

Online Supplementary Appendices

Information Transparency, Choice Confidence, and Customer Retention: Field and Lab Evidence from Markets for Credence Goods

Aaron Cheng (LSE), Ting Li (Erasmus University), Paul A. Pavlou (University of Miami)

Appendix A. Context Details, Data, and Descriptive Statistics

A.1. “Get a Quote” via a Comparison Website

Customers who go through the “get a quote” process on comparison sites need to first type in the information about themselves, their cars, and their preferences for insurance coverage (e.g., limited liability or a comprehensive protection). Then the customers will be directed to a comparison page where all insurers will be displayed in a list. Customers will not only see the prices, but also the coverage details, the number of ratings by other customers, the average star rating (e.g., 8.1/10), and the insurance policy conditions (e.g., 7.5/10) rated by the third-party comparison site. These non-price attributes (with or without customer clicking) would signal market demand for each insurer’s offer and, therefore, affect customer choices. Customers on the last comparison page can click “+ more information” to know more about each offer, including the basics (daily cancellability, telephone contact, paperless policy), details about the coverage, and customer reviews one by one. [Figure A-1](#) below presents an example of the last offer comparison page.

Insurer	Price-quality	Deductible	Policy conditions	Star Rating (Reviews)	Monthly Premium	Additional Info
Insurer 1	Excellent	€ 0,-	7.1	83 reviews (4.5 stars)	52.39 per month	Free physio after a car accident
Insurer 2	Good	€ 0,-	7.5	605 reviews (8.1 stars)	68.26 per month	Free physio after a car accident
Insurer 3	Unfavorable	€ 0,-	7.8	27481 reviews (7.8 stars)	76.77 per month	Free physio after a car accident

Figure A-1. An Offer Comparison Page of a Third-Party Comparison Website

Note: We redacted some information so as not to compromise the identities of the insurers and the comparison website in the above screenshot. Another important note is that the field data for our analysis is obtained from the partner insurer, rather than a third-party comparison website. Hence, this dataset does not allow us to observe and measure how exactly each customer interacted with information elements (e.g., clicking a button to see more product details) on comparison websites. While we use data from a “get a quote” simulation to measure prices and customer ratings, we still cannot identify whether a customer actually used the information in the above screen at the point of purchase, nor can we determine how they used it on the comparison website.

A.2. Model Free Evidence

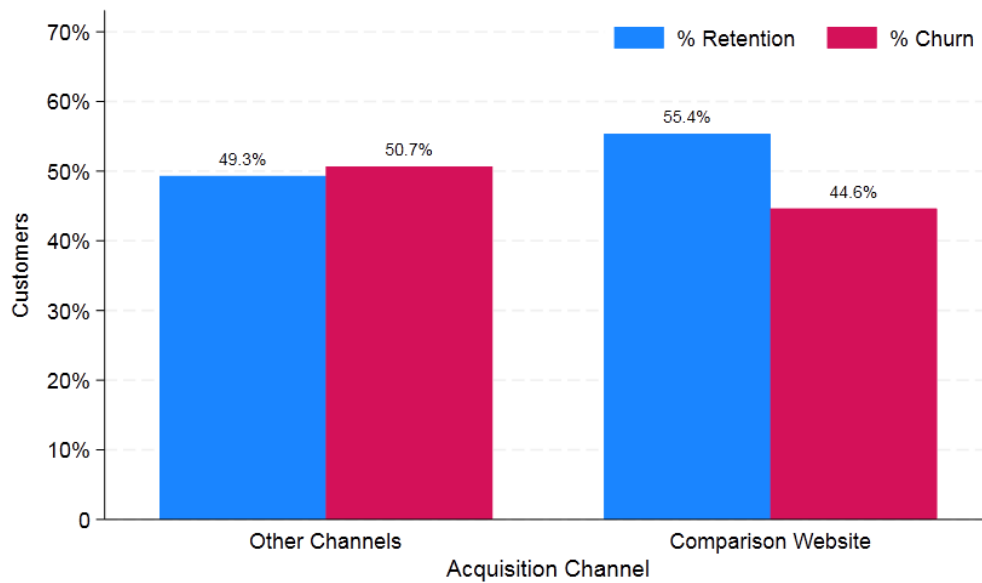


Figure A-2. Overall Retention and Churn Rates across Acquisition Channels (More Transparent Comparison Website vs. Less Transparent Other Channels)

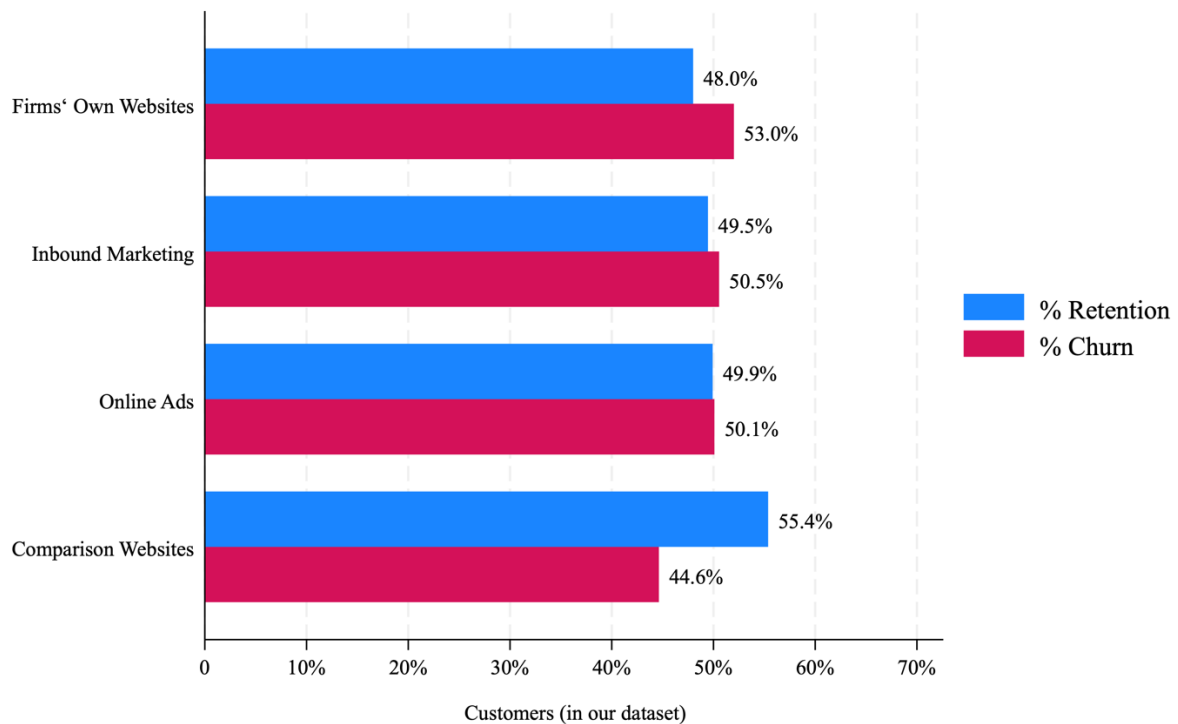


Figure A-3. Retention and Churn Rates across the Four Major Acquisition Channels (Firms' Own Websites, Inbound Telemarketing, Online Ads, and Comparison Websites)

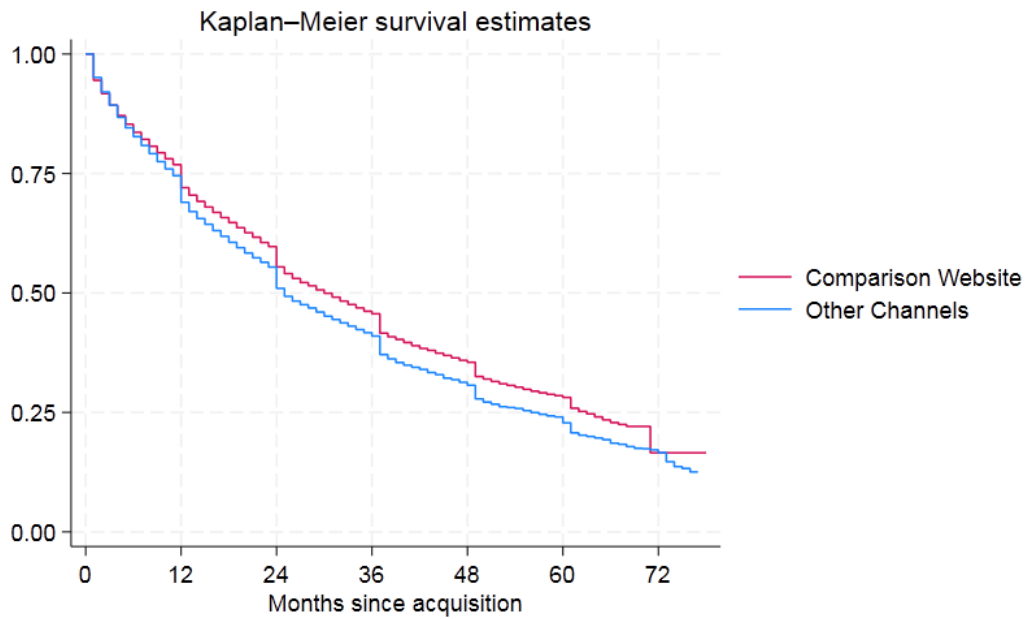


Figure A-4. Kaplan-Meier Survival Rates across Acquisition Channels Over Time (More Transparent Comparison Website vs. Less Transparent Other Channels)

