Institute of Insurance Economics

University of St.Gallen

BETWEEN-GROUP ADVERSE SELECTION: EVIDENCE FROM GROUP CRITICAL ILLNESS INSURANCE

MARTIN ELING
RUO JIA
YI YAO

WORKING PAPERS ON RISK MANAGEMENT AND INSURANCE NO. 141

EDITED BY HATO SCHMEISER

CHAIR FOR RISK MANAGEMENT AND INSURANCE

OCTOBER 2014

Between-Group Adverse Selection:
Evidence from Group Critical Illness Insurance

Martin Eling, Ruo Jia, Yi Yao

Abstract

This paper demonstrates the presence of adverse selection in the group insurance market. Conventional wisdom suggests that group insurance mitigates adverse selection because it minimizes individual choice. We complement this conventional wisdom by analyzing a group insurance scenario in which individual choice is excluded, and we find that group insurance alone is not effective enough to eliminate adverse selection, i.e., between-group adverse selection exists. Between-group adverse selection, however, disappears over time if the group renews with the same insurer for a certain period. Our results thus indicate that experience rating and underwriting based on information that insurers learn over time are important in addressing adverse selection.

Keywords

Adverse Selection, Information Asymmetry, Learning over Time, Group Insurance, Health Insurance

----

Martin Eling and Ruo Jia are affiliated with the University of St. Gallen. Yi Yao is affiliated with Peking University. The authors can be reached via email: martin.eling@unisg.ch, ruo.jia@unisg.ch, and yao.yi@pku.edu.cn.

Introduction

Group insurance constitutes a substantial portion of global insurance markets, and its importance to life and health insurance is increasing.[1] Administrative efficiency and low performance volatility are strong motivations for the insurance industry to develop group insurance products (Bickelhaupt, 1983). Many policyholders also favor group insurance because it allows them to avoid the difficulty and anxiety of shopping for insurance (Pauly and Percy, 2000).

Consistent with conventional wisdom and widely accepted industry practice, insurers use group insurance to mitigate adverse selection because (1) the mixture of high- and low-risks decreases the variance of grouped losses compared to individual losses, (2) individual choice is minimized, and (3) individuals do not act strategically on information advantage regarding risk types when group insurance is tangential to other factors that influence an employment decision (Mayers and Smith, 1981; Browne, 1992). Our paper defines a group as a set of five or more people formed for reasons other than purchasing insurance, e.g., firms, unions, families, or other social groups.[2] Thus, between-group adverse selection refers to the adverse selection in the group insurance market, which results from group strategic action on behalf of collective welfare, rather than from individual choice within a group (Mayers and Smith, 1981; Hanson, 2005).

Mayers and Smith (1981) predict that the group insurance market should have no adverse selection if the group insurance does not allow for individual choices within a group and if the group is formed for purposes other than purchasing insurance. Browne (1992) uses the group insurance market as the benchmark market free of adverse selection and concludes that individual insurance suffers more from adverse selection than group insurance. In contrast to Mayers and Smith's (1981) prediction, Hanson (2005) proves that the equilibrium in the group insurance market with no individual choice is not materially different from the equilibrium in the individual insurance market. Hanson's (2005) model implies that groups, like individuals, act strategically on their information advantages, which yields between-group adverse

---

1 In 2012, the direct written premiums of group insurance in U.S. totaled USD 295 billion, accounting for 41.9% of total premiums in the life and health sector; in particular, group insurance dominates the U.S. health insurance market, accounting for 53.8% of total health premiums (Insurance Information Institute, 2013). In Europe, group insurance accounted for 36% of total premiums in life insurance in 2012, whereas the percentage was 29% in 2010 and 31% in 2008 (Insurance Europe, 2014).

2 We use two datasets to test the robustness of our conclusions on different types of groups. Our core model uses employment-based groups, such as firms and unions; in a robustness test, we explore another dataset of family-based groups.

selection. The two competing theoretical predictions motivate us to empirically test for the existence of and, if found, the persistence of adverse selection in a group insurance market.

The existing empirical evidence on adverse selection in group insurance concentrates on the U.S. health insurance market, where individual choice among competing health insurance plans is an important driver of adverse selection (see, e.g., Cutler and Zeckhauser, 2000; Handel, 2013). Simon (2005) analyzes adverse selection in the U.S. small-group health insurance market, which combines the effects of individual choice and group selection. To our knowledge, no empirical research has tested for the existence and/or persistence of adverse selection in a group insurance market in which no individual choice is allowed.

Following Mayers and Smith (1981), we split the individual choice and group decision effects in group adverse selection by differentiating within- and between-group adverse selection. We use a new and comprehensive dataset of group critical illness (CI) insurance policies.[3] Our findings show that group insurance cannot eliminate adverse selection, even if individuals within a group are not allowed to choose their participation and/or coverage, supporting Hanson's (2005) prediction. The information advantage of the group over the insurer may come from the underreporting of claim histories when the group is a new customer.

We also find evidence that between-group adverse selection, together with the group's information advantage, disappears over time if a group renews with the same insurer for a certain period. We attribute this disappearance to insurer learning over time (Kunreuther and Pauly, 1985; D'Arcy and Doherty, 1990; Hendel and Lizzeri, 2003; De Garidel-Thoron, 2005). Once the insurer has experience with the group, it will learn the risk type and then either adjust the premium or elect not to renew the high-risk group based on each group's claim experience. Therefore, the information advantage that the group possesses as a new client diminishes with the renewal process. The asymmetric information on which the group can act strategically disappears and thus no between-group adverse selection occurs. Cohen (2012) documents the evidence of an insurer learning over time in an Israeli automobile insurance portfolio. Cutler (1994) documents the wide existence of experience rating in the U.S. small-group health insurance market. He shows that the variation in small-group insurance premiums results not from demographic or

[3] CI insurance covers the 25 critical diseases listed in Appendix 1 and pays the insurance amount if any of these diseases is first diagnosed during the policy period, following a predefined waiting period for first-time purchasers.

benefit differences but from experience ratings. This paper complements the extant empirical evidence by demonstrating the disappearance of between-group adverse selection over time.

The remainder of this paper is structured as follows. In Section 2, we introduce the theoretical background and derive our hypotheses. Section 3 presents the data and some background information regarding the CI insurance product and the Chinese insurance market. Section 4 explains our empirical models. Section 5 discusses the results. Section 6 presents several robustness tests. Sections 7 and 8 provide the discussion and concluding remarks, respectively.

Theoretical Background and Hypotheses

Adverse selection is the tendency of high-risks to purchase insurance or to purchase more insurance coverage than do low-risks (Cummins, et al., 1983). Adverse selection results from asymmetric information that favors the insurance buyer over the insurer (Akerlof, 1970; Rothshild and Stiglitz, 1976). The existence of adverse selection has been widely documented in many types of individual insurance (see Cohen and Siegelman, 2010, for a review).

Regarding adverse selection in the group insurance market, Cutler and Zeckhauser (2000) review 14 empirical studies that examine the selection of group health insurance with individual choices, all of which find some type of adverse selection; however, none of these studies distinguish within from between-group adverse selection. Simon's (2005) study of the U.S. small-group health insurance market reveals the relationship between individual choice and within-group adverse selection. However, whether between-group adverse selection exists and, if so, the extent of its persistence is yet unknown. We aim to fill this gap. Table 1 summarizes the theoretical framework, empirical literature, and contribution of this paper.

Table 1 Three Types of Adverse Selection

Types of adverse selection	Insurance decisions on participation and coverage	Literature providing empirical evidence
Type I: Individual adverse selection	Individual	See Cohen and Siegelman (2010) for a review of empirical studies that focus on individual insurance markets.
Type II: Within-group adverse selection	Individual and group: The participation of group members is voluntary, and/or coverage choice is allowed within a group.	See Cutler and Zeckhauser (2000) for a review of empirical studies that focus on the U.S. group health insurance market.
Type III: Between-group adverse selection	Group: The participation of group members is mandatory, and the coverage is identical within a group.	The aim of this paper, not covered by the existing literature

Two competing theoretical predictions were developed regarding between-group adverse selection, as mentioned in the Introduction. Mayers and Smith (1981) predict that if a group is formed for purposes other than purchasing insurance, the average risk for that group is less likely to deviate from the relevant population average, which solves between-group adverse selection problems. However, Hanson's (2005) theoretical model yields the opposite prediction. Hanson (2005) compares the equilibria in an individual insurance market with that in a group insurance market with no individual choice and concludes that a profit-maximizing employer[4] will choose contracts off the same equilibrium contract curve as would a purchaser of individual insurance, suggesting that group insurance cannot eliminate adverse selection. This conclusion is subject to the conditions that (1) the group insured[5] makes all decisions on behalf of group members, (2) it is a uniform group insurance policy, (3) there is no wealth effect, (4) there are no administrative costs, and (5) the group is formed for purposes other than purchasing insurance. In Hanson's (2005) model, although the pooling of high- and low-risk members reduces the variance of group losses, this effect turns out to be irrelevant to the equilibrium. Between-group adverse selection is independent of the individual choices within a group. These two competing predictions yield our first hypothesis.

