overline{p}(1-F(x))q}_{\text{cost when claiming}} \tag{60}$$

$$+\underbrace{\left(u'(w_0-q)+\overline{p}\int_0^x u'(w-c(l)-q)f(l)dl+\overline{p}u'(w_1-(1+\theta)q)(1-F(x))(1+\theta)+(1-\overline{p})u'(w_1-q)\right)(-\frac{\partial q}{\partial\theta})}_{\text{benefit from decrease in base premium}} \tag{61}$$

$$=u'\left[\underbrace{-p(1-F(x^*)q}_{\text{loss from penalty}}+\underbrace{(2+\overline{p}\theta(1-F(x^*)))\left(-\frac{dq(\theta)}{d\theta}\right)}_{\text{benefit from prm}\downarrow}\right]+ \tag{62}$$

$$u''\left[\underbrace{\theta p(1-F(x^*)q}_{\uparrow\text{ risk from penalty}\uparrow}+\underbrace{\frac{dq(\theta)}{d\theta}p\left(\int_0^x c(l)f(l)dl+\theta q(1-F(x))(1+\theta)\right)}_{\downarrow\text{ risk from prm}\downarrow}\right] \tag{63}$$

$$=-u'\overline{p}\left[\underbrace{\frac{\partial}{\partial\theta}\int_{x^*}^{\infty}c(l)f(l)dl}_{<0;\downarrow\text{expected cost}}+\underbrace{\theta q\frac{\partial F(x^*)}{\partial\theta}}_{>0,\downarrow\text{ penalty income}}\right]+ \tag{64}$$

$$u''\left[\underbrace{\theta p(1-F(x^*)q}_{\uparrow\text{ risk from penalty}\uparrow}+\underbrace{\frac{dq(\theta)}{d\theta}p\left(\int_0^x c(l)f(l)dl+\theta q(1-F(x))(1+\theta)\right)}_{\downarrow\text{ risk from prm}\downarrow}\right]. \tag{65}$$

E Details of Sufficient Statistic

E.1 Observable Parts

Several components in Equation 18 can be observed from data directly. We will discuss each part separately.

Claim rate $p_r(1-F(x_r))$. $p_r(1-F(x_r)) \equiv m_r$ represents claim rate of risk class r, which can be directly observed in our claim data, that is, the ratio of number of claims to number of policies. However, terms, p_r and $(1-F(x_r))$, cannot be directly observed, where p_r represents the probability of an accident, and $(1-F(x_r))$ represents the conditional probability of claiming given an accident occurs. We use the pre-reform monthly claim rate of risk class r multiplied by 12 as the estimate for m_r.

Base premiums and experience rating factors. Base premium and experience rating factors can also be observed in our data. However, the base premium in the data is defined as the premium in the risk class of one claim in the past year, denoted as $\tilde{q}$. Premium in the case of zero claim in the past year is a percentage, denoted as ϕ_r , of $\tilde{q}$. Note that the discount factor ϕ_r depends on the current risk class r. To match our model, the base premium q_r in the case of no claim and experience rating factor θ_r in our model is re-normalized as $q_r = \phi_r \tilde{q}$, and $\theta_r = \frac{1-\phi_r}{\phi_r}$. When we consider a 1 p.p. increase in θ_r for each risk classes, this would be induced by ϕ_r^2 p.p. decreases in ϕ_r, as implied by $d\phi_r = -\phi_r^2 d\theta_r$.

One additional issue is that there are multiple coverage levels for different types of coverage. For simplicity, we set $\tilde{q}$ at the average of base premiums across all individuals and periods, which is 2,718 yuan. This is relatively good approximation as most parts of equations are linear in the premium.

E.2 Bounding $p_r^*(\theta) \int_0^{x_r^*} l dF(l)$

$p_r^* \int_0^{x_r^*} l dF(l)$ measures the average unclaimed cost, which is the total unclaimed loss sizes over the effective number of policy. This unclaimed cost is unobserved in the claim data, but we can bound it.

One one hand, the smallest possible value of $p_r^* \int_0^{x_r^*} l dF(l)$ is zero if we assume that no claims are underreported. On the other hand, the largest possible value is $p_r^* F(x_r^*) x_r^*$, though all three components are unobserved. A practical approach is by $p_r^* \int_0^{x_r^*} l dF(l) \leq p_r^* \int_{x_r^*}^{\infty} l dF(l)$, which is the average cost, defined as the total claim amount over number of effective policy, denoted as $\bar{x}_r$.