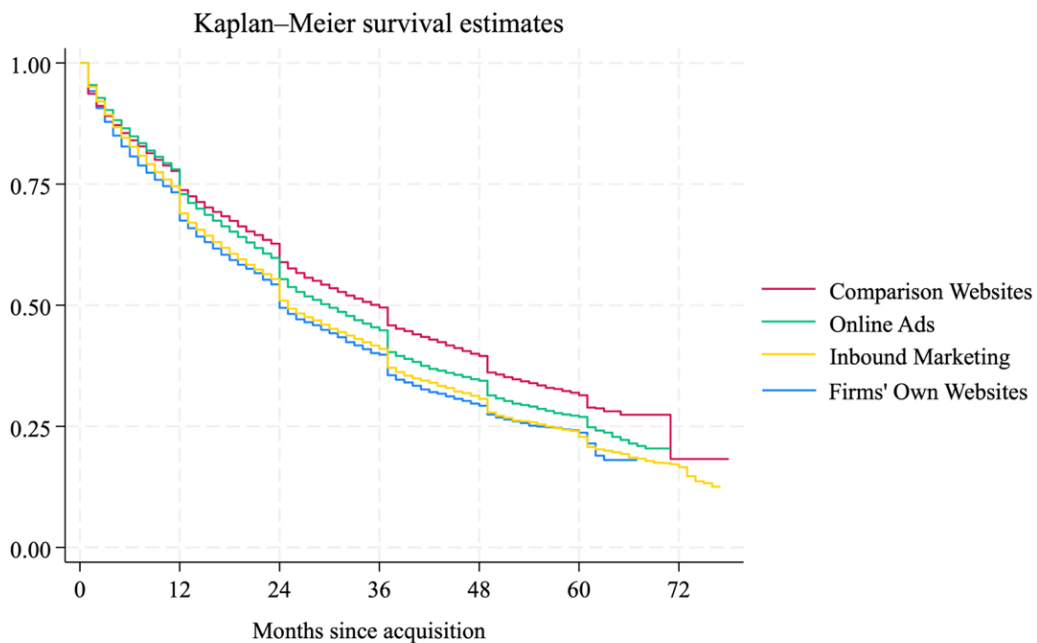


Figure A-5. Kaplan-Meier Survival Rates across the Four Acquisition Channels Over Time (Firms' Own Websites, Inbound Telemarketing, Online Ads, and Comparison Websites)

Appendix B. Additional Details for Main Analysis using Field Data

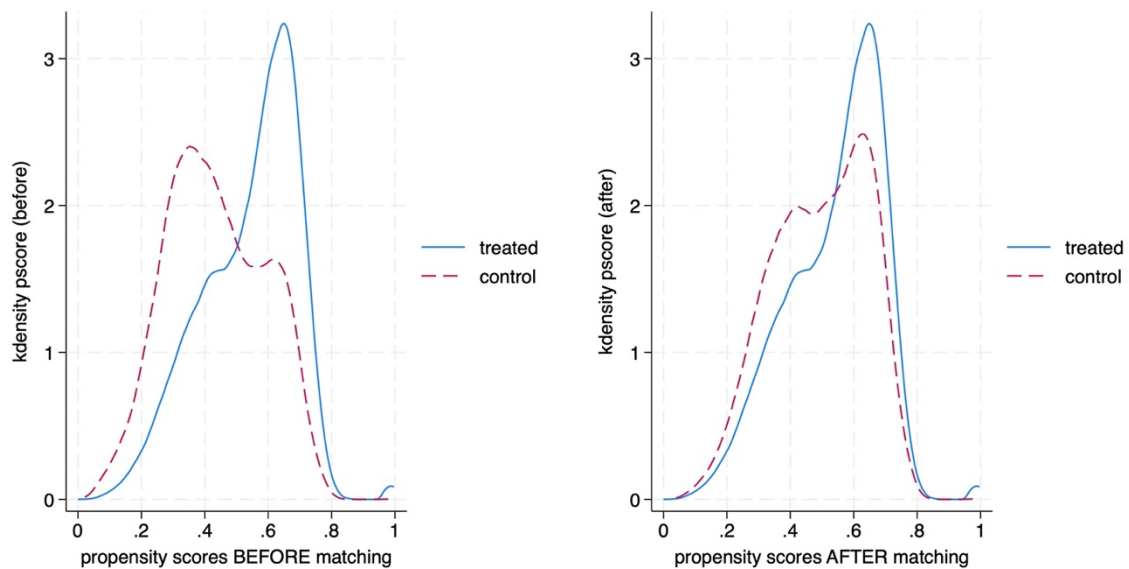
B.1. Analysis Using Matching

Table B-1. Covariate Balance Before and After Matching on Covariates from *Treatment Group* (i.e., Customers Acquired from Comparison Websites)

	Before Matching			After Matching		
	Treatment Group	Control Group	Standard difference	Treatment Group	Control Group	Standard difference
Driver gender	0.309	0.336	5.7	0.309	0.309	0.1
Driver age	43.810	46.525	19.3	43.813	43.749	0.5
Past purchase history*	0.156	0.452	23.9	0.156	0.173	1.3
Vehicle age	12.401	12.002	7.3	12.393	12.212	3.3
Original vehicle value*	10.002	9.879	22.8	10.003	10.010	-1.3
Total kilometers driven*	8.053	7.000	26.3	8.065	8.136	-1.8
Damage-free years	5.599	7.250	-22.5	5.599	5.698	-1.3
Limited comprehensive	0.323	0.339	-3.3	0.324	0.321	0.6
Comprehensive	0.254	0.356	-22.3	0.254	0.257	-0.7
Monthly premium	31.949	32.200	-1.5	31.956	32.316	-2.2
Total damage*	0.937	1.308	-14.7	0.937	0.966	-1.2
Travel insurance	0.023	0.053	-11.4	0.022	0.028	-2.4
Accident insurance	0.001	0.005	-6.4	0.001	0.002	-1.3
Legal aid insurance	0.008	0.026	-13.3	0.009	0.010	-1.2
Home insurance	0.026	0.066	-19.2	0.026	0.031	-2.5

Note: Standardized difference is the mean difference divided by the pooled standard deviation. We also use t-statistics for the comparisons, and the results are qualitatively similar. Variables with * are log-transformed.

Figure B-1: Propensity Score Density Before and After Matching on Covariates from *Treatment Group*



Notes: After matching, there exists a substantial overlap (right figure) of the propensity score densities for the two groups (treated: customer from comparison website, control: customers from less transparent channels), indicating matching helps generate reliable estimates to control observed heterogeneity that may drive the channel selection.

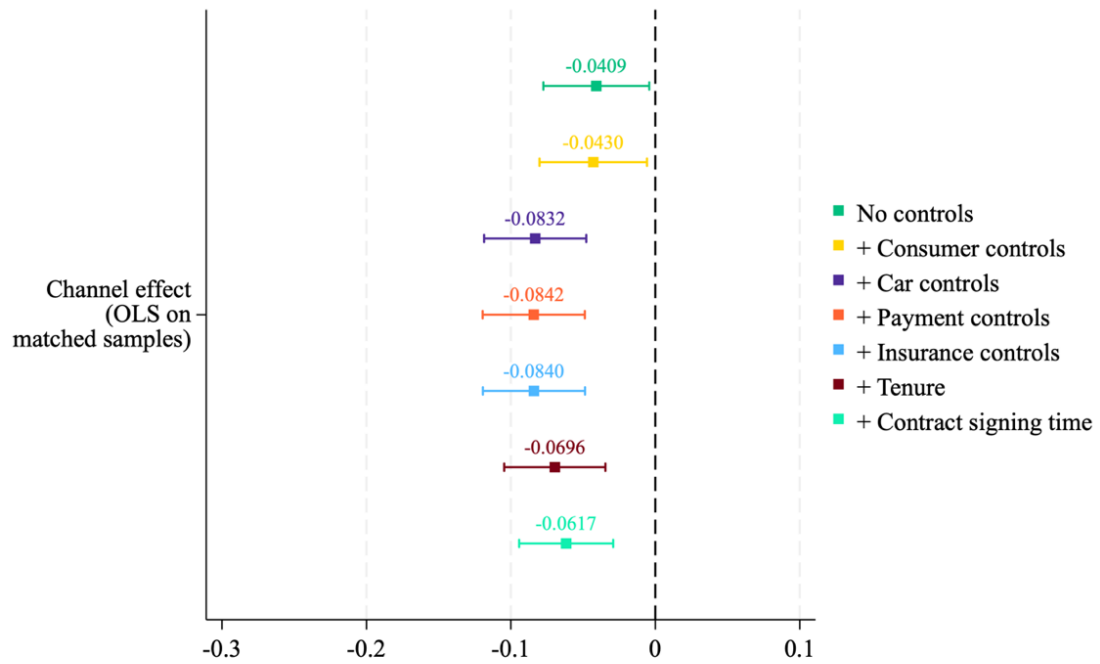


Figure B-2. OLS Estimates Using Propensity Score Matched Sample Using Field Data
(Treatment and Control Groups Matched on Covariates of Treatment Group)

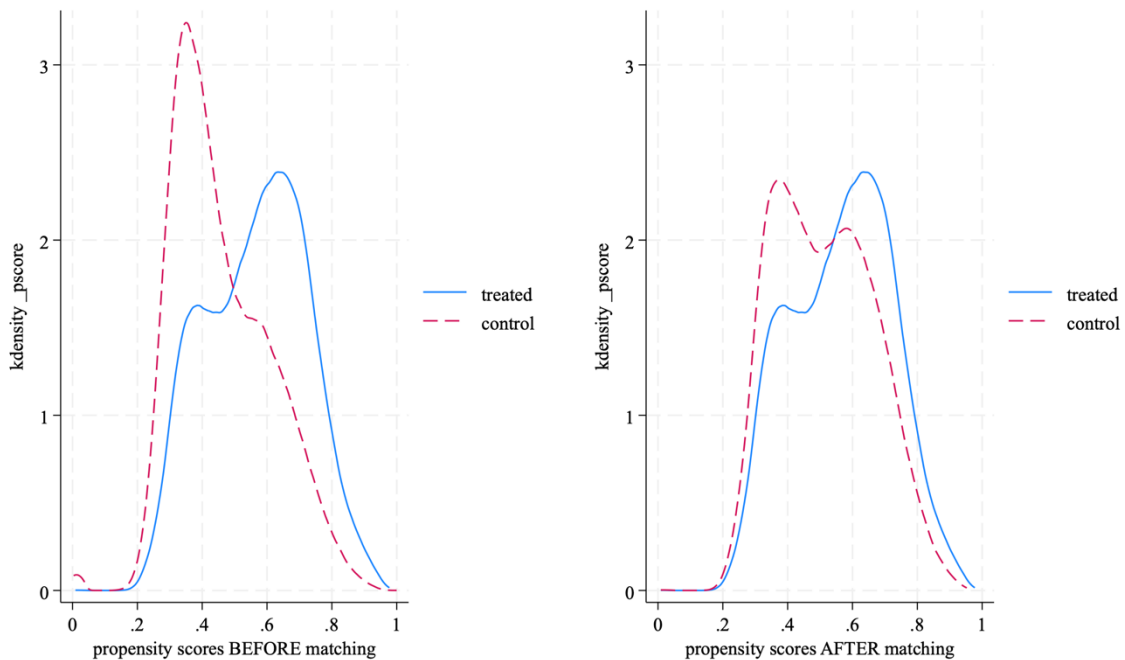
Notes: We also use matched samples on covariates of control group to replicate the estimation, and results remain consistent (See more details in [Table B-2](#) and [Figure B-2 below](#)).