Hypothesis I: If group insurance does not allow for individual choices within a group, and if the group is formed for purposes other than purchasing insurance, group insurance eliminates adverse selection.

Under the two conditions of *Hypothesis I*, two explanations suggest how and why adverse selection may still exist. Mayers and Smith (1981) argue that individual insured may switch employers due to differences in health insurance coverage among similar jobs; therefore, high-risk individuals choose jobs that offer more comprehensive health insurance coverage. This explanation is unlikely to apply to our case because the expected benefit of group CI insurance is small. On average, 95% of individuals have an expected benefit[6] of equal to or less than CNY 109 (USD 17) per year, which does not suggest a strong motivation for employees to switch jobs. Hanson (2005) provides another explanation for the existence of between-group adverse selection: a profit-maximizing employer behaves similarly to individual insurance buyers when choosing insurance coverage. The employer will choose

[4] A profit-maximizing employer is defined as an employer who wishes to maximize the sum of the benefits its employees receive from group insurance minus the price paid for the insurance.

[5] A group insured is an employer or other type of group purchasing coverage, e.g., a union.

[6] The expected benefit is calculated as the insurance amount multiplied by the average claim frequency (see summary statistics in Table 2).

coverage that is more comprehensive if its employees are high-risk and vice versa. This explanation requires that the employer, specifically the human resources (HR) or other responsible department, possesses an information advantage over the insurer; otherwise, the insurer would be able to charge a risk-adequate rate to eliminate adverse selection.

Under the condition that the group insured makes all insurance decisions, there are two potential sources of group insured's information advantage. Cohen (2005, 2012) empirically shows that the insurer is not able to fully observe a new customer's past claim records; thus, cannot accurately assess the risk type of a new customer because self-reporting of past claims is incomplete or inaccurate (Insurance Research Council, 1991). In other words, a client is most likely to have an information advantage when he/she is a new customer to an insurer (Cohen, 2005). This argument applies to both individual and group insurance. In our sample, most group insureds are established firms that can afford voluntary employee benefits. Their HR departments know more about their employees than does the insurer, including knowledge of prior health claims and critical diseases. Thus, we will use the subsample of new policies to test *Hypothesis I,* the existence of between-group adverse selection.

Hanson (2005) offers another explanation for group insured's information advantage; that is, the group insured usually knows the health condition of its employees at the individual level, but the insurer knows only the group average. In reality, the HR department usually possesses individual-level health information that the insurer does not, such as whether an employee smokes or has a history of serious diseases (e.g., heart attack). In some circumstances, the employer is aware that some of its employees have symptoms of critical diseases, and it purchases insurance based on such information (Monheit and Schone, 2004), particularly for small groups. Employers obtain their information advantage via daily interactions with employees and/or via employee health examination results consolidated by HR departments, if permitted by regulations. As many U.S. small-group health insurance studies show, small groups are expected either to have more information advantages or to make better use of those advantages than large groups do; thus, small groups will exhibit stronger between-group adverse selection (Cultler, 1994; Monheit and Schone, 2004).

Kunreuther and Pauly (1985) and Watt and Vazquez (1997) emphasize that observing the realization of a policyholder's risk during a given period enables an insurer to update its prior beliefs concerning the risk posed by that policyholder in a future policy period. Kunreuther and Pauly (1985) and Cohen and Siegelman (2010) call this phenomenon "learning over time." Jean-Baptiste and Santomero (2000) construct a model of the reinsurance market, supported by evidence in Garven, Hilliard and

Grace (2014), showing that asymmetric information between insurers and reinsurers declines over time with the tenure of the insurer-reinsurer relationship.

Insurer learning over time occurs when the insurer is able to use the observed respective insured's claim experience to adjust the premium or reject the renewal. However, this necessary condition does not apply to many individual insurance products. For instance, individual health insurance usually includes an insurer's commitment as a guaranteed renewable clause, which prevents a premium increase and renewal rejection based on an individual's past claim experience with the insurer. Individual life insurance usually involves a long-term commitment from insurer, and often, policy termination by the insurer or premium rate adjustment is not allowed during the entire policy period. Thus, insurer learning over time is limited by various contractual, regulatory, and/or market conditions in individual insurance.

In contrast, an insurer that offers group insurance, of almost all lines, is free to adjust the premium rate upon renewal or to reject renewals based on each group's past claim experience. Therefore, after a few policy periods with the same insurer, the group insured's information advantage from the initial claim underreporting disappears. The group insured's information advantage regarding its employees' individual risk type (Hanson, 2005) also fails to persist because high-risk individuals reveal their risk type over time by making claims. The insurer can thus identify high-risk groups and adjust the premium of (or elect not to renew) high-risk groups based on their claim experience. In other words, the group-average risk type, which the insurer learns over time, is enough to accurately assess group risk.[7] Mayers and Smith (1981) summarize that frequent contract renegotiation controls adverse selection as long as information is revealed over time and the insurer is able to monitor and apply that information accordingly in pricing and underwriting.

Because information asymmetry is a necessary condition for the existence and persistence of adverse selection (Akerlof, 1970; Rothshild and Stiglitz, 1976), adverse selection should cease to be a problem as soon as the insurance buyer's information advantage over the insurer disappears. We phrase our second hypothesis as follows.

Hypothesis II: Between-group adverse selection in the group insurance market, if there is any, will disappear if the group renews with the same insurer for a few policy periods.

There is a substantial body of empirical literature that shows informational asymmetries and adverse selection decrease over time due to repeated contracting

[7] This argument is subject to the mandatory participation requirement, i.e., low-risk individuals cannot exit the group simply because the premium rises while other members stay. We discuss this issue in the Discussion section.

(see, e.g., Cohen, 2012; Garven, Hilliard and Grace, 2014). This paper further documents such a process for between-group adverse selection.

Data and Insurance Product

CI insurance is a type of loss-occurrence health insurance that was offered for the first time in 1983 (Barnard, 2004). The full insurance amount is paid as long as an insurer-recognized hospital provides the first-time diagnosis of the covered disease during the policy period. Usually, there is a 30- to 90-day waiting period for first-time purchasers. The claim benefit always equals the insurance amount and is paid to the insured in a lump sum without additional benefits, such as medical service. The claim payment does not require an actual medical expenditure or hospitalization. Thus CI insurance is immunized from many common problems observed in medical expense health insurance, such as choices between private and public hospitals. The product works as a finance tool rather than a cost reimbursement tool. Cochrane (1995) proposed a time-consistent health insurance plan that provides a lump sum payment and enables those with long-term critical diseases to afford future health insurance coverage. This solution supports the provision of CI insurance as extra funding for insureds with critical diseases.

We obtain the dataset from a life and health insurance company in China. The company has over 15 years of nationwide operations, with a broad spatial range that covers over 90% of the Chinese population. It has ranked among the top 10 largest life insurers in China over the past 15 years in terms of premium market share and assets. Its core business comes from the open market and thus is not concentrated in any particular industry or region. Its operational model, growth path, risk portfolio, and performance are representative in the Chinese insurance market. In 2012, 68 life and health insurers and 62 property and liability insurers operated in the Chinese insurance market, and most of them are legally eligible to issue group CI insurance policies, yielding a very competitive market.