To summarize, the bound can be illustrated as

$$0 \leq p_r^* \int_0^{x_r^*} l dF(l) \leq \bar{x}_r, \tag{66}$$

Combining with $p^*(\theta)(1-F(x^*))(1+\theta)\theta q(\theta)$, which can be quantified directly, we have

$$\underbrace{p^*(1-F(x^*))(1+\theta)\theta q(\theta)}_{\equiv MLB} \leq \underbrace{p^* \int_0^{x^*} l f(l) + p^*(1-F(x^*))(1+\theta)\theta q(\theta)}_{\equiv M} \leq \underbrace{\bar{x} + p^*(1-F(x^*))(1+\theta)\theta q(\theta)}_{\equiv MUB}. \tag{67}$$

E.3 Estimating the Equilibrium Changes in Premium $\frac{dq(\theta)}{d\theta}$

We quantify changes in equilibrium premiums with actual increase in θ, i.e. Δθ. We start with single risk class as a demonstration, and extend to multiple classes as in practice. It takes two steps to get $\frac{dq_r}{d\theta_r}$. First, we use the equilibrium condition to estimate $\frac{d\tilde{q}}{d\phi_r}$, that is, how base premium of one claim $\tilde{q}$ changes with discount factor ϕ_r; Second, we obtain $\frac{dq_r}{d\theta_r}$ from $\frac{d\tilde{q}}{d\phi_r}$ using $\frac{dq_r}{d\theta_r} = \frac{dq_r}{d\phi_r}\frac{d\phi_r}{d\theta_r} = \left[\tilde{q} + \frac{d\tilde{q}}{d\phi_r}\right] \times (-\phi_r^2)$.

Single Risk Class. To start, we consider a single risk class or setting premium separately without pooling for each risk class, which helps us illustrate the primary components to estimate the premium changes. First, we use the actual base premium $\tilde{q}$ and discount factor ϕ_r to calculate changes in premiums. The equilibrium condition can be rewritten as:

$$p_r^* \int_{x_r^*}^{\infty} l dF(l) = p_r^*(1-F(x_r^*))\tilde{q} + p_r^* F(x_r^*)\phi_r \tilde{q} + (1-p_r^*)\phi_{1r}\tilde{q}, \tag{68}$$

which leads to changes in base premium as

$$\frac{d\tilde{q}}{d\phi_r} = \frac{1}{((p_r^*(1-F(x_r^*)) + \phi_r(1-p_r^*(1-F(x_r^*)))))} \times \tag{69}$$

$$\left(\underbrace{\frac{d}{d\phi_r}\left[p_r^* \int_{x_r^*}^{\infty} l dF(l)\right]}_{\text{changes in expected cost}} - \underbrace{\frac{d}{d\phi_r}\left[p_r^*(1-F(x_r^*))\right]}_{\text{changes in claim rate}}(1-\phi_r)\tilde{q} - \underbrace{(1-p_r^*(1-F(x_r^*)))}_{\text{mechanical; observed}}\tilde{q} \right) \tag{70}$$

- $\frac{d}{d\phi_r}\left[p_r^* \int_{x_r^*}^{\infty} l dF(l)\right]$ is the changes in the average cost. As vehicles can have more than one accidents, the incentives of the second accident/claim is different from the first one. More importantly, with multiple possible accidents, it is difficult to set number of zero dollar claims. Therefore, we only consider the first claim in our setting. Therefore, we can set the unclaimed vehicles with claims with zero.

- $-\frac{d}{d\phi}\left[p_r^*(1-F(x_r^*))\right]$ is the changes in (first) claim rate.

To estimate the average treatment effect on claim rate and average cost due to reform, we use the TWFE estimating equation:

$$y_{rt}^{\text{type}} = \alpha\ \text{anticipation}_{rt} + \beta \text{post}_{rt} + \mu_r + \tau_t + \varepsilon_{rt}, \tag{71}$$

where in the first case, y_{rt}^{type} is average cost, and in the second case, it is claim rate. The estimated parameters of interest are denoted as $\hat{\beta}^{\text{size}}$ and $\hat{\beta}^{\text{rate}}$, representing effects of 1 p.p. decrease in ϕ_r. Plugging the estimates, we have $\Delta\tilde{q}_r = \frac{\hat{\beta}^{\text{cost}} - \hat{\beta}^{\text{rate}}(1-\phi_r)\tilde{q} - (1-m_r)\tilde{q}\times\Delta\phi_r}{m_r + \phi_r(1-m_r)}$.

Second, in order to obtain $\frac{dq_r}{d\theta_r}$, we use $\frac{dq_r}{d\theta_r} = \frac{dq_r}{d\phi_r}\frac{d\phi_r}{d\theta_r} = \left[\tilde{q} + \frac{d\tilde{q}}{d\phi_r}\right] \times (-\phi_r^2)$. Therefore, an change in θ_r by $\Delta\theta$, induced by a $\Delta\phi$ change in ϕ leads to premium change $\Delta q_r = \left[\phi_r^2\tilde{q}\Delta\theta + \Delta\tilde{q}_r\right] = \tilde{q}\Delta\phi_r + \Delta\tilde{q}_r$.