Table B-2. Covariate Balance Before and After Matching on Covariates from *Control* Group (i.e., Customers Acquired from Other Channels)

	Before Matching			After Matching		
	Treatment Group	Control Group	Standard difference	Treatment Group	Control Group	Standard difference
Driver gender	0.309	0.336	5.7	0.336	0.336	0.1
Driver age	43.810	46.525	19.3	46.036	46.521	3.4
Past purchase history*	0.156	0.452	23.9	0.396	0.451	4.4
Vehicle age	12.401	12.002	7.3	11.989	12.002	0.2
Original vehicle value*	10.002	9.879	22.8	9.871	9.879	1.6
Total kilometers driven*	8.053	7.000	26.3	7.000	7.000	3.9
Damage-free years	5.599	7.250	-22.5	7.371	7.249	-1.7
Limited comprehensive	0.323	0.339	-3.3	0.334	0.339	1.0
Comprehensive	0.254	0.356	-22.3	0.364	0.356	-1.9
Monthly premium	31.949	32.200	-1.5	32.440	32.203	-1.4
Total damage*	0.937	1.308	-14.7	1.313	1.308	-0.2
Travel insurance	0.023	0.053	-11.4	0.048	0.053	2.2
Accident insurance	0.001	0.005	-6.4	0.004	0.005	1.1
Legal aid insurance	0.008	0.026	-13.3	0.021	0.026	3.7
Home insurance	0.026	0.066	-19.2	0.062	0.066	2.1

Note: Standardized difference is the mean difference divided by the pooled standard deviation. We also use t-statistics for the comparisons, and the results are qualitatively similar. Variables with * are log-transformed.

Figure B-3: Propensity Score Density Before and After Matching on Covariates from *Control* Group



Notes: After matching, there exists a substantial overlap (right figure) of the propensity score densities for the two groups (treated: customer from comparison website, control: customers from less transparent channels), indicating matching helps generate reliable estimates to control observed heterogeneity that may drive the channel selection.

**Table B-3. Comparison of Estimates from Approaches
Using Covariates from Treatment and Control Groups for Matching**

	LPM	Probit Model (marginal effect)	Cox Proportional Hazard Model
Using sample matching using covariates from treatment group (T)	-0.062*** (0.017)	-0.059*** (0.014)	-0.210*** (0.065)
Using sample matching using covariates from control group (C)	-0.053*** (0.014)	-0.047*** (0.013)	-0.217*** (0.056)
Z-stats for comparing estimates	0.009 (0.022)	0.012 (0.019)	-0.016 (0.085)

Notes: The results above from comparing estimates from two matching approaches demonstrate the observed effect is not driven by the self-selection of customers with innate differences (e.g., willingness to engage with comparison websites or other channels for insurance search and purchase).

B.2. Quasi-Natural Experiment and DiD Estimation

Table B-4. DiD Estimates of the Effect of Comparison Websites on Customer Churn

<i>Dependent Variable: Churn_{it}</i>	LPM	Probit Model	Cox Proportional Hazard Model
	(1)	(2)	(3)
Comparison _{it}	-0.118** (0.050)	-0.159** (0.073)	-0.256* (0.138)
All covariates	YES	YES	YES
Individual customer fixed effects	YES	NO	NO
Time period dummies	YES	YES	YES
# of observations	760	760	760

Notes: LPM includes customer fixed effects, whereas Probit and proportional-hazards models cannot fit a fixed effects estimation. Robust standard errors (clustered at the consumer level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-5. Comparison of Estimates from Different Identification Strategies

<i>Dependent Variable: Churn</i>	Baseline Regression	Matching	DID	Differences between Estimates (Standard Error)		
	(1)	(2)	(3)	(2)–(1)	(3)–(1)	(3)–(2)
LPM	-0.046*** (0.008)	-0.062*** (0.017)	-0.118** (0.050)	-0.016 (0.018)	-0.072 (0.051)	-0.056 (0.053)
Probit Model	-0.049*** (0.007)	-0.059*** (0.014)	-0.159** (0.073)	-0.010 (0.016)	-0.110 (0.073)	-0.100 (0.074)
Cox Proportional Hazard Model	-0.194*** (0.024)	-0.210*** (0.065)	-0.256* (0.138)	-0.016 (0.069)	-0.062 (0.140)	-0.046 (0.152)

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix C. Robustness Checks

C.1. Alternative Model Specifications

Linear Probability Models (LPMs) have two main limitations: (i) they assume a linear relationship between customer churn and acquisition channel, potentially leading to misspecification and sensitivity to data variations; and (ii) they may produce probability predictions outside the 0 to 1 range by treating the dependent variable (churn) as continuous, which can lead to significant underestimation or overestimation of effects. To address these issues, we use a Probit model.

The results shown in [Column 2](#) of [Table C-1](#) indicate that the estimate is consistent with the LPM in terms of sign and significance, with a calculated marginal effect of 4.7% ($p < 0.01$), closely aligning with the baseline estimate of 4.6% ($p < 0.01$). This suggests that the potential under- or over-estimations by the LPM are not a major concern in our modeling approach.

Both the LPM and Probit models do not account for the nonlinear effects of contract duration. There is a marked difference between customers who hold insurance policies for two months versus two years. Furthermore, our data are right-censored because we cannot track all customers until they churn.

To address these issues, we adopt the Cox proportional hazards model, which handles the non-linearity of contract duration and right-censoring ([Cox 1972](#)). The results from the Cox model, reported in [Column 3](#) of [Table C-1](#), align with our expectations (-0.198 , $p < 0.01$), corroborating the finding that customers from transparent channels tend to stay with the insurer longer. To ensure the robustness of our finding, we replicate the analysis but collapsed data for customers with multiple purchases to avoid over-representation in the sample, as shown in [Column 4](#) of [Table C-1](#). The estimates remain consistent.

Table C-1. Effects of Transparent Channel Acquisition Under Different Specifications

<i>Dependent Variable: Churn</i>	LPM	Probit	Cox Proportional Hazard Model:	
			<i>Continuous Time</i>	<i>Collapse Repeat Purchases</i>
	(1)	(2)	(3)	(4)
Comparison	-0.046*** (0.008)	-0.047*** (0.007)	-0.198*** (0.024)	-0.236*** (0.028)
covariates	YES	YES	YES	YES
# of observations	83,835	83,835	83,835	73,600

Notes: Robust standard errors (clustered at the individual customer level for columns 1, 2, 3, but not 4, as each observation in the sample for column 4 is a unique customer) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.2. Alternative Transparency Measures

Alternative Operationalization for Transparent Channels. We check whether the baseline estimates are robust to alternative measures of the independent variable ($Comparison_i$). Notably, 82.9% of customers using third-party comparison websites chose a dominant market player. Holding the control group (i.e., customers not from comparison websites) constant, we use a subsample of customers solely from the dominant comparison website as the treatment group to replicate the baseline LPM model. We then rerun this test using customers from other comparison websites as the treatment group. The observed effect still holds across different comparison websites (Table C-2), regardless of the specific features or functions of the comparison websites.

Table C-2. Comparison of Churn Probability across Different Comparison Website Measures

<i>Dependent Variable: Churn</i>	<i>Treatment groups: those who purchase insurance from</i>		
	All comparison websites	Comparison website “A”	Other comparison websites except “A”
	(1)	(2)	(3)
Comparison	-0.046*** (0.008)	-0.044** (0.008)	-0.056*** (0.010)
All covariates	YES	YES	YES
# of observations	83,835	77,189	52,074
Adjusted R-squared	0.381	0.384	0.418

Notes: Robust standard errors (clustered at the consumer level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Alternative Baseline Channels as the Control Group. To address whether our baseline estimates are robust to alternative baseline channels associated with lower information transparency, we vary the baseline channel by testing each of the less transparent channels individually. For instance, although the quote functions of the firm’s own website are similar to a third-party comparison website by providing product and price information, they show only the focal firm’s products. However, compared to the firm’s website or a comparison website, online advertising displays such information quite differently. The results consistently show the advantage of comparison sites in reducing customer churn (Table C-3).

Table C-3. Comparison of Churn Probability Using Different Baseline Channels

<i>Dependent Variable: Churn</i>	<i>Baselines Channels:</i>			
	All other channels	Firm’s own quote website	Inbound telemarketing	Online Ads
	(1)	(2)	(3)	(4)
Comparison	-0.046*** (0.008)	-0.018*** (0.003)	-0.032*** (0.004)	-0.069*** (0.011)

Notes: Robust standard errors (clustered at the customer level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Subjective Measurement of Information Transparency. We have so far used whether a customer was acquired from comparison websites to proxy for a transparent channel. While this approach allows us to distinguish between channels based on acquisition costs and other factors, it does not directly measure information transparency. To address this, we conduct a survey and introduced a subjective measure of information transparency in our survey

The survey involved 500 Dutch participants (drivers who reside in the focal country and have insurance purchase experiences) in the same auto insurance market to understand their channel preferences. Participants were asked to rate the likelihood of using multiple channels, their satisfaction, and the search and acquisition costs associated with different channels for their insurance purchases. Additional information about themselves, their insured vehicles, and insurance products was collected. [Table C-4](#) shows the main variables of the survey data, question items, and coding values.

Table C-4. Variables, Definitions, and Values in the Survey Data (N=500)

Variables	Question Items	Values
(1)	(2)	(3)
Transparency	Please rate the level of information transparency for the following channels for purchasing auto insurance.	0=the LEAST transparent; 10=the MOST transparent
Quote Duration	Please estimate how long it takes for you to go through the auto insurance quote process when using the following channels.	1=less than 5 min; 2=5-10 min; 3=10-20 min; 4=more than 20 min
Ease of Use	Please rate the ease of using the following channels.	0=very EASY 10=very DIFFICULT
Channel Satisfaction	Please rate how you are satisfied with the auto insurance quote process when using the following channels.	0=very DISSATISFIED, 10= very SATISFIED
Acquisition Cost	Please rate the likelihood of stopping the process without getting a quote when using the following channels.	0=LEAST likely 10=MOST likely
Multi-Channel Information Gathering	Please rate the likelihood of using the channel A to search for information and get an auto insurance quote, while using the channels B to make the actual purchase.	0=the LEAST likely 10=the MOST likely
Post Search	Please rate the likelihood of searching or obtaining information of other insurance providers and offerings after purchasing the current insurance.	0=LEAST likely 10=MOST likely
Driver Age	How old are you?	1=Below 25; 2=25-35; 3=36-45; 4=46-55; 5=56-65; 6=Above 65
Driver Gender	What's your gender?	0=Male; 1=Female
Vehicle Year of Production	What is the year of production of your car?	1=Before 1990; 2=1990...; 30=2018
Vehicle Price	What was the list price of your car?	1=less than €5000; 2=€5001~€10000...; 16= more than €100,000
Total Kilometer Driven	How many kilometers do you drive your car per year?	1=less than 5000; 2=5001~7500; ...; 9=more than 30,000
Insurance Type	What is the type of your current car insurance?	1=Liability; 2=Liability and Collision; 3=Comprehensive
Monthly Premium	What is your monthly premium for your car?	0=Less than €25; 1=€25 ~ €50; ...;12=More than €300
Other Insurance	Have you purchased any other types of insurance from the current car insurance provider?	0=No; 1=Travel accident insurance; 2=Home Insurance; 3=Life insurance; 4=Legal assistance insurance; 5=Others insurance

Although survey participants may not necessarily be customers of the focal company, the sample is not statistically significantly different from that in our main dataset. [Table C-5](#) demonstrates the representativeness of the survey data: main variables of customers, vehicles, and contracts follow similar distributions as those in the main dataset.