In 2007, the Insurance Association of China and the Chinese Medical Doctor Association issued guidelines that define 25 types of critical diseases. In our case, and in most cases in the Chinese CI insurance market, the insurer strictly follows the CI coverage guideline, which standardizes CI insurance products. In our sample, all group policies and insured individuals have the same coverage for the 25 critical diseases listed in Appendix 1. In 2012, the total premiums written for CI insurance was CNY 40.6 billion (USD 6.5 billion) in China, accounting for 38% of health insurance premiums and covering more than 90 million people (China Insurance Regulatory Committee, 2013; Su, 2013). Both group and individual CI insurance are available in the Chinese market. The group CI insurance market is dominated by employee benefits for which the employer pays the premium and the employee contribution is minimal. The Chinese group CI insurance market has no restrictions regarding risk classification based on age, gender, occupation, region or other possible pricing factors. The insurer has sole discretion to determine the price offered for both new and renewed contracts. The market is commercial and voluntary; thus, the concerns regarding risk reclassification, availability, and affordability of such insurance are minimal.

Our dataset includes all information that the insurer uses to make underwriting and pricing decisions, which minimizes the possibility of spurious adverse selection due to information asymmetry between insurer and researcher (Cohen and Siegelman, 2010). Claims records are also included. The dataset covers all group CI policies issued between January 2008 and June 2013 and all claims settled between January 2008 and August 2012 under the corresponding group CI policies.[8] The business nature of the insurance portfolio is largely, but not restricted to, employee benefits.

The original data are at individual policy level. For each individual policy entry, the dataset provides (1) policy information, including individual policy number, group policy number, insurance amount, premium, policy inception date, expiration date, and issuance date, (2) individual demographic information, including age, gender, and occupation category[9], and (3) group insured's demographic information, including group name, group location, and group size. The dataset also contains claim amounts and claim settlement dates for individual policies with claims.[10]

We organize the individual policy entries into group policies according to the group policy number. Because this paper focuses on between-group adverse selection, we select only those group policies with identical insurance coverage, i.e., identical insurance amounts[11], for each individual insured in the group. Group policies with

---

8 The claims information is electronically recorded in real time but only retrieved and organized by the actuarial team once per year. When we obtained our data, the claim information for September 2012 to June 2013 were not yet available. In a later analysis, to avoid a potential truncation problem, we code the claim status of polices expiring after August 2012 as missing values; thus, these observations are excluded from our regressions.

9 The occupation category is based on the accident tendency of each occupation instead of the illness tendency. We acknowledge that an industry classification reflecting the illness tendency would be a better indicator; however, such an indicator is available neither to the insurer nor to us.

10 We do not have the information of rejected claims, which is a common issue in empirical research that uses real market data (Cohen and Siegelman, 2010). However, because we care about the actual risk type of insureds, it is reasonable to assume that rejected claims fall outside the policy coverage and thus are irrelevant to the risk type and to our conclusions regarding adverse selection.

11 In group CI insurance, the only possible coverage difference within a group policy is the insurance amount. This product involves no deductible and all insureds are covered for the same 25 critical diseases.

identical coverage, i.e., with no individual choice, account for more than 75% of the portfolio. This produces 7,784 group policy-year observations purchased by 3,453 groups[12], representing more than 2,230,000 individual policies. Missing information is present in our dataset, particularly related to missing claims after August 2012 and missing renewal status for some policies in 2008. Our final sample is thus reduced to 3,540 group policy-year observations purchased by 1,957 groups after excluding observations with missing values.[13] As shown in Table 2, the portfolio is characterized by a low claim frequency, a relatively small insurance amount for most policies, and a mixture of different group sizes, occupations, ages, and genders.

Table 3 compares new policies with renewed policies; it also shows subsamples of policies with two or more consecutive renewals and with three or more consecutive renewals. ANOVA mean difference F-tests are reported. We observe no significant difference in claim frequency among subsamples, implying that the risk quality does not materially change over renewal times. Thus, we expect that our conclusions will not be influenced by inherent risk quality difference among subsamples. We find, however, significant trends in most demographic variables and policy features with an increase in renewal times. The portfolio with more renewal times contains groups with larger size, in richer region, with younger members, with more accident-prone occupations, and with more women. The portfolio with more renewal times has smaller average insurance amounts per person, lower premium rates, and shorter policy durations. The policy duration decreases with renewal times because short-term policies renew more times than long-term policies during a certain observation period. The observed trends highlight the importance of controlling for policy and demographic features or premium rates[14] in later regression analyses.

Our CI insurance portfolio is a good approximation of the mandatory participation of group members. According to the insurer's underwriting guidelines, for small groups of no more than 50 people, the participation ratio must be 100% to issue a group policy. For larger groups, the participation ratio can be reduced to a minimum

[12] All groups are recognized as independent entities and assigned a unique group reference code by the insurer. The majority of groups in our sample are independent firms. The rest include subsidiaries, branches or other types of operating units but purchase insurance policies independently. We follow the insurer's practice in treating all entities as independent decision-making units because they buy group policies separately from their affiliates.

[13] We also exclude policies that are rejoined, i.e., the group insured came back to the insurer after a gap period (496 observations).

[14] The premium rate is defined as the policy premium divided by the insurance amount.

of 75%[15]. Employee family participation is sometimes offered and voluntary; however, because the employer pays the premiums, such participation is also very high if the contract is open to family members. The insurer's underwriting guidelines forbid offering group insurance policies to groups formed for the purpose of purchasing insurance. We expect that groups in our sample pool different types of risks and different levels of risk attitudes. Groups may vary in levels of risk aversion; however, we find no reason to sustain a systemic pattern. As mentioned above, our sample excludes the possibility of within-group adverse selection by ensuring identical coverage for every individual insured in a group, thus any group adverse selection identified is between-group adverse selection.

[15] It is unlikely that large firms use this advantage to exclude high-risk members from group coverage to keep premiums low. Group CI coverage is provided as an employee benefit voluntarily offered by the employer, and it is hard for an employer to justify internally offering such benefits to low-risk employees but not to high-risk employees who actually need it more. The employer can stop providing such benefits to all employees at any time if the premium is too high. The voluntary and employee benefit nature of group CI product also minimize the incentive of claim underreporting, aiming at avoiding premium increases.

Table 2 Summary Statistics: Insurance Portfolio Overview

Variables	Descriptions	Obs.	Mean	Min	p5	Median	p95	Max
ClaimDummy	1 if any claim(s) under the group policy	3,540	0.086	0	0	0	1	1
ClaimCount	Number of claims under the group policy	3,540	0.17	0	0	0	1	60 [d]
ClaimFrequency	Average number of claims per insured, i.e., the fraction of insureds making a claim	3,540	0.00064	0	0	0	0.0026	0.11
InsuranceAmount	Insurance amount per insured in CNY	3,540	62,231.4	1,900	3,000	50,000	170,000	1,000,000
PolicyDuration	Group policy duration in days	3,540	284.2	15	30	362	366	485 [e]
GroupSize	Number of individual insureds in the group	3,540	305.7	5	7	60	1,094	28,691
PremiumRate	Annualized premium rate per insured	3,540	0.0024	0.000033	0.00036	0.0016	0.0066	0.066
Area[a]	Indicator of relative wealth and level of insurance market development of the group's location	3,540	1.68	1	1	1	3	4
Sex	Fraction of women in the group	3,540	0.42	0	0.072	0.40	0.83	1
Age	Average age in the group	3,540	33.4	0 [c]	23.1	32.5	45.4	57.6
Work[b]	Group average occupation accident tendency	3,432	1.96	1	1	1.91	3	6
N	Total number of group policies	3,540						

Note:

a. 1 represents the most developed regions in China, while 4 represents the least developed regions. The *area* is based on the insurer's branch categories, which consider not only regional wealth level but also regional insurance development level. It is a better control variable than pure wealth measurement.

b. 1 represents the safest occupations, while 6 represents the most dangerous. The *work* is measured by the accident tendency of an occupation, e.g., office workers are 1, while coal mine workers are 6.

c. Only one group policy has a minimum age of zero. Our results are robust if we exclude this policy.

d. Only one group policy has 60 claims, and only two policies have more than 20 claims. Our results are robust if we exclude these policies.

e. Only 19 group policies have policy durations longer than 366 days. Such policies reflect negotiated conditions for special clients. Our results are robust if we exclude these policies.