Multiple Risk Classes. Now we demonstrate how base premiums changes with multiple risk classes and changes in discount factors of all risk classes. Suppose that each risk class has shares ρ_r and, with subscript to denote the risk class r. As in the single risk class case, we still use two steps to obtain $\frac{dq_r}{d\theta_r}$ by estimating $\frac{d\tilde{q}}{d\phi_r}$ first and then moving to $\frac{dq_r}{d\theta_r}$. The only additional complication is from changes in discount factors of all risk classes in the first step.

First, we show $\frac{dq_r}{d\theta_r}$ with multiple risk classes. To start, we write the equilibrium premium condition as:

$$\sum_{r=1}^{R} \rho_r \left(p_r^* \int_{x_r^*}^{\infty} l dF(l) \right) = \sum_{r=1}^{R} \rho_r \Big(p_r^*(1-F(x_r^*))\tilde{q} + p_r^* F(x_r^*)\phi_r\tilde{q} + (1-p_r^*)\phi_r\tilde{q} \Big), \tag{72}$$

where ρ_r represents the share of insureds in risk class r, and R represents the total number of risk classes. The change in base premium with changes in all discount factors can obtained through total differentials of

$\tilde{q}$ w.r.t. ϕ_r:

$$d\tilde{q}\left[\sum_{r=1}^{R}\rho_r\left(\left(p_r^*(1-F(x_r^*))+\phi_r(1-p_r^*(1-F(x_r^*)))\right)\right)\right]= \tag{73}$$

$$\sum_{r=1}^{R}\rho_r\left(\frac{d}{d\phi_r}\left[p_r^*\int_{x_r^*}^{\infty} ldF(l)\right]-\frac{d}{d\phi_r}\left[p_r^*(1-F(x_r^*))\right](1-\phi_r)\tilde{q}-(1-p_r^*(1-F(x_r^*)))\,\tilde{q}\right)d\phi_r. \tag{74}$$

If all experience rating factors θ_r change by $d\theta_r$, the change in actual base premium $\tilde{q}$ can be expressed as

$$d\tilde{q}=\sum_{r=1}^{R}\omega_r\frac{d\tilde{q}}{d\phi_r}\frac{d\phi_r}{d\theta_r}d\theta_r, \tag{75}$$

where $\omega_r=\frac{\rho_r(\phi_r(p_r^*(1-F(x_r^*))+p_r^*(1-F(x_r^*))))}{\sum_{r'}\rho_{r'}\left(\phi_{r'}\left(p_{r'}^*(1-F(x_{r'}^*))+p_{r'}^*(1-F(x_{r'}^*))\right)\right)}$ and $\frac{d\tilde{q}}{d\phi_r}$ is from Equation 69. As such, to compute the change in original base premium $\tilde{q}$ for risk class r with $\Delta\theta_r$, we use

$$\Delta\tilde{q}=\sum_{r=1}^{R}\omega_r\Delta\tilde{q}_r, \tag{76}$$

where the change in the re-defined base premium for risk class r, i.e. $\Delta\tilde{q}_r$, is

$$\Delta\tilde{q}_r=\frac{\hat{\beta}_r^{\text{cost}}-\hat{\beta}_r^{\text{rate}}(1-\phi_r)\tilde{q}-(1-m_r)\tilde{q}\times\Delta\phi}{m_r+\phi_r(1-m_r)}. \tag{77}$$

Therefore, the changes in the base premium q_r is as follows:

$$dq_r=\tilde{q}\frac{d\phi_r}{d\theta_r}d\theta_r+\phi_r d\tilde{q}=-\phi_r^2\tilde{q}d\theta_r+\phi_r d\tilde{q}=-\phi_r^2\tilde{q}d\theta_r+\phi_r\left[\sum_{r'=1}^{R}\omega_{r'}\frac{d\tilde{q}}{d\phi_{r'}}\frac{d\phi_{r'}}{d\theta_{r'}}d\theta_{r'}\right]. \tag{78}$$

To compute the changes in base premiums q_r with $\Delta\theta_r$, we use

$$\Delta q_r=\tilde{q}\Delta\phi_r+\phi_r\Delta\tilde{q}_r. \tag{79}$$

E.4 Combining All Components

We have quantified or bounded each part in Equation 18. Combining all parts for each risk class, we have

$$\Delta V_r = \left(-m_r q_r \times \Delta\theta_r + (-\Delta q_r)\left(1 + m_r \theta_r\right)\right) - \gamma\left(m_r \theta_r q_r^2 \times \Delta\theta_r - \Delta q_r M_r\right). \tag{80}$$

This provides us with a (bounded) quantification of welfare changes with 1 p.p. increase in θ.

F Timeline of the Reform