Table C-5. Comparison of Covariates in the Survey and Main Dataset

	Survey Data (N=500)	Main dataset (N=83,835)	Differences in Mean
	(1)	(2)	(3)
Driver gender	0.301 (0.460)	0.315 (0.464)	-0.014 (0.031)
Driver age	3.446 (1.177)	3.533 (1.377)	-0.087 (0.078)
Vehicle age	20.258 (7.004)	20.736 (7.367)	-0.478 (0.467)
Vehicle price	5.407 (4.275)	5.215 (2.589)	0.192 (0.284)
Total kilometers driven	5.305 (2.247)	5.575 (2.640)	-0.269 (0.150)
Insurance type	2.301 (0.815)	2.272 (0.817)	0.028 (0.054)
Monthly premium	3.841 (3.117)	3.601 (1.983)	0.239 (0.207)
Other insurance	1.243 (1.481)	1.196 (0.712)	0.047 (0.099)

[Table C-6](#) shows the mean and standard deviations of new variables and compares them between treatment and control channels. The values of these variables are generated through prediction (in a way that is similar to generating the subjective transparency measure) using a multinomial regression of new variables on covariates in the baseline analysis.

Table C-6. Replicating Main Analysis by Using Subjective Transparency Measure

Predicted Variable	Definition	Total (N=83,835)	From a comparison website (N=38,396)	From other channels (N=45,439)	Difference in Mean (4)-(5)
(1)	(2)	(3)	(4)	(5)	(6)
Multi-Channel Information Gathering	Likelihood of getting info. from other channels before purchase	5.219 (0.405)	5.116 (0.381)	5.339 (0.396)	-0.223 (0.549)
Search cost (quotation duration)	Duration of getting a quote via the acquisition channel	2.528 (0.183)	2.405 (0.137)	2.672 (0.115)	-0.266 (0.179)
Ease of use	Ease of use of the acquisition channel	4.741 (0.474)	4.761 (0.553)	4.724 (0.394)	0.037 (0.679)
Satisfaction	Satisfaction of the acquisition channel	6.193 (0.347)	6.378 (0.264)	5.976 (0.298)	0.403 (0.398)
Acquisition cost	Likelihood of stopping quoting without purchase	5.149 (0.474)	5.076 (0.335)	5.212 (0.558)	-0.136 (0.651)
Post search	Likelihood of searching other offerings after acquisition	5.476 (0.507)	5.564 (0.592)	5.401 (0.406)	0.163 (0.718)
Transparency	Subjective transparency of the acquisition channel	6.023 (0.327)	6.474 (0.352)	5.640 (0.304)	0.834* (0.465)

In the survey, we gave participants the definition of information transparency (i.e., the extent to which you have ready access to the information about insurance product/services, and price offerings from all available insurance *providers in the market*), and asked them “How would you rate the level of information transparency for different channels for purchasing auto insurance? (0 = least transparent, 10 = the most transparent)”

We then performed several steps to integrate the survey data into our analysis. First, we estimated a linear regression model to predict perceived transparency values in the survey data, mapping respondents’ characteristics to their transparency ratings. Using the estimated coefficients, we next generated a continuous transparency score for each customer in the field data based on the same covariates and standardized it to obtain the subjective measure $Transparency_i$. Finally, we replaced the original independent variable $Comparison_i$ with $Transparency_i$ and replicated the baseline analysis.

The results confirm that the main result: [Table C-7](#) cross-validates our baseline estimates using both transparency measures, *Transparency* and *comparison*, as the independent variable. *Transparency* denotes the degree of subjective transparency for each acquisition channel, which is an alternative measure for transparency.

Table C-7. Replicating Main Analysis by Using Subjective Transparency Measure

	<i>Dependent Variable: Churn</i>	
	(1)	(2)
Comparison	-0.046*** (0.008)	
Transparency		-0.040*** (0.008)
All covariates	YES	YES
Adjusted R-squared	0.384	0.383

Note: Robust standard errors (clustered at the consumer level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

C.3. Further Controlling Channel Selection Bias

Despite various model specifications and robustness checks, certain factors related to customer channel selection (either before, during, or after the initial insurance purchase) could still affect churn decisions. *Before acquisition*, customers may gather information from multiple channels. The estimate may be biased if such a multichannel interference exists. Additionally, some customers may be more (or less) satisfied with specific acquisition channels. This predisposition may influence their churn decisions. *During acquisition*, customers' time duration for getting a quote, ease of use, and the likelihood of stopping the quote without actual purchase may be different across channels. These factors represent the search costs of purchase via different channels. *After acquisition*, customers may search for or obtain information about alternative offerings. The extent of such post-acquisition activities may vary across customers who used different channels for initial purchase.

Using responses from survey described earlier in [Appendix C.2](#), we measured each participant's likelihood of using multiple channels, their satisfaction, and the search and acquisition costs associated with different channels for their insurance purchases. Additional information about themselves, their insured vehicles, and insurance products was collected.

We incorporated these new variables into our main dataset and then predicted and generated values for these variables (in a way that is similar to generating subjective transparency measurement) using multinomial regressions based on covariates linked to customers, vehicles, and contracts. These variables were used as covariates to replicate the baseline analysis. The results confirm that the main effect ([Table C-8](#)) is robust, even after controlling potential confounders (e.g., multichannel information gathering, satisfaction, and search and acquisition costs across channels). Notably, this approach is inherently limited compared with stronger causal identification strategies (e.g., difference-in-differences or experimental manipulation) and is intended only as a sensitivity *check*. (0.022)

Table C-8. Replicating Main Analysis by Controlling Potential Confounders

	<i>Dependent Variable: Churn</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comparison	-0.046*** (0.008)	-0.057*** (0.008)	-0.091*** (0.011)	-0.046*** (0.008)	-0.023*** (0.008)	-0.033*** (0.009)	-0.047*** (0.008)	-0.054*** (0.011)
Multi-Channel Information Gathering		0.068*** (0.005)						0.041*** (0.008)
Search cost (quote duration)			0.167*** (0.011)					0.073*** (0.005)
Ease of use				0 (0.009)				0.008 (0.005)
Satisfaction					-0.054*** (0.008)			-0.060*** (0.003)
Acquisition Cost						0.058*** (0.005)		0.052*** (0.004)
Post Search							0.021*** (0.004)	0.028*** (0.004)
All covariates	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.384	0.384	0.384	0.384	0.384	0.384	0.384	0.385

Notes: Robust standard errors (clustered at the consumer level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

C.4. Alternative Dependent Variable and Effect Persistence

Repeated Contract Renewals and Effect Persistence. The dependent variable we have so far used is a customer's churn status recorded in the cross-sectional dataset. It is also sensible to explore the persistent effect of comparison websites as a transparent acquisition channel on the repeated renewal/churn decision that a customer makes each time when a contract is expiring (i.e., each contract term is one year in our case). Empirically, we use several fixed lengths of observation periods (e.g., 1 year, 2 years, ..., after each customer was acquired) and adjust the measurement of the dependent variable to correspond to each customer's churn status at the end of each contract period. Then we replicate our baseline analysis.

Figure C-1 presents the results. We find the initial transparent acquisition channels continue to affect repeated renewal decisions over time, with the strongest effect on reducing churn in the first year and smaller but still statistically significant effect in the longer term (e.g., six years). This further supports the observed main effect and its persistence over time.

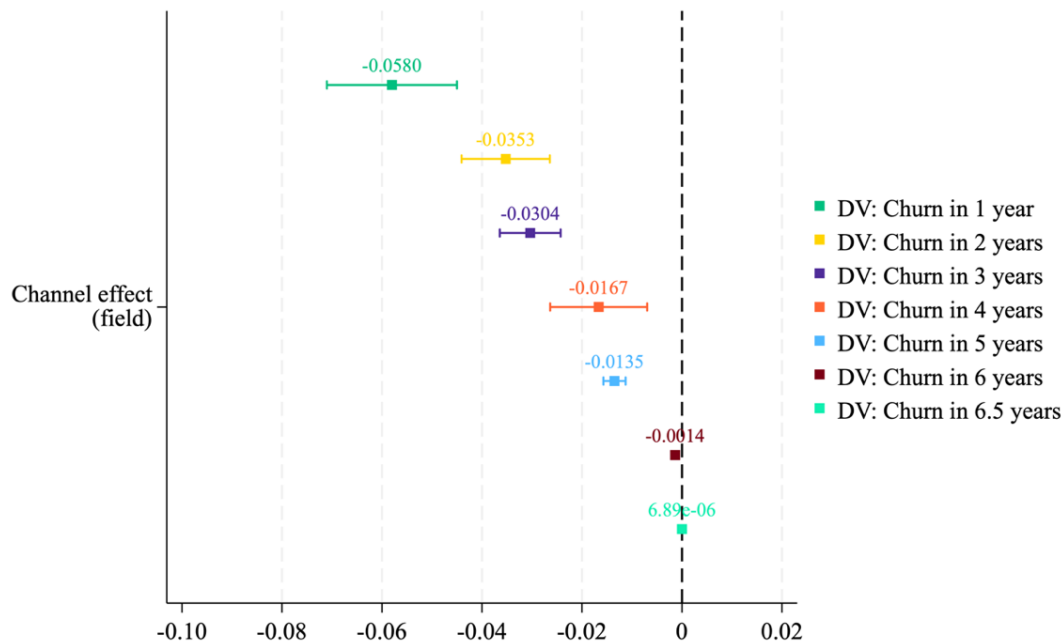


Figure C-1. Effects of Comparison Websites on Customer Churn in Each of the Subsequent Years after the Initial Contract Signup

C.5. Sampling Sensitivity

First-Time versus Repeat Customers. In the baseline estimation, we pooled repeat and first-time customers. Generally, repeat customers may be more loyal and thus less likely to churn than first-time buyers. To determine if the larger observed effect is primarily due to repeat customers' stronger loyalty, we analyze the subsamples of first-time and repeat customers separately. [Table C-9](#) reports the results, showing that the effects of using a comparison website for both customer groups are consistent in terms of sign and significance, albeit they are more pronounced for repeat customers ($0.078 > 0.041$, $p < 0.01$), supporting the notion that loyalty contributes to heterogeneity in the likelihood of churn, while transparent channels reduce churn across all customer types.