Table 3 Comparison between Subsamples

Variables	New Policies			Renewed Policies			Policies renewed two or more times			Policies renewed three or more times			ANOVA F-Test for Mean Difference	
	Mean	Median	Obs.	Mean	Median	Obs.	Mean	Median	Obs.	Mean	Median	Obs.	F-Value	P-Value
avClaimDummy[a]	0.0000019	0	1,690	0.0000016	0	1,850	0.0000012	0	1,088	0.0000010	0	745	1.29	0.275
avClaimCount[a]	0.0000020	0	1,690	0.0000018	0	1,850	0.0000013	0	1,088	0.0000010	0	745	1.71	0.163
InsuranceAmount	64,204.6	50,000	1,690	60,428.8	50,000	1,850	59,717.0	50,000	1,088	55,483.5	50,000	745	5.64	0.001
PolicyDuration	334.9	365	1,690	237.8	333	1,850	171.5	31	1,088	107.8	31	745	670.57	0.000
GroupSize	240.3	40	1,690	365.5	95	1,850	395.4	104	1,088	378.8	113	745	4.02	0.007
PremiumRate	0.0026	0.0020	1,690	0.0021	0.0013	1,850	0.0018	0.0012	1,088	0.0015	0.0012	745	61.17	0.000
Area	1.92	2	1,690	1.45	1	1,850	1.32	1	1,088	1.20	1	745	248.59	0.000
Sex	0.41	0.39	1,690	0.42	0.40	1,850	0.44	0.42	1,088	0.46	0.44	745	10.80	0.000
Age	34.1	33.7	1,690	32.6	31.6	1,850	31.2	30.0	1,088	29.7	28.7	745	83.69	0.000
Work	1.93	1.67	1,597	1.98	2	1,835	2.19	3	1,079	2.41	3	739	49.07	0.000
N			1,690			1,850			1,088			745		

Note:

a. *av* represents the average per insured per day. It scales claim performance indicators to a comparable level.

Empirical Models

To identify adverse selection, we use the classical risk-coverage correlation model shown in Equation (1) (Cohen and Siegelman, 2010). A positive correlation between risk and coverage is the necessary condition of adverse selection, implying that high risks buy more insurance coverage. An insignificant correlation between risk and coverage suggests no adverse selection. Alternatively, Chiappori and Salanié (1997, 2000) suggest the reduced form model shown in simultaneous Equations (2.1) and (2.2) to detect adverse selection. A positive correlation between the residuals ε_i and ω_i is the necessary condition of adverse selection, implying that the coverage choice and claim occurrence are not independent phenomena after controlling for observables. An insignificant correlation between ε_i and ω_i suggests no adverse selection.

Both models have advantages and limitations. Chiappori and Salanié (1997, 2000) suggest that the reduced form model may be more robust but less efficient than the classical approach. Chiappori and Salanié (2013) argue that the advantage of the reduced form model is that it does not require the estimation of the insurer's pricing policy, which, in our case, is not a problem because we know and control for the actual prices charged for each group policy. Cohen and Siegelman (2010) summarize that the two models are equivalent under general conditions, and the major differences between them depend on the distributional assumptions made conditional on the explanatory variables. Therefore, we use the classical model as our core model (Table 4) and the reduced form as a robustness test (All robustness tests are discussed in the Section of Robustness Tests and available from the authors upon request).

$$Risk_i = \alpha + \beta Coverage_i + \gamma X_i + \varepsilon_i \quad (1)$$

$$Risk_i = g(X_i) + \varepsilon_i \quad (2.1)$$

$$Coverage_i = f(X_i) + \omega_i \quad (2.2)$$

We measure the actual *Risk* of a group insured by its ex post claim performance in group policy i. The claim performance is measured by three claim indicators commonly used in existing literature: (1) *ClaimDummy*, which equals 1 if there is any claim under group policy i; (2) *ClaimCount*, which records the total number of claims under group policy i; and (3) *ClaimFrequency*, which equals the total number of claims divided by the number of individual insureds under group policy i. We do not use the total claim amount or claim severity as claim performance indicators because the claim severity of CI insurance always equals the insurance amount on an individual basis. The claim amount thus adds no additional information to *ClaimFrequency*.

We measure the insurance *Coverage* for group policy i as the natural logarithm of the insurance amount per insured. The insurance amount under group policy i is identical for every individual insured in the group. In our sample, group CI insurance always has a zero deductible and covers the same 25 critical diseases for all group policies and individual insureds. There are concerns that the insurer may price discriminate clients with different insurance amounts, i.e., provide discounted premium rates for groups buying higher insurance amounts and/or for large groups. However, we find no indications in the insurer's underwriting guidelines that such a pricing strategy exists. Moreover, we control for both group size and premium rate in our model to minimize the potential effects of a price discrimination strategy.

X_i is a vector of control variables that includes policy features, risk classification, and time effects. We control for policy features such as duration and group size. Risk classification refers to the use of observable characteristics by insurers to compute the corresponding premiums and thereby reduce asymmetric information (Dionne and Rothschild, 2014). Thus, any appearance of adverse selection in the market must reflect residual asymmetric information after controlling for risk classification. There are two alternative ways to control for risk classification: (1) observable demographic features and (2) premium rates computed based on observables. The premium rate is preferred to demographic features because it not only incorporates all demographics but also reflects the insurer's reaction to the respective group's claim experience (Finkelstein and McGarry, 2006). The premium rate represents the insurer's up-to-date best estimation of risks; therefore, we use the average annualized premium rate per person, *PremiumRate*, to control for risk classification. We use demographic features in a robustness test. We use year dummies to control for time effects.

We use logistic regression to estimate the model with *ClaimDummy* as the dependent variable and estimate a probit model as a robustness test. We estimate a negative binomial regression with *ClaimCount* as the dependent variable, legitimated by the likelihood ratio tests shown in Table 4, and use Poisson, zero-inflated negative binomial, and zero-inflated Poisson models as robustness tests. We estimate a Tobit model to fit the zero-censored dependent variable *ClaimFrequency*. We use cross-sectional models to test our hypotheses because one group usually has only one new policy and one nth time renewed policy, but use panel data models as robustness tests where applicable.

We examine the potential issues of multicollinearity, endogeneity and heterogeneity. The variance inflation factors (VIF) of the independent variables range from 1.13 to 1.51

for the new policy portfolio and between 1.14 and 1.63 for the renewed policy portfolio. All VIFs are below 5, suggesting that multicollinearity is not a problem. We test for potential endogeneity of the primary explanatory variable, *lnInsuranceAmount*, using the instrumental variable (IV) approach. The IVs used are demographic features (*age*, *work*, *sex*, and *area*) that determine the demand for insurance and thus correlate with *lnInsuranceAmount*. These demographic features are exogenous, and their effects on *Risk* are captured by another control variable, *lnPremiumRate*; thus, are not related to the error term of Equation (1). Wald tests of endogeneity (see Wooldridge, 2002, pp. 472-477 for a detailed discussion) are conducted for nonlinear models. For the subsample of new (renewed) policies, the p-values of Wald tests are 0.84 (0.31) for probit[16] regression with *ClaimDummy* as the dependent variable and 0.79 (0.09) for Tobit regression with *ClaimFrequency* as the dependent variable. Durbin-Wu-Hausman (DWH) tests of endogeneity are conducted using a two-stage least square (2SLS) model[17] with *ClaimCount* as the dependent variable. The DWH tests yield p-values of 0.25 for the new policy portfolio and 0.64 for the renewed policy portfolio. All tests suggest the acceptance of the null hypothesis of the exogenous *lnInsuranceAmount* at the 95% confidence level. The use of simple, linear functional forms, such as logit or probit models, should be restricted to homogeneous populations (Chiappori and Salanié, 2000). Our dataset approximates homogeneity because (1) the business nature of our sample is largely the same as employee benefits, (2) the insurer sources its business nationwide in China, (3) we control for potential heterogeneity among different group insureds using either premium rates or demographic features, and (4) we use robust standard errors clustered by group insureds in all specifications to further control for heterogeneity.

Results

Table 4 presents the results for Equation (1). The three panels show estimations with three different claim indicators, i.e., *ClaimDummy*, *ClaimCount*, and *ClaimFrequency*. The six columns show specifications with different subsamples.

To test *Hypothesis I*, we use the subsample of new policies because new customers are most likely to have information advantages over the insurer (Cohen, 2005). The results in Column 1 of Table 4 show that the insurance amount positively and significantly

[16] We use a probit model to perform the Wald test with *ClaimDummy* as the dependent variable because the error term of the logit model is not normally distributed.