Table C-9. Comparison of Churn Probability between Those Who Purchased *Single* Auto insurance and Those Who Purchased *Multiple* Auto insurances respectively

<i>Dependent Variable: Churn</i>	<i>Customers who purchased</i>	
	<i>A single auto insurance</i>	<i>Multiple auto insurances</i>
	(1)	(2)
Comparison	-0.041 *** (0.009)	-0.078 *** (0.019)
All covariates	YES	YES
# of observations	67,071	16,764
Adjusted R-squared	0.395	0.351

Note: Robust standard errors clustered at the customer level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Alternative Sampling Strategies. While we used survival analysis to handle the right-censoring nature of our dependent variable, one would be concerned whether the observed effect is an artifact of the observation window (from late October 2008 to early March 2015). Thus, it is meaningful to assess if the main estimate is sensitive to different ways of truncating our field data. Specifically, we arbitrarily change the end of the sample period to November 2009 (so that customers acquired since October 2008 have at most one year of observed relationship history with the focal insurer), November 2010, ... March 2015 (i.e., the window in our baseline estimation), and we create subsamples accordingly to replicate the baseline analysis.

Figure C-2 below presents the results. The transparency effect remains negative across subsamples with different end dates, suggesting that the main finding is not driven by our use of the entire observation window.

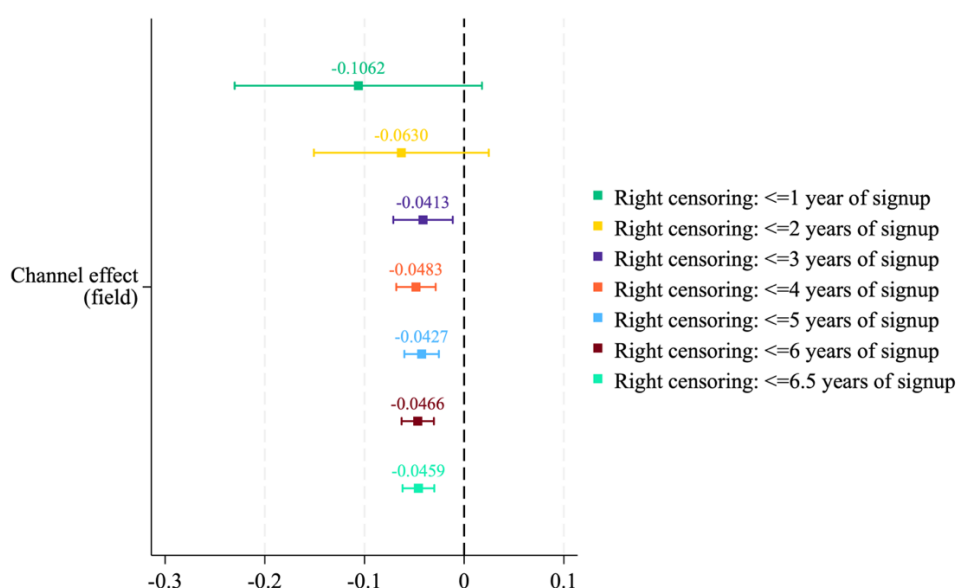


Figure C-2. Effects of Comparison Websites on Customer Churn Using Subsamples with Different End Dates

A related and important issue is the large sample issue that even minor or negligible effects appear statistically significant (with very small p-values) when estimated using a large sample (Lin et al. 2013). To address this, we use a consistent approach throughout the paper, namely using graphs to present point estimates and confidence intervals, rather than relying on statistical significance alone to report the results.

Here, we move one step further to test sensitivity of the effect estimates and p-values when the sample size increases. Specifically, we draw random samples of different sizes from our full dataset and we replicate our estimation.

Figure C-3 below presents the results, indicating that while standard error estimates from smaller samples are obviously larger, the point estimates are highly consistent across different sample sizes, which reassures us that the effect of comparison websites holds.

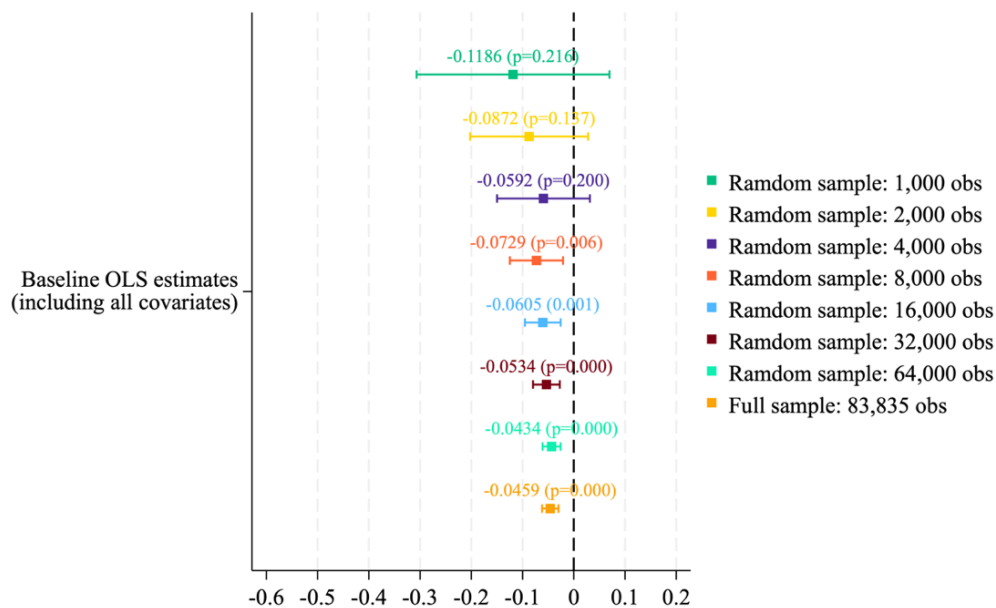


Figure C-3. Baseline OLS Estimates in Random Samples with Different Sizes

Appendix D. Mechanism Exploration Using an Online Experiment

D.1. Experimental Procedure

Table D-1. The Steps and Tasks Involved in the Randomized Controlled Experiment

Step	Item	Action
0	Introduction	Introduction to the Study and Informed Consent
1	Registration	Enter the Prolific ID
2	Familiarity	Familiarity check (see Table D-2)
3	Task description	Brief description of the insurance market (three insurers and their names), and the tasks (insurance search, contract sign-up, and contract renewal or switching providers decision).
4	Random assignment	Participants <i>randomly assigned</i> to one of two conditions: either comparison site (Treatment groups A1-A5) or the firms' own websites (Control group B), and they are told that they need to buy an auto insurance for their car.
5	Questions for a quote	Participants answer several questions (about themselves, their cars, their insurance coverage preferences) to get a quote (see Table D-3).
6	Search and compare insurance offerings	Participants in Group A see all offerings (with different transparency regarding price, product specification, and other customers' assessment; see Figures D-1 and D-2). Participants in Group B see the first insurers' offer, go through Step 5 again to see the second insurers' offer, and do it yet again to see the third insurer's offer (see Figure D-3).
7	Purchase confirmation	After seeing the three insurers' offerings (given the insurance coverage they chose), participants confirm their purchase by choosing the insurer.
8	Manipulation Checks	Manipulation checks, including the acquisition channel assigned, the insurance coverage chosen, and the insurer chosen for the purchase.
9	Post-purchase survey	Rate the importance of factors influencing their <u>purchase</u> decision (price, coverage, insurer name, review rating/#ratings/policy evaluation ratings if they are assigned to Groups A1-A5) and rate statements about search costs, choice confidence, product fit, product quality, and brand trust
10	Experience an accident scenario claim compensation	Participants <i>randomly assigned</i> to one of the three scenarios (see Appendix D.4) Scenario 1: accident on self, with damage compensation 100% claimed Scenario 2: accident on self, with damage compensation 50% claimed Scenario 3: accident on others, without need to claim damage
11	Post-scenario survey	After experiencing the scenario, participants rate again statements about search costs, choice confidence, product fit, product quality, and brand trust
12	Recheck the market offering	Participants are reminded that their insurance contract will be expired soon, they will need to decide to renew the contract or switch to other providers for their auto insurance. They are given the chance to recheck the market offerings. If choosing No, participants will automatically renew the contract with the current insurer they chose before (and skip Step 13). If choosing Yes, they will be directed to Step 13.
13	Make decision to stay or switch providers	Participants who chose Yes in Step 12 will be directed to the most transparent channel (Group A4) and view all offerings (with the insurance attributes including prices and coverage specifications unchanged). They will choose whether to stay or switch (and switch to which insurer)?
14	Post-renewal or -switching survey	Rate the importance of factors influencing their <u>renewal/switching</u> decision (price, coverage, insurer name, review rating/#ratings/policy evaluation ratings if they are assigned to Groups A1-A5)
15	Comments	Participants given the chance to comment on their experience in this study

D.2. Treatment Designs for Acquisition Channels

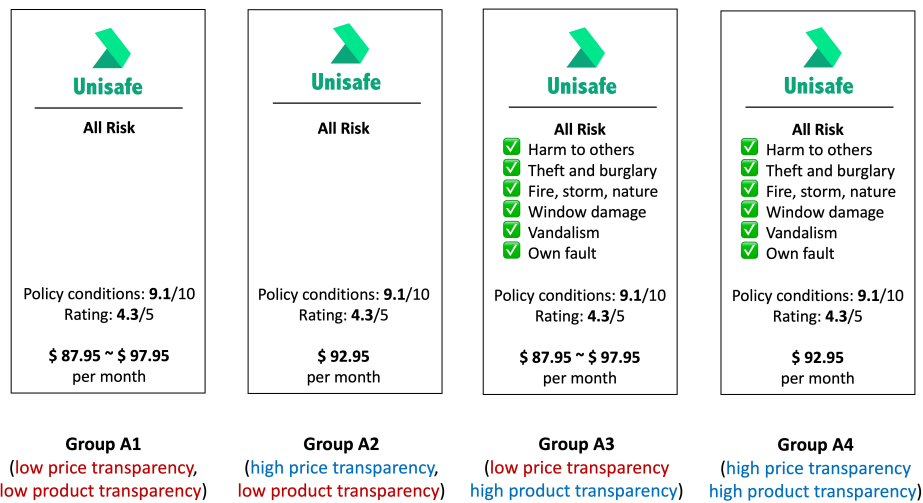


Figure D-1. Manipulation of the *Attribute* Transparency
(An example using insurer “Unisafe” and insurance coverage “All Risk”)