[17] We use a 2SLS model to perform the DWH test with *ClaimCount* as the dependent variable because the counted number of claims approximates continuous; thus, the DWH test for linear model applies.

correlates with all three claim indicators conditional on premium rate. We interpret the coefficients as follows: if the insurance amount per person increases by 1%, the probability that the group will make a claim increases by 0.85% (Panel A) and the group claim frequency increases by 0.34 percentage points (Panel C). The positive correlation between risk and coverage suggests the existence of between-group adverse selection.[18] This evidence is against *Hypothesis I.* Group insurance cannot eliminate adverse selection, even if it does not allow for individual choices and even if the group is formed for purposes other than purchasing insurance. Such adverse selection is, by definition, between-group adverse selection. The results support Hanson's (2005) prediction.

To test *Hypothesis II*, we use subsamples of renewed policies to examine the extent of persistence (or lack thereof) of between-group adverse selection. We structure our sample in two alternative ways: (1) three layered sub-portfolios of renewed policies, i.e., all renewed policies, policies renewed two or more times, and policies renewed three or more times (Columns 2-4, Table 4); and (2) three exclusive sub-portfolios of renewed policies, i.e., first-time renewed policies, second-time renewed policies (Columns 5-6, Table 4), and policies renewed three or more times (Column 4, Table 4). The results in Columns 2 and 5 of Table 4 show positive and significant risk-coverage correlations in first-time renewed policies for all three claim indicators[19]. The results in Columns 3, 4, and 6 show that for policies renewed two or more times, the positive risk-coverage correlation disappears. The results suggest that there exists between-group adverse selection in new policies and that it persists for the first-time renewal, but disappears for policies renewed two or more times. Because the risk-coverage correlation is a necessary condition of adverse selection[20] (Chiappori and Salanie, 2000), we conclude that between-group adverse selection disappears over time as group insured renews with the same insurer for a certain period. The evidence supports *Hypothesis II*. The disappearance of between-group adverse selection can be explained by insurer learning

---

[18] The finding in cross-sectional data that coverage is correlated with risk does not suffice to tell us whether it is caused by adverse selection alone, moral hazard alone, or both (Cohen and Siegelman, 2010). We address this issue in the Discussion section.

[19] We conduct Z-statistic tests to compare the claim-coverage coefficients of the new policy portfolio in Column 1 with the claim-coverage coefficients of the (first-time) renewed policy portfolio in Column 2 (Column 5). All of them are not statistically different from each other at the 95% confidence level.

[20] Finkelstein and McGarry (2006) argue that the positive risk-coverage correlation may be neither a necessary nor a sufficient condition for the presence of asymmetric information about risk type, but instead, indicates a risk preference (attitude). We address this issue in the Discussion section.

over time[21], that is, insurer experience rating and underwriting based on a group's claim experience mitigate the group insured's information advantage, and thus mitigate between-group adverse selection.

The control variables indicate that there are more claims with longer policy durations and larger groups. Group size positively correlates with claim frequency, suggesting that large groups have a higher claim frequency than small groups. Premium rates positively correlate with number of claims and claim frequency, suggesting that the insurer successfully captures the group risk quality with premium. High-risk groups must pay higher premiums due to observables and poor claim experience.

Robustness Tests

We conduct six robustness tests, the results of which are available from the authors upon request. First, we estimate the reduced form model (Chiappori and Salanié, 1997, 2000) and monitor the correlation coefficients of the residuals from Equations (2.1) and (2.2). There are in total six specifications with both separate and system estimations of the two equations and with three claim indicators, respectively. We allow for potential error term correlations between the two equations when doing system estimation with the full information maximum likelihood (FIML) method. Alternatively, we use the three-stage least squares (3SLS) method to replace FIML, the results of which are very similar. In five of six specifications (except system estimation with claim frequency), the residual correlations are positive and significant for the subsamples of new and renewed policies; and insignificant for the subsamples of policies renewed two or more times. The results suggest that between-group adverse selection occurs with new customers but not with customers who renew for a certain period.

[21] We discuss alternative explanations in the Discussion section.

Table 4 Core Model Results

	(1)	(2)	(3)	(4)	(5)	(6)
Subsamples	New Policies	Renewed Policies	Policies Renewed 2 or More Times	Policies Renewed 3 or More Times	1st-time renewed policies	2nd-time renewed policies
Panel A	**Dependent Variable: *ClaimDummy*, Model: *Logistic***					
lnInsuranceAmount	0.00853***	0.00463*	0.00116	0.00160	0.0194***	0.00115
	(0.00286)	(0.00246)	(0.00145)	(0.00117)	(0.00665)	(0.00651)
lnPremiumRate	0.0127***	0.00603*	0.00145	0.00169	0.0246***	-0.00165
	(0.00393)	(0.00321)	(0.00164)	(0.00126)	(0.00869)	(0.00907)
lnGroupSize	0.0293***	0.0165**	0.00712	0.00438**	0.0541***	0.0332***
	(0.00379)	(0.00647)	(0.00453)	(0.00189)	(0.00739)	(0.0128)
lnPolicyDuration	0.0671***	0.0369***	0.0166**	0.00774***	0.165***	0.174***
	(0.0226)	(0.00711)	(0.00731)	(0.00270)	(0.0427)	(0.0403)
Pseudo R^2	0.307	0.362	0.440	0.427	0.279	0.419
Panel B	**Dependent Variable: *ClaimCount*, Model: *Negative Binomial***					
lnInsuranceAmount	0.00863***	0.00539***	0.00105	0.0000658	0.0211***	-0.00293
	(0.00236)	(0.00196)	(0.00147)	(4.37e-05)	(0.00570)	(0.0107)
lnPremiumRate	0.0127***	0.00924***	0.00372**	0.000135***	0.0305***	0.00653
	(0.00352)	(0.00254)	(0.00163)	(4.68e-05)	(0.00692)	(0.00991)
lnGroupSize	0.0284***	0.0201***	0.00953***	0.000220***	0.0599***	0.0455***
	(0.00432)	(0.00416)	(0.00258)	(4.07e-05)	(0.00955)	(0.00951)
lnPolicyDuration	0.0776***	0.0380***	0.0209***	0.000293***	0.150***	0.140***
	(0.0223)	(0.00494)	(0.00475)	(9.07e-05)	(0.0441)	(0.0341)
Pseudo R^2	0.319	0.312	0.364	0.396	0.259	0.310
P-value of LR test [a]	0.000	0.000	0.055	0.059	0.019	0.500
Panel C	**Dependent Variable: *ClaimFrequency* (scaled up by 1,000), Model: *Tobit***					
lnInsuranceAmount	3.439**	1.292*	0.901	2.402	1.822*	0.0203
	(1.510)	(0.708)	(0.802)	(1.768)	(1.016)	(0.848)
lnPremiumRate	5.108**	2.134***	1.875**	3.213*	2.240*	0.777
	(2.054)	(0.780)	(0.937)	(1.646)	(1.250)	(1.200)
lnGroupSize	10.45***	4.925***	3.840***	4.828***	5.629***	3.220***
	(1.415)	(0.487)	(0.592)	(1.026)	(0.587)	(0.725)
lnPolicyDuration	26.91***	11.25***	8.625***	7.519**	19.19**	16.33***
	(10.11)	(2.670)	(2.182)	(3.035)	(8.884)	(4.757)
Pseudo R^2	0.091	0.124	0.182	0.211	0.072	0.128
Intercepts and year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,690	1,850	1,088	745	762	343

Note:

We present the marginal effects of logistic and negative binomial models at the means of the independent variables and present the estimated coefficients of Tobit model. Robust standard errors clustered by group insureds are presented in parentheses. We also present *, **, ***, indicating significant differences of coefficients from 0 at the 10%, 5%, and 1% levels, respectively.

a. The log-likelihood ratio test discriminates negative binomial model (H1: alpha≠0) from Poisson model (H0: alpha=0).

Second, we use demographic features (*age*, *sex*, *work* and *area*) to replace *lnPremiumRate* as the risk classification control. We introduce dummies *area1* to *area4* to control for regional differences in levels of wealth and insurance market development. We use fraction variables of different age ranges to reflect the age mixture within one group and the fraction variables of occupation categories *work1* to *work6* to reflect the mixture of occupations within a group[22]. For both dummy and fraction control variables, we omit the largest category to avoid collinearity. The results are similar to our core models.