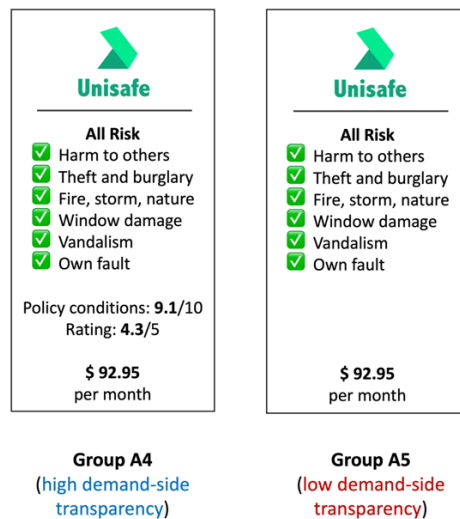
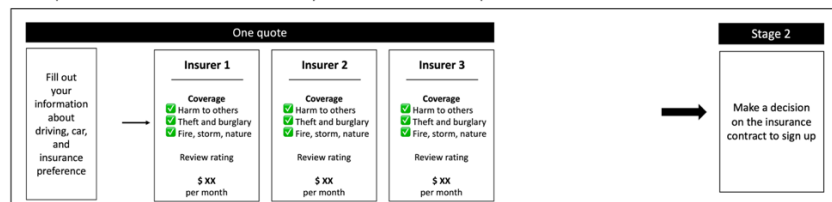


Figure D-2. Manipulation of the *Evaluative* Transparency
(An example using insurer “Unisafe” and insurance coverage “All Risk”)

Group A: Simultaneous Search with Only One Quote on a Comparison Site



Group B. Sequential Search with Three Quotes on Multiple Insurers' Websites

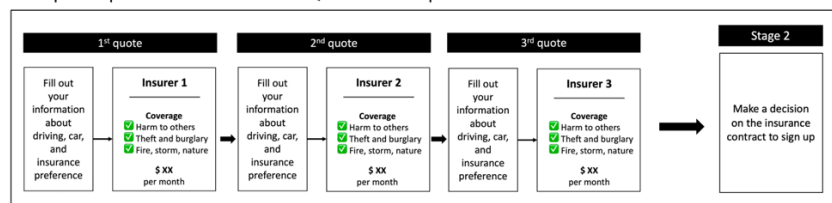


Figure D-3. Manipulation of *Comparative* Transparency
(Simultaneous vs. Sequential)

D.3. Survey Questions in the Online Experiment

Table D-2. Familiarity Checks

	Questions
1	How familiar are you with searching and shopping car insurance?
2	How often do you use the following channels when looking for car insurance (third-party comparison website; insurance company's website, talk to an insurance agent)?
3	How easy do you find the above channels when comparing insurance policies?
4	How sufficient is the information provided in these channels to help you make an informed decision?
5	How likely do you use the same channel again to search market offerings before your current insurance contract expires?
6	When was the last time you bought car insurance?
7	Please indicate the importance of each factor below affecting your insurance choices in the past (price, coverage specification, deductible amount, customer review, company reputation).
8	How confident are you that the insurance policy you choose is the best one for your needs?
9	How often do you switch car insurance providers?

Table D-3. Demographics Survey before Insurance Search and Shopping

Demographics	Survey items (choices in the parentheses)
Driving age	How many years have you driven a car? (0-3; 4-7; 8-10; over 10)
Driver Gender	What is your gender? (Male; Female)
Driver Age	How old are you? (18-25; 26-35; 46-55; 56-65; over 65)
Driving distance per year	How many kilometers do you drive per year? (Up to and including 7,500; 7,501 km – 10,000; 10,001 – 12,000; 12,000 – 15,000; 15,001 – 20,000; 20,001 – 25,000; 25,001 – 30,000; over 30,000)
Accumulated damage-free years	How many years since your last damage claim? (0; 1; 2; 3; 4-7; 8-10; more than 10)
Car type	What is the type of your car? (Hatchback; Sedan; SUV; Crossover; Coupe; Convertible; Pickup Truck)
Location	Where do you live? (Urban area; Suburban area; Rural area)
Insurance coverage preference	Choose your insurance coverage: Liability (i.e., “harm to others”) Liability plus (i.e., “harm to others,” “theft and burglary,” “fire, storm, nature,” “window damage.”) All risks (i.e., “harm to others,” “theft and burglary,” “fire, storm, nature,” “window damage,” “Vandalism,” “Own fault.”)

Table D-4. Switching Costs Measured After Purchase and Before Renewal Decisions

Switching Cost	Survey items (choices in the parentheses)
Search cost (Subjective; One's perception)	It was easy for me to find the information of different car insurance policies and prices.
	I did not spend a significant amount of time searching for car insurance policies and prices before finding the information I needed.
	I did not encounter difficulties or obstacles while searching for car insurance policies and prices.
Search cost (Objective; One's actual time spent on search)	The objective search cost is measured by the <i>time</i> spent on information filling and offer search and comparing before they confirmed purchase. The online survey questionnaire (on Qualtrics) has the feature to count seconds on activities, which we leverage for this measurement.
	Notably, customers assigned to a comparison website (Group A) filled their information to get a quote for just once and saw all three quotes, while those assigned to the firms' websites (Group B) had to fill the same information three times to get three quotes one by one and then made a purchase decision.
Choice confidence	I am confident about the choice I made earlier.
	I am sure that I have a good understanding of my insurance policy and its terms, and don't feel the need to review it again.
	I believe that this is the best insurance for me, and I don't feel the need to question about it.
	I have been with my current insurance company for a while, and I am still comfortable sticking with them even if other options are available.
Product Fit	The insurance policy still meets my specific needs and preferences.
	The coverage options and limits offered by the insurance policy adequately protect me in the event of an accident.
	The insurer is a good fit for me, given its pricing, reputation, and customer service.
	I believe Insurance policy is flexible enough to accommodate changes in my driving or personal circumstances.
Product quality	The insurance I purchased works exactly as it was described at the point of sale.
	The insurance performs as I expected it to perform.
	The insurance has maintained the same level of quality, or even exceeded it, compared to when I initially purchased it.
Brand trust	The insurer I chose is still the one I trust
	The insurer I chose still delivers what it promises

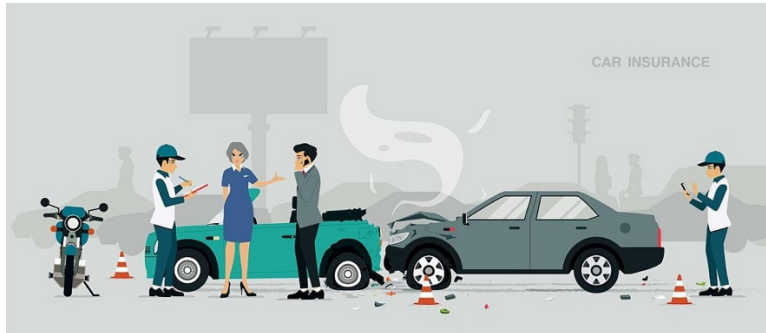
Notes: We adopt the existing definitions or survey instruments for above concepts in the literature and adapt them to the insurance context. For example, the survey items for product fit and quality are created based on [Hong and Pavlou \(2014\)](#), items for choice confidence are based on [Tsai and McGill \(2011\)](#) and references they cited, and items on brand trust are from [Rajavi et al. \(2019\)](#). All multi-item constructs exhibit strong internal consistency (Cronbach's $\alpha \geq 0.80$), indicating satisfactory measurement reliability ([Hair et al., 2018](#)). Specifically, choice confidence ($\alpha = 0.95$), product fit ($\alpha = 0.88$), product quality ($\alpha = 0.97$), brand trust ($\alpha = 0.96$), and subjective search cost ($\alpha = 0.81$) all exceed conventional thresholds.

D.4. The Design of Claim Scenarios

[Scenario 1]

Imagine that you are involved in a car accident. It is fortunate that nobody gets hurt. However, the accident cause identification result suggests that you are liable for this incident and should compensate for the property loss of another vehicle.

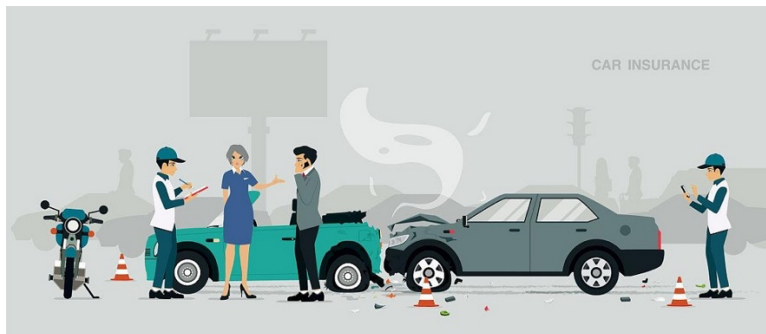
The estimated loss is **€8,000**. The insurance you purchased will cover **100%** of it.



[Scenario 2]

Imagine that you are involved in a car accident. It is fortunate that nobody gets hurt. However, the accident cause identification result suggests that you are liable for this incident and should compensate for the property loss of another vehicle.

The estimated loss is **€8,000**. The insurance you purchased will cover **50%** of it.



[Scenario 3]

Imagine that you witness a car accident nearby while you are driving. It is fortunate that nobody gets hurt. You recall that half a year ago you purchased insurance that could cover the loss if the incident happens to you.



Note: the figure is identical and displayed for all scenarios.

D.5. Experimental Data

Table D-5. Summary Statistics of Key Variables Used in the Experimental Analysis

ID	Variable	Mean (1)	SD (2)	Min (3)	Max (4)
1	Driving age	2.646	1.198	1	4
2	Driver Gender	0.487	0.500	0	1
3	Driver Age	2.194	1.064	1	6
4	Driving distance per year	3.041	1.908	1	8
5	Accumulated damage-free years	3.762	1.681	0	6
6	Damage claimed or not	0.367	0.482	0	1
7	Car type	2.029	1.356	1	7
8	Driver Location	1.595	0.667	1	3
9	Insurance coverage (liability=1, otherwise 0)	0.136	0.343	0	1
10	Insurance coverage (liability plus=1, otherwise 0)	0.390	0.488	0	1
11	Insurance coverage (all risk=1, otherwise 0)	0.474	0.499	0	1
12	Transparent channel (Comparison site=1, firm's quoting site=0)	0.726	0.446	0	1
13	Insurer chosen (Allinsur=1, otherwise 0)	0.408	0.492	0	1
14	Insurer chosen (AutoCare =1, otherwise 0)	0.383	0.486	0	1
15	Insurer chosen (Unisafe =1, otherwise 0)	0.209	0.407	0	1
16	Choice confidence (standardized)	-0.034	1.027	-3.117	3.700
17	Brand trust (standardized)	-0.024	1.030	-3.343	3.583
18	Product fit - perceived (standardized)	-0.032	1.032	-3.323	3.993
19	Product quality -percieved (standardized)	-0.034	1.025	-2.300	1.209
20	Objective search cost – time duration in search (standardized)	0.014	1.016	-0.888	18.330
21	Subjective search cost - perceived (standardized)	-0.026	0.975	-3.551	4.253
22	Scenario 1 (accident involved, 100% claimed)	0.359	0.480	0	1
23	Scenario 2 (accident involved, 50% claimed)	0.358	0.480	0	1
24	Scenario 3 (accident not involved)	0.283	0.451	0	1
25	Recheck (reconsider and recheck offers=1, otherwise 0)	0.572	0.495	0	1
26	Churn (switch to another provider=1, otherwise 0)	0.284	0.451	0	1

Notes: Variables 1–11 are control variables, with definitions and coding provided in [Table D-3](#). Variable 12 is the treatment variable, indicating whether a participant was assigned to a transparent channel. Variables 13–14 denote the insurer brand selected during the purchase task. Variables 16–21 represent switching-cost constructs, with definitions and coding in [Table D-4](#). Variables 22–24 capture post-purchase experience conditions, reflecting whether participants encountered different service quality outcomes. Variable 25 indicates whether a participant chose to recheck market offerings, and Variable 26 records whether the participant switched to another provider or remained with the previously chosen one. The final analytical sample includes 1,357 participants, after excluding 158 of the original 1,515 who failed at least one manipulation or attention check related to channel assignment, insurer names, or coverage choice.

Appendix E. Market Reconstruction Data and Market Standing Analysis

E.1. “Get a Quote” Simulation: Data and Analysis

To test the public signal boundary conditions (market standing of firms), we compiled a supplementary dataset from the dominant third-party comparison website responsible for 82.9% of the customers’ auto insurance comparisons in the studied country. Specifically, from the primary dataset, we selected a random sample of 1,200 customers. For each customer, we simulate her “get a quote” process: She goes to the comparison website, types in the information about herself, the vehicle to be insured, and preferred contract coverage, browses the search results, compares different offerings, selects the best offering, and purchases auto insurance. We manually collected information about product attributes and price offerings from all available insurance providers for each customer. This includes data about the brand names, prices (and rankings), policy conditions (e.g., 7.9 in the range of 0–10), customer review ratings (out of 10), and the number of ratings for all available insurance providers (see [Figure A-1](#) in [Appendix A](#)). In sum, the rich market information for each customer allows us to perform a rich examination of how the effect of transparency on customer churn decisions vary across firms’ market standings. [Table E-1](#) below reports the summary statistics of the new variables.

We acknowledge two limitations about the data and the setting. First, there is a time lag between the actual quote by a customer and the simulated quote by us, and the lag varies across individual customers. This is because our simulation was done between March and May 2018, but the customers may have been acquired anytime from 2008 to 2015. Therefore, the market-offerings information we collect for each customer may not precisely reflect the situations they faced at the time of purchase. It may also be overly strong to assume that such information remains fixed over time, as insurance providers often adjust their products and pricing in response to evolving demand and competitive conditions. However, the change is driven mainly by their market strategy, which does little to alter customers’ preferences of channels for insurance search and purchase. Moreover, in the studied insurance market, insurance providers make the same price and product updates across different channels. Hence, the inability to incorporate dynamic market information does not systematically bias our estimate. Nevertheless, to mitigate this concern, we leverage the time stamps in our data for both the date of the simulation and the date of the actual acquisition (shortly after the original quote) for the same customer. Specifically, we’ve included the time lag as a covariate to account for the aggregated market changes after the customer’s purchase, and the results remain consistent with those without considering this lag.

Second, and relatedly, as the market environment changes, the market position of the focal firm (regarding product and price ranking) might change. It is a legitimate concern that the new data cannot retrospectively capture the firm’s market position when a customer was initially acquired. Although the changes in the provider’s market position determine both customer acquisition and churn, they have little influence over which channel a customer makes her purchase. In other words, the provider’s market position, even if it changes, is independent of the customer’s channel selection and thus not a concern for the internal validity of our study. It may still raise concerns about external validity (or generalizability), especially when the focal provider is not a representative insurance company (e.g., a quality leader). However, this is not the case in our setting. Among 1,200 customers, the focal insurance provider was recommended among the top three providers with only a 17.1%

probability. This fact shows that the focal provider neither offers the best price nor is the highest rated provider at all times for all customers in the market.

Table E-1. Summary Statistics for Key Variables in the Market Reconstruction Data

	Mean	SD	Min	Max
	(1)	(2)	(3)	(4)
Within the top 3 recommended	0.171	0.377	0	1
Price advantage	0.868	0.339	0	1
Rating score advantage	0.608	0.489	0	1
# Ratings advantage	0.733	0.443	0	1
Price differential	10.059	11.154	-68.860	125.390
Rating score differential	0.216	0.326	-0.200	1.200
# Ratings differential	304.874	5,027.262	-14,110	3,955

Notes: Price advantage and rating score advantage, # ratings advantage, are dummy variables (0-1), measuring whether or not the price, product quality, and brand image of the insurance from the focal firm (we collaborated) is higher/lower, or better/worse than that of its best competitors. Price differential, rating score differential, # ratings differentials, are continuous variables, measuring to what extent the price, product quality, and brand image of the insurance from the focal firm (we collaborated) is higher/lower, or better/worse than that of its best competitors. The number of observations is 1,200.

E.2. Alternative Measures of Relative Market Standing

Brand image represents an additional dimension of public signals, capturing a firm's market visibility rather than product-level quality. Prior research shows that firms with stronger brand image typically command higher prices because market visibility, reputation, and brand awareness signal superior quality (Smith and Brynjolfsson 2001). Accordingly, the congruence between price and brand image serves another indicator of coherent market standing, whereas incongruence (e.g., a low-reputation brand charging a high price) introduces uncertainty about credibility and weakens transparency's effect.

To test this idea, we constructed an indicator variable for brand image, equal to 1 if the focal insurer had a larger number of customer ratings than its closest competitor. Using the reconstructed market data (We leverage the market re-construction data for this analysis, as the number of ratings vary a lot more than what we can manipulate in the experiment), we classified observations by the joint configuration of price and brand image (2×2) and re-estimate the transparency-churn relations for each of the four groups. The negative effect of transparency is stronger and statistically significant when the focal insurer signals congruence between price and image, but weaker and statistically insignificant otherwise (Figure E-1). This finding aligns with the signal-congruence condition under which transparency enhances customer retention and further corroborates H3a.

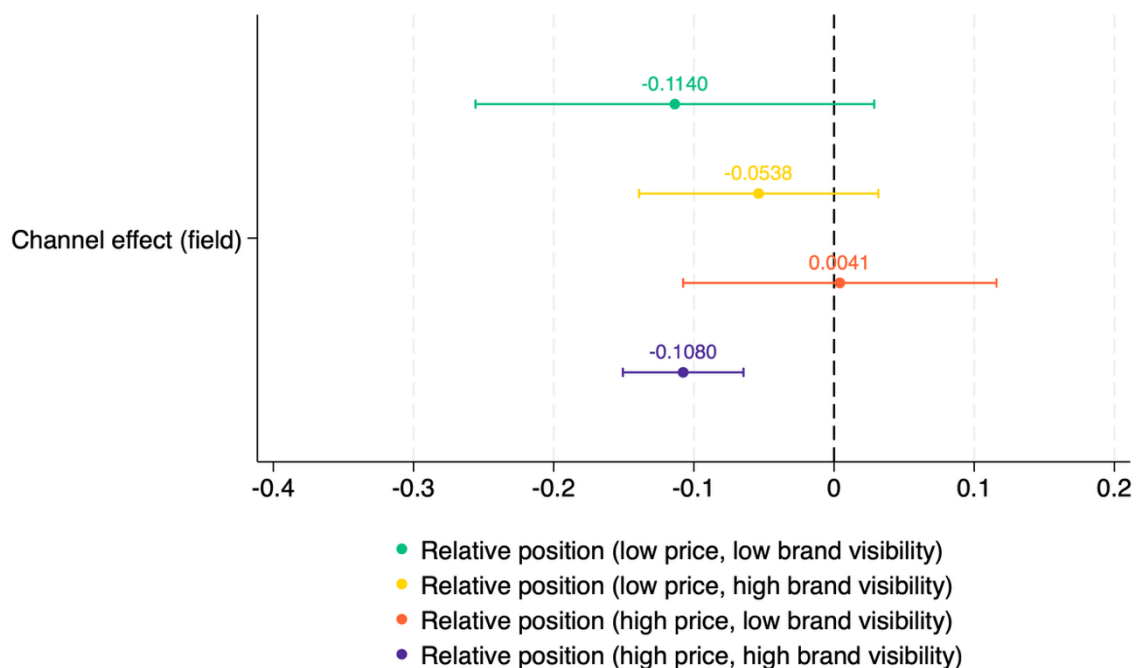


Figure E-1. Channel Effects by Firm's Market Position Measured by the Focal Firm's Relative (Dis)Advantage in Its Brand Image and Price (Field Evidence)

Appendix F. Heterogeneous Effects across Customers

The average effect estimates of the effect of comparison websites (the transparent channel) may not render enough detail to craft an effective churn management strategy. As insurance is a highly differentiated product, we examined the potential effect heterogeneity across customers' characteristics, vehicles, and contracts. [Figures F-1](#) and [F-2](#) present the results from the field and the experiment data analyses, respectively. We find the effect of comparison websites is negative and statistically significant across almost all customer subgroups, with few exceptions.

First, the effect is statistically insignificant for drivers who recently made claims (zero years of damage free) at acquisition. This is reasonable because these customers face high insurance prices due to their high accident risks, but market prices for them are likely to decrease significantly after the first year (as they are less likely to get involved in accidents twice in a year or so), increasing their propensity to churn.

Second, customers who purchased other insurance policies from the focal insurer are less likely to be influenced by comparison websites. This may be because these customers have built loyalty through buying multiple insurance products from the same firm, and the increased switching cost makes transparent channels less important to determine the churn likelihood of a customer.