Third, we consider the potential effect of group size on between-group adverse selection. The literature on U.S. group health insurance suggests that between-group adverse selection may only exist or be more problematic for small groups because small firms may seek coverage simply because an employee or dependent is ill; large groups may have less of an information advantage, if any (Monheit and Schone, 2004). We test whether between-group adverse selection exists for large groups and, if so, whether small groups have stronger adverse selection than large groups. Our dataset contains both small groups with 50 or fewer members and large groups with more than 50 members.[23] We thus introduce the interaction term *lnInsuranceAmount**small* to test these arguments. The results show (1) the same existence and persistence pattern of between-group adverse selection for large groups and (2) insignificant coefficients of interaction terms for all specifications (except the negative binomial model with policies renewed two or more times), suggesting that the level of between-group adverse selection for small groups is not materially different from that of large groups.

Fourth, we consider the potential nonlinear effects of our primary explanatory variable, *lnInsuranceAmount*. Dionne, Gouriéroux, and Vanasse (2001) show that the significant risk-coverage correlation may disappear after adding the projected primary explanatory variable to the classic model. We mirror their methodology by (1) regressing *lnInsuranceAmount* on other policy features, risk classification, and year dummies; (2) predicting the *projected lnInsuranceAmount* from the step 1 regression; and (3) including the *projected lnInsuranceAmount* as an additional control variable in Equation (1). The

---

[22] Other studies also consider schooling, income, or number of dependents as control variables; however, that information is not captured by the insurer and thus is not considered in pricing CI risks. Moreover, these variables are more determinants of insurance demand, but relatively less influence the risk of critical illness.

[23] The 50-person cutoff point for small and large groups has been adopted by most states in U.S. (Simon, 2005) and is used in many group health insurance studies (e.g. Cutler, 1994; Monheit and Schone, 2004).

risk-coverage correlation remains positive and significant for new policy portfolio, suggesting the existence of between-group adverse selection is robust.

Fifth, we find that our conclusions hold when estimating alternative econometric models, including a probit model on *ClaimDummy*; Poisson, zero-inflated Poisson, and zero-inflated negative binomial models on *ClaimCount*; and panel random effects models for renewed portfolios, which contain multiple continuous policies for one group insured.

Sixth, we consider the potential model misspecification risk of our core linear models; thus, adopt a semi-parametric approach (Robinson, 1988) as shown in Equation (3). We test whether a semi-parametric model is significantly different from our core models using Hardle and Mammen's (1993) specification tests. The results suggest that the model misspecification risk of our core linear models is minimal, i.e., for subsamples of new and renewed policies, the semi-parametric model is equivalent to the linear model with first-order *lnInsuranceAmount*; and for subsamples of policies renewed two or more times, the semi-parametric model is equivalent to the linear model without *lnInsuranceAmount* (zero order). We also provide an example semi-parametric estimation, which illustrates that the risk-coverage correlation is positive for new and renewed portfolios but less significant for policies renewed two or more times.

$$Risk_i = \alpha + \delta(Coverage_i) + \gamma X_i + \varepsilon_i \tag{3}$$

Discussion

Adverse Selection vs. Moral Hazard

A critical issue to the positive risk-coverage correlation test is disentangling adverse selection from moral hazard. Both concepts predict that insured (fully insured) should have a higher probability of accident than do uninsured (partially insured) (Richaudeau, 1999). The positive correlation found in cross-sectional models does not suffice to tell whether the relationship is caused by adverse selection alone, moral hazard alone, or both (Cohen and Siegelman, 2010). We thus disentangle between-group adverse selection from moral hazard from statistical, theoretical and product design perspectives.

Statistically, we use longitudinal data to conduct the disentangling test introduced by Abbring et al. (2003). They suggest that moral hazard should lead to a negative correlation between historical claims and claims in the subsequent period under the experience rating system because the insured’s behavior changes corresponding to experience rating and insurance coverage. In contrast, adverse selection should result in a positive correlation between past and future claims because claims are determined by the insured's risk type. We find that new policy claims and first-time renewed policy claims are positively related to each other, which is consistent with adverse selection. The results are available from the authors upon request. Theoretically, the disappearance pattern of risk-coverage correlation itself provides direct evidence that at least some new customers have private information about their risk type, which cannot be explained by moral hazard (Cohen and Siegelman, 2010).

Moreover, we consider potential risk-bearing and claim-reporting moral hazard in group insurance (Butler and Worrall, 1991; Ruser and Butler, 2010; Butler, Gardner and Kleinman, 2013). Previous research explains the pattern of risk-coverage correlation in workers' compensation (WC) with moral hazard and argues that the incentives of employers and employees under experience-rated group insurance coverage are different. Employees tend to care less and report more claims than employees with no insurance, where the moral hazard story holds. Employers, however, face incentives to improve safety and report fewer claims to keep the premium low in the following year. The employer's incentive, termed risk-bearing moral hazard, mitigates any moral hazard and claim-reporting moral hazard of employees, which may also show a disappearance pattern of risk-coverage correlation. This explanation was developed in the WC context, in which adverse selection is minimal due to compulsory insurance. However, in group CI insurance, neither the employer nor the employee is likely able to systemically influence the frequency of CI incidents and the market is commercial and voluntary. Thus, the moral hazard in WC market may not apply.

The product design of group CI insurance also minimizes moral hazard because the insurer pays the full insurance amount once the covered critical disease is diagnosed, which avoids the over-utilization problem commonly observed in medical expense insurance (i.e., *ex post* moral hazard). Moreover, considering the very small expected benefit, the insured's incentive to change his/her lifestyle due to CI coverage is minimal. Wang et al. (2011) find adverse selection in Taiwan's cancer insurance market, which is similar to our CI insurance product. They conclude that purchasing extended cancer insurance does not reduce insureds' efforts to prevent cancer. We also note the potential for claim fraud. However, the claim payment is subject to the verification procedure that the claimant is always asked to obtain a second opinion on the diagnosis from a different hospital approved by the insurer, which significantly reduces the risk of claim fraud. In general, group CI insurance is much less affected by moral hazard and claim fraud than is medical expense health insurance.

Alternative Explanations for Adverse Selection Disappearance

Our explanation for the disappearance of adverse selection is insurer learning over time. We discuss alternative explanations below. First, the disappearance of risk-coverage correlation could be driven by risk quality changes or selections of the group. The covered individuals may drop their coverage in the subsequent period when their groups renew, due to death, dismissal from the group, or other reasons (Yao, 2013). We thus compare the reason for coverage drop-offs between claimants (individuals filing a claim) and non-claimants (individuals not filing a claim). For claimants, 65% of drop-offs are due to group drop-offs, and 35% are standalone drop-offs when their groups renew in the subsequent period. For non-claimants, 62% of drop-offs are due to group drop-offs, and 38% are standalone. Clearly, most individual drop-offs are attributable to group drop-offs, and claimants are not more likely to leave the portfolio on a standalone basis than are non-claimants. Moreover, we examine whether claimants are more likely to leave the portfolio under the condition that the group renews in the subsequent period. For claimants, 28% leave the portfolio when their groups stay; for non-claimants, 25% leave. At the 95% confidence level, claimants are not more likely to leave the portfolio than are non-claimants when their groups stay, subject to a binomial probability test. We find no evidence on group risk quality change over time or risk selection by the group insured. The disappearance of adverse selection is unlikely to be driven by risk quality changes within a group.

Another explanation is multi-dimensional information advantages. Finkelstein and McGarry (2006) argue that there are multiple dimensions of private information related not only to insureds' risk types but also to their risk attitudes (or risk preferences). They suggest that empirical studies based on risk-coverage correlations should control for insureds' risk attitudes because risk-averse insureds are often associated with low levels of risk, which will blur the positive risk-coverage correlation test due to "advantageous selection" (De Meza and Webb, 2001; Fang, Keane, and Silverman, 2008). However, it is hard to argue that such a correlation between risk attitudes and risk coverage applies to group insureds. It has been noted that risk aversion provides much less of a motivation for corporate insureds to purchase insurance, particularly for stock companies, because stakeholders could instead manage idiosyncratic losses through diversified portfolios (Mayers and Smith, 1982, 1990). For Chinese corporate insurance, Zhu, Kui, and Fang (2011) find that risk aversion is not a significant factor in insurance demand. In our analysis, group insureds are largely corporations and decision makers are usually in HR departments; therefore, we assume that group insureds' risk attitudes are not the major driver of insurance demand and will not distort the risk-coverage correlation.[24]

Generalizability of Low Loss Frequency Portfolio and of Chinese Market

Medical expense health insurance and automobile insurance are common types of insurance products used to test for adverse selection. The high loss frequency of such products enables econometricians to detect adverse selection more easily than do low loss frequency products, partially because the small number of claims in a dataset may bias the measurement of actual risk when using ex post losses and partially because the insured's perception of risk types is biased when the event probability is low (Cawley and Philipson, 1999). We thus discuss the potential effect of loss frequency on between-group adverse selection.

Theories of adverse selection do not distinguish between predictions of high and low probability risks. However, some theories suggest that people exhibit biases in judgments about risks and probabilities when the probabilities of events are small (Camerer and Kunreuther, 1989). In the context of group CI insurance, the insurance decision maker, such as the HR department of the group insured, usually has experience purchasing insurance and possesses relevant information about employee health.

Empirically, in our sample, there is one loss in every six group-year policies or every 1,500 individual-year policies, as shown in Table 2. This loss frequency is lower than in automobile and health insurance but not materially different from many common insurance products, such as fire or term life insurance. There are also empirical studies that detect adverse selection in low loss frequency markets where longitudinal data increase the power of detection (He, 2008). We perform a robustness test using a bootstrap resampling method (200 replications) to correct for potential bias in estimated standard errors (Efron and Tibshirani, 1993) due to a low loss frequency. The bootstrap resampling method constructs a number of resamples of the observed dataset and thus provides a way to account for distortions caused by a small sample that may not be fully representative of the population. The results confirm our conclusions and are available from the authors upon request.

---

[24] In addition, Chiappori and Salanié (2013) suggest that risk preference alone should have negligible consequences on the positive risk-coverage correlation in competitive markets because insureds of all types of risk aversion prefer full coverage in a competitive market. The Chinese group CI insurance market can be considered as a competitive market with standardized coverage and no rate regulation.

In addition to the evidence drawn from Chinese group CI insurance, the U.S. small-group health insurance market provides another example of between-group adverse selection, in which high prices and low coverage resulting from adverse selection are notorious (Simon, 2005). However, two major obstacles jeopardize the empirical conclusions pertaining to between-group adverse selection in the U.S. small-group health insurance market. First, it is difficult to separate the effect of individual choices from group strategic actions; second, small-group reforms in the early 1990s restricted insurer experience rating and redlining. In contrast, the Chinese group CI insurance market provides almost laboratory conditions to test between-group adverse selection, considering 1) standardized products with competing prices, 2) voluntary purchase and renewal for both buyers and insurers, and 3) commercial markets without considerations of broad availability and affordability.

In principle, any group insurance arrangement without individual choices is suitable to test between-group adverse selection.[25] For example, we expect to find similar results from employer-provided group health insurance in the U.S., if policies that allow individual plan choices are excluded, from crop insurance at the village- or county-level, and from motor fleet insurance including both liability and own damage coverage. We might also find between-group adverse selection in the supplementary coverage of social insurance. However, we do not argue for the universality of between-group adverse selection. Product lines with little adverse selection in the individual insurance market may not exhibit between-group adverse selection in the group insurance market, e.g., life insurance.

Concluding Remarks and Future Research

We find evidence for the existence of between-group adverse selection. Our dataset allows us to separate the effect of individual choice within a group from the effect of group strategic actions. The empirical findings complement conventional wisdom and support Hanson's (2005) theoretical prediction that adverse selection may well exist in a group insurance market even if no individual choices are allowed and even if the group is formed for purposes other than purchasing insurance. This paper also complements

[25] We use another health insurance portfolio from Pakistan to generalize our conclusions regarding *Hypothesis I*. The results confirm that between-group adverse selection exists (1) in the high loss frequency portfolio, (2) when the group insureds are families. The dataset cannot be used to test *Hypothesis II* due to the limited number of policy periods. These results are available from the authors upon request.

empirical works based on U.S. (small-) group health insurance, where the individual choice within a group is an important driver of adverse selection. In addition, we find that between-group adverse selection disappears over time as the group insured renews with the same insurer for a certain period. This evidence is consistent with the explanation of insurer learning over time. The combination of two pieces of evidence suggests that group insurance alone is not necessarily sufficient to mitigate adverse selection; in the repeated contracting setup, learning from the performance of past contracts and taking corresponding actions based on the information observed help to mitigate adverse selection problems (Dionne, 1983).

As discussed in the section of generalizability, we expect our conclusions to hold for multiple group insurance products in various markets of a similar nature, e.g., U.S. small-group health insurance. Our results may also shed light on markets other than insurance characterized by information asymmetry, where the existence and persistence of adverse selection are also relevant (see, e.g., Chari, Shourideh, Zetlin-Jones, 2014).

One implication of our results is that group insurance with no individual choice cannot be considered a market free of adverse selection, even if the group is formed for purposes other than purchasing insurance. Our results thus also have important business implications. Insurers must be aware that group insurance policyholders strategically act on their information advantages and should therefore monitor group claim experience over time, first to learn and then to apply that knowledge to their renewal underwriting and pricing decisions. The more efficiently insurers acquire and use such knowledge, the sooner they will overcome adverse selection.

Our dataset does not allow us to compare the level of group adverse selection to the level of individual adverse selection. An interesting next step would be to compare the magnitude and/or the persistent time of adverse selection between group and individual insurance. Moreover, there is, to date, limited theoretical work on why and how between-group adverse selection is generated via internal group processes. It would be useful to know how group insurance decision processes differ from individual insurance decisions. Finally, we recommend additional empirical studies to explore other eligible group insurance products in other markets to verify our empirical results.

Appendix 1 Critical diseases covered by the group CI policy

The complete list of 25 critical diseases covered under the group CI policy is shown below. The list is recommended by the Insurance Association of China and the Chinese Medical Doctor Association and has been adopted by most players in the Chinese CI insurance market. The standard and binding definitions of these diseases can be found in Insurance Association of China and Chinese Medical Doctor Association (2007).

1. Malignant Tumor
2. Acute Myocardial Infarction
3. Sequelae of Stroke
4. Major Organ / Hematopoietic Stem Cells Transplant
5. Coronary Artery Bypass Graft
6. End Stage Renal Disease (Chronic Kidney Failure)
7. Loss of Limbs
8. Acute or Subacute Severe Hepatitis
9. Benign Brain Tumor
10. Chronic Liver Failure (End Stage)
11. Encephalitis Sequelae or Meningitis Sequelae
12. Deep Coma
13. Deafness in Both Ears
14. Blindness in Both Eyes
15. Paralysis
16. Heart Valve Surgery
17. Severe Alzheimer's Disease
18. Major Head Trauma
19. Severe Parkinson's Disease
20. Major Third Degree Burn
21. Severe Primary Pulmonary Hypertension
22. Severe Motor Neuron Disease
23. Loss of Speech
24. Severe Aplastic Anemia
25. Aorta Surgery

References

Abbring, J. H., Chiappori, P. A., Heckman, J. J., and Pinquet, J., 2003, Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish? *Journal of the European Economic Association*, 1(2–3): 512–521.

Akerlof, G. A., 1970, The Market for "Lemons": Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, 84(3): 488-500.

Barnard, M., 2004, Critical Illness Insurance: Past, Present and Future, downloaded from http://www.actuaries.org.uk/sites/all/files/documents/pdf/Barnard.pdf.

Bickelhaupt, D. L., 1983, *General Insurance*, 11th revised Ed. (New York NY: McGraw-Hill).

Browne, M. J., 1992, Evidence of Adverse Selection in the Individual Health Insurance Market, *Journal of Risk and Insurance*, 59(1): 13-33.

Butler, R. J., Gardner H. H., and Kleinman N. L., 2013, Chapter 16 Workers' Compensation: Occupational Injury Insurance's Influence on theWorkplace, *Handbook of Insurance*, (Springer), pp. 449-469.

Butler, R. J., and Worrall, J. D., 1991, Claims Reporting and Risk Bearing Moral Hazard in Workers' Compensation, *Journal of Risk and Insurance*, 58(2): 191–204.

Camerer, C. F., and Kunreuther, H., 1989, Decision Processes for Low Probability Events: Policy Implications, *Journal of Policy Analysis and Management*, 8(4): 565-592.