Third, for some age groups (e.g., above 65 years), the effects are negative but not statistically significant ([Figure F-1](#)). This is likely due to the smaller sample sizes for these groups, which is indicated by larger standard deviations. Additionally, for the 25-35 age group, the lesser significance could be due to this demographic's heightened price sensitivity and lower tendency to rely on choice confidence for retention or switching decisions. Lastly, it is notable that point estimates from the experimental data are nearly all negative ([Figure F-2](#)), though more statistically insignificant results appear than those from the field. The larger standard errors are reasonable due to the large difference in sample size (1,515 customers in experiments versus 83,835 in the field) and higher representativeness in the field data.

In sum, the heterogeneous effect analysis suggests that while the overall effect of transparency channels on reducing churn is generalizable across various customer subgroups, there exist notable boundary conditions (e.g., multi-contract holders, damage-free years, age, kilometers driven) that influence the magnitude and statistical significance of this effect.

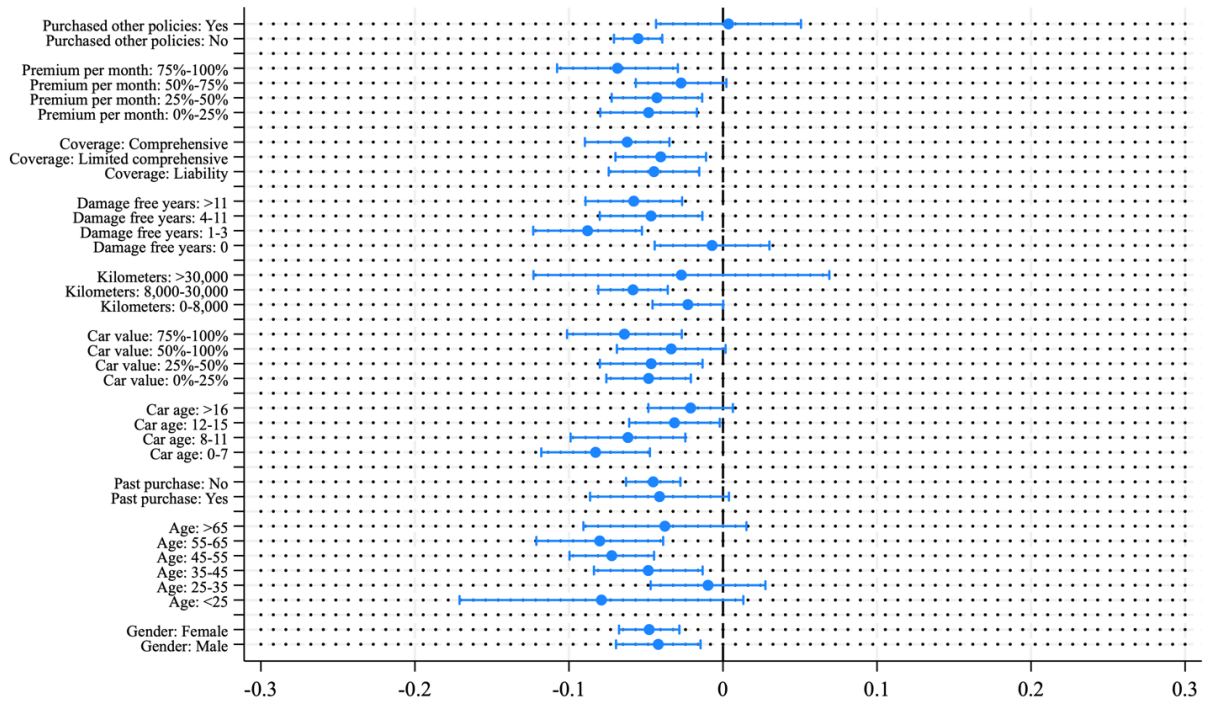


Figure F-1. Consumer Heterogeneity and Channel Effect Generalizability (Field Evidence)

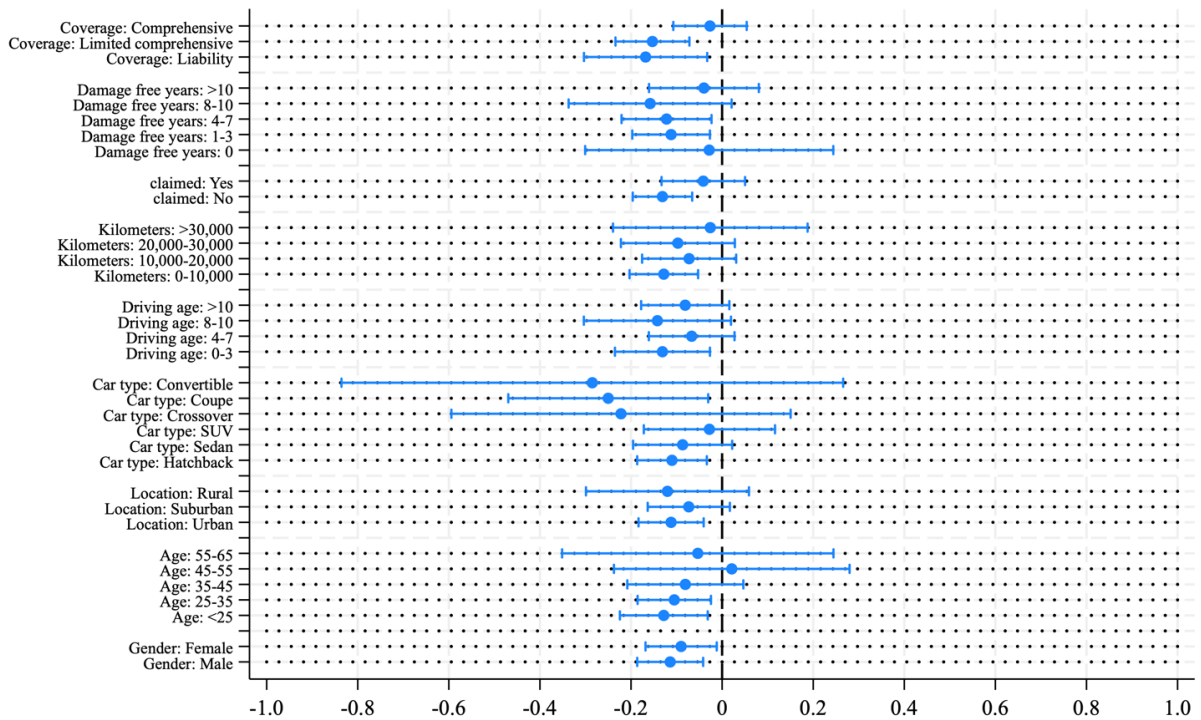


Figure F-2. Consumer Heterogeneity and Channel Effect Generalizability (Lab Evidence)

Appendix G. Economic Significance

The information transparency literature has discussed consumer surplus by using transparent/opaque channels ([Granados et al. 2010](#)), but the firm surplus from a transparency strategy perspective is unclear, despite its long-term economic value. We integrate the estimate of the churn-reducing effect and tenure of churners for back-of-the-envelope calculations of the economic value of allocating customers to transparent acquisition channels. We focus on three managerially relevant factors, *overall churn rate*, *number of churners (who could have been retained)*, and *profit gain/loss* if churners from low transparent channels had purchased from a comparison website.

We use the 95% confidence interval (CI) of the estimate (ρ_I) of the LPM estimator since it is easier to interpret. As the mean and standard errors of the reduced churn rate of customers acquired from comparison websites (relative to other channels) are 0.046 and 0.008, the 95% CI of a relative churn rate (ρ_I) is between 0.030 and 0.062.

First, we calculate the *overall churn rate* if churners were previously acquired from other channels, had they purchased from comparison websites, using the following formula: (the number of churners from a comparison website + $(1 - \rho_I) \times$ the number of churners from other channels) / (total number of customers). Hence, we obtain the *overall churn rate* in the interval between [46.24%, 47.09%], which is lower than the overall actual churn rate (47.91%) in our dataset, reduced by 0.82% ~ 1.67%.

Second, we calculate the *number of churners that could have been retained*, had churners who were previously acquired from other channels purchased via a comparison website using the following formula: total number of customers $\times$ *overall churn rate*. We obtain the *number of churners* in the interval of [38,764, 39,478], which would be smaller than the 40,169 actual churners in our sample. This indicates that a comparison website could have retained about 691 to 1,405 additional customers.

Third, we calculate the *profit gain or loss* if non-comparison churners could not have churned using the formula: number of saved churners $\times$ (total premium paid per person – total damage claimed per person) using data of customers from comparison (here note an interesting fact that each customer acquired from comparison websites is, on average, more profitable than those from other channels (€240.94 vs. €205.65) in their tenure with the focal insurer. Thus, we obtain the profit gain in the interval between [€166,489, €338,521] for the focal firm. Notably, the profit gain is a conservative estimate since it is calculated based on the extended customer lifetime value until March 2015 for customers acquired between October 2008 and March 2015. We expect those retained customers will stay longer with the firm and bring in additional profits. This finding further underscores the fact that our proposed transparent strategy (i.e., using comparison websites to acquire customers) is not only to retain customers but also to potentially retain profitable ones. This aligns with recent research that calls for profit-oriented customer retention (e.g., [Lemmens and Gupta 2020](#)).

REFERENCES (for Online Appendix):

- Cox DR (1972) Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* 34(2):187–202.
- Granados N, Gupta A, Kauffman RJ (2010) information transparency in business-to-consumer markets: Concepts, framework, and research agenda. *Information Systems Research* 21(2):207–226.
- Hair JF, Black WC, Babin BJ, Anderson RE (2018) *Multivariate Data Analysis* (8th ed.). Cengage.
- Hong Y, Pavlou PA (2014) Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research* 25(2):328–344.
- Lemmens A, Gupta S (2020) Managing churn to maximize profits. *Marketing Science* 39(5):956–973.
- Lin M, Lucas HC, Shmueli G (2013) Too big to fail: Large samples and the p -value problem. *Information Systems Research* 24(4):906–917.
- Rajavi K, Kushwaha T, Steenkamp JBE (2019) In brands we trust? A multicategory, multicountry investigation of sensitivity of consumers' trust in brands to marketing-mix activities. *Journal of Consumer Research* 46(4):651–670.
- Singer JD, Willett JB (1993) It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics* 18(2):155–195.
- Smith MD, Brynjolfsson E (2001) Consumer decision-making at an Internet shopbot: Brand still matters. *The Journal of Industrial Economics* 49(4):541–558.
- Tsai CI, McGill AL (2011) No pain, no gain? How fluency and construal level affect consumer confidence. *Journal of Consumer Research* 37(5):807–821.