Cawley, J., and T. Philipson, 1999, An Empirical Examination of Information Barriers to Trade in Insurance, *American Economic Review*, 89: 827-846.

Chiappori, P. A., and Salanié, B., 1997, Empirical Contract Theory: The Case of Insurance Data, European Economic Review, 41(3–5): 943-950.

Chiappori, P. A., and Salanié, B., 2000, Testing for Asymmetric Information in Insurance Markets, *Journal of Political Economy*, 108(1): 56-78.

Chiappori, P. A., and Salanié, B., 2013, Chapter 14 Asymmetric Information in Insurance Markets: Predictions and Tests, *Handbook of Insurance*, (Springer), pp. 397-422.

Chari, V. V., Shourideh, A., and Zetlin-Jones, A., 2014, Reputation and Persistence of Adverse Selection in Secondary Loan, *American Economic Review*, forthcoming.

China Insurance Regulatory Committee, 2013, *Yearbook of China's Insurance*. (Beijing: Yearbook of China's Insurance Press).

Cochrane, J., 1995, Time-Consistent Health Insurance, *Journal of Political Economy*, 103(3):445-473.

Cohen, A., 2005, Asymmetric Information and Learning: Evidence from the Automobile Insurance Market, *Review of Economics and Statistics*, 87(2): 197-207.

Cohen, A., 2012, Asymmetric Learning in Repeated Contracting: An Empirical Study, *Review of Economics and Statistics,* 94(2): 419-432.

Cohen, A., and Siegelman, P., 2010, Testing for Adverse Selection in Insurance Markets, *Journal of Risk and Insurance,* 77(1): 39-84.

Cummins, J. D., Smith, B. D., Vance, R. N., and VanDerhei, J. L., 1983, *Risk Classification in Life Insurance*, (Springer), softcover reprint at 2013.

Cutler, D. M., 1994, Market Failure in Small Group Health Insurance, National Bureau of Economic Research, No. w4879.

Cutler, D. M., and Zeckhauser, R. J., 2000, Chapter 11 The Anatomy of Health Insurance, *Handbook of Health Economics*, Vol. 1A, (Amsterdam: Elsevier), pp. 563-643.

D'Arcy, S. P., and Doherty, N. A., 1990, Adverse Selection, Private Information, and Lowballing in Insurance Markets, *Journal of Business*, 63(2): 145-164.

De Garidel-Thoron, T., 2005, Welfare-Improving Asymmetric Information in Dynamic Insurance Markets, *Journal of Political Economy*, 113(1): 121-150.

De Meza, D., and Webb, D.C., 2001, Advantageous Selection in Insurance Markets, *RAND Journal of Economics*, 32: 249-262.

Dionne, G., 1983, Adverse Selection and Repeated Insurance Contracts, *Geneva Papers on Risk and Insurance*, 29: 316-332.

Dionne, G., Gouriéroux, C., and Vanasse, C., 2001, Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment, *Journal of Political Economy*, 109(2): 444-453.

Dionne, G., and Rothschild, C. G., 2014, Economic Effects of Risk Classification Bans, *Geneva Risk and Insurance Review,* 39(2): 184-221.

Fang, H., Keane, M.P., and Silverman, D., 2008, Sources of Advantageous Selection: Evidence from the Medigap Insurance Market, *Journal of Political Economy*, 116(2): 303-349.

Efron, B., and Tibshirani, R., 1993, *An Introduction to the Bootstrap*. (New York, NY: Chapman and Hall).

Finkelstein, A., and McGarry, K., 2006, Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market, *American Economic Review*, 96(4): 938-958.

Garven, J. R., Hilliard, J. I., and Grace, M. F., 2014, Adverse Selection in Reinsurance Markets, *Geneva Risk and Insurance Review*, 39(2): 222-253.

Handel, B. R., 2013, Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts, *American Economic Review*, 103(7): 2643-2682.

Hanson, R., 2005, Adverse Selection in Group Insurance: The Virtues of Failing to Represent Voters, *Economics of Governance,* 6(2): 139-157.

Hardle, W., and Mammen, E., 1993, Comparing Nonparametric Versus Parametric Regression Fits. *Annals of Statistics*, 21(4), 1926-1947.

He, D. (2008). The Life Insurance Market: Adverse Selection Revisited. Working paper, Department of Economics, Washington University in St. Louis. Downloaded from: http://artsci.wustl.edu/~ econgr/gradconference/08/DaifengHe. pdf.

Hendel, I., and Lizzeri, A., 2003, The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance, *Quarterly Journal of Economics*, 118(1): 299-328.

Insurance Association of China, and Chinese Medical Doctor Association, 2007, *Regulation on Definitions in Critical Illness Insurance*, downloaded from http://www.iachina.cn/.

Insurance Europe, 2014, *European Insurance — Key Facts*, downloaded from http://www.insuranceeurope.eu/.

Insurance Information Institute, 2013, *Insurance Fact Book 2014* (New York, NY: Insurance Information Institute), downloaded from http://www.iii.org/.

Insurance Research Council, 1991, *Adequacy of Motor Vehicle Records in Evaluating Driver Performance*, (Malvern, PA: Insurance Research Council), downloaded from http://www.insurance-research.org/.

Jean-Baptiste, E. L., and Santomero, A. M., 2000, The Design of Private Reinsurance Contracts, *Journal of Financial Intermediation*, 9(3): 274-297.

Kunreuther, H., and Pauly, M., 1985, Market Equilibrium with Private Knowledge: An Insurance Example, *Journal of Public Economics,* 26(3): 269-288.

Mayers, D., and Smith, C. W., 1981, Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets, *Journal of Business*, 54(3): 407-434.

Mayers, D., and Smith, C. W., 1982, On the Corporate Demand for Insurance, *Journal of Business,* 55(2): 281-296.

Mayers, D., and Smith, C. W., 1990, On the Corporate Demand for Insurance: Evidence from the Reinsurance Market, *Journal of Business,* 63(1): 19-40.

Monheit, A. C., and Schone, B. S., 2004, How Has Small Group Market Reform Affected Employee Health Insurance Coverage? *Journal of Public Economics*, 88(1), 237-254.

Pauly, M. V., and Percy, A. M., 2000, Cost and Performance: A Comparison of the Individual and Group Health Insurance Markets, *Journal of Health Politics Policy and Law,* 25(1): 9-26.

Richaudeau, D., 1999, Automobile Insurance Contracts and Risk of Accident: An Empirical Test Using French Individual Data, *Geneva Papers on Risk and Insurance Theory*, 24(1), 97-114.

Robinson, P. M., 1988, Root-N-Consistent Semiparametric Regression, .*Econometrica*, 56(4), 931-954.

Rothshild, M., and Stiglitz, J., 1976, Equilibrium in Competitive Insurance Market: An Essay on the Economics of Imperfect Information, *Quarterly Journal of Economics*, 90(4): 629-647.

Ruser J., and Butler R., 2010, The Economics of Occupational Safety and Health, *Foundations and Trends in Microeconomics,* 5(5): 301-354.

Simon, K., 2005, Adverse Selection in Health Insurance Markets? Evidence from State Small-Group Health Insurance Reforms, *Journal of Public Economics*, 89(9): 1865-1877.

Su, N., 2013, The First Critical Illness Table from CIRC Has Limited Short Term Impact, caijing.com.cn, http://finance.caijing.com.cn/2013-12-04/113648773.html.

Watt, R., and Vazquez, F. J., 1997, Full Insurance, Bayesian Updated Premiums, and Adverse Selection, *Geneva Papers on Risk and Insurance Theory,* 22(2): 135-150.

Wang, K. C., Peng, J., Sun, Y., and Chang, Y., 2011, The Asymmetric Information Problem in Taiwan's Cancer Insurance Market. *Geneva Risk and Insurance Review,* 36(2), 202-219.

Wooldridge, J. M., 2002, *Econometric Analysis of Cross Section and Panel Data.* (Cambridge, MA: MIT Press).

Yao, Y., 2013, Development and Sustainability of Emerging Health Insurance Markets: Evidence from Microinsurance in Pakistan, *Geneva Papers on Risk and Insurance-Issues and Practice*, 38(1): 160-180.

Zhu, M., Kui, C., and Fang, Y., 2011, Demand for Corporate Insurance in China: Empirical Analysis with Province-Level Panel Data, *Chinese Economy,* 44(6): 15-29.