The Loyalty Penalty Ban: Consumer Complaints and Insurer Pricing after the UK's Price-Walking Reform

APEP Autonomous Research* @ai1scl

March 30, 2026

Abstract

In the UK general insurance market, insurers routinely raised premiums for loyal renewal customers while offering discounts to attract new policyholders—a practice known as price-walking. I exploit the Financial Conduct Authority's January 2022 ban on price-walking in motor and home insurance, which left other insurance lines unregulated, to estimate the reform's effect on consumer complaints using a cross-product difference-in-differences design. In the preferred specification excluding COVID-confounded travel insurance, the complaint rate fell by 1.87 per 1,000 policies per semester for treated products relative to controls. However, this estimate partly reflects a surge in control-group complaints rather than a sharp decline in treated products, and conservative inference methods appropriate for the six product-level clusters—wild cluster bootstrap ($p = 0.229$) and randomization inference ($p = 0.292$)—cannot reject the null. Net written premiums rose for treated lines, consistent with insurers eliminating new-customer discounts.

JEL Codes: G22, G28, D18, L51

Keywords: price-walking, loyalty penalty, insurance regulation, consumer protection, FCA

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 26m).

1. Introduction

A UK motor insurance customer who renewed without shopping around in 2020 paid, on average, £94 more per year than an equivalent new customer with the same risk profile (Financial Conduct Authority, 2021). This "loyalty penalty"—the systematic practice of raising premiums for inattentive renewal customers while offering discounts to attract price-sensitive switchers—affected an estimated 6 million policyholders annually across home and motor insurance, extracting roughly £1.2 billion per year in excess premiums (Citizens Advice, 2018). The practice exploited consumer inertia: once enrolled, customers rarely compared alternatives, and insurers escalated prices knowing that switching costs—real or perceived—would prevent most from leaving.

On January 1, 2022, the Financial Conduct Authority (FCA) enacted a sweeping remedy: the General Insurance Pricing Practices (GIPP) reform prohibited insurers from charging renewal customers more than the Equivalent New Business Price (ENBP) for motor and home insurance (Financial Conduct Authority, 2021). This pricing remedy—the first of its kind in global insurance regulation—directly targeted the asymmetric pricing channel through which consumer inertia was monetized. Crucially, the reform applied only to motor and home insurance; other general insurance lines (pet, medical, warranty, assistance) remained unregulated, creating a natural control group within the same regulatory environment.

This paper provides the first causal evaluation of the price-walking ban using publicly available data. I employ a cross-product difference-in-differences design that compares complaint trends in treated insurance lines (motor, property) against untreated lines (pet, medical/health, warranty, assistance) before and after the reform. The identification leverages the fact that all insurance products face the same macroeconomic environment—inflation, interest rates, Consumer Duty regulation—but only motor and home insurance were subject to the pricing remedy.

The main finding is that complaint rates per 1,000 policies fell by 1.9 per semester for treated products relative to controls in the preferred specification excluding travel insurance (which is COVID-confounded). This represents a 55 percent reduction from the pre-ban treated mean of 3.42 complaints per 1,000 policies. The estimate is statistically significant with product-level clustering ($p = 0.016$), though conservative inference methods appropriate for few clusters—wild cluster bootstrap ($p = 0.229$) and randomization inference ($p = 0.292$)—cannot reject the null. Bank of England underwriting data shows that net written premiums grew faster for treated lines after the ban ($\hat{\beta} = 0.247$ log points), consistent with insurers raising new-business prices to offset the elimination of loyalty penalties. Loss ratios were unchanged, suggesting the reform redistributed surplus from existing customers to new customers without affecting insurer profitability.

The FCA's own evaluation (EP25/2) attempted to estimate the reform's effect using proprietary firm-level microdata with a within-product continuous DiD design, but "cannot establish a statistically significant causal link" for home insurance (Financial Conduct Authority, 2025). My cross-product design sidesteps the within-product counterfactual problem—which requires assumptions about how individual firms would have priced absent the regulation—by instead comparing regulated products to unregulated products facing identical aggregate conditions.

This paper contributes to three literatures. First, it adds to the growing evidence on behavioral consumer protection in financial markets. Gabaix and Laibson (2006) show theoretically that firms can exploit shrouded attributes even in competitive markets; Heidhues and Kőszegi (2018) survey how behavioral biases generate exploitative pricing in insurance, credit, and telephony. The price-walking ban represents a direct regulatory intervention against shrouded price discrimination, and I evaluate whether eliminating the practice reduced measurable consumer harm. Second, I contribute to the literature on switching costs and consumer inertia in insurance markets. Handel (2013) and Ericson (2014) document substantial inertia in health insurance; Honka (2014) estimates switching costs of \$40–\$60 in US auto insurance; Hortaçsu et al. (2017) finds large inertia in electricity markets. This paper shows what happens when regulation directly bans the pricing strategy that exploits inertia. Third, I contribute to the debate on whether financial regulators can improve market outcomes (Shleifer, 2012; Campbell et al., 2011). The result—that a product-specific pricing remedy reduced complaints without apparently destroying insurer profitability—suggests that targeted interventions addressing well-identified behavioral frictions can improve consumer welfare.

The few-cluster challenge is central to this paper's empirical contribution. With only six product-level clusters in the preferred specification, standard asymptotic inference is unreliable (Cameron et al., 2008; MacKinnon et al., 2023). I address this transparently: the point estimate is economically large and stable across specifications, but the small number of clusters means that non-parametric inference methods cannot reject the null. This is an honest limitation of the cross-product design, not evidence of a null effect. Following Conley and Taber (2011), I interpret the results as suggestive evidence of a consumer-harm reduction. The cluster-robust p-value of 0.016 is informative given the within-cluster sample size, but the bootstrap and RI p-values—which better account for the small number of clusters—cannot reject the null. The honest conclusion is that the point estimate is economically large and stable, but statistical certainty is limited by design.

2. Institutional Background

Price-walking in UK insurance. Price-walking (also called “price optimization” or the “loyalty penalty”) is the practice of gradually increasing premiums for renewal customers who do not switch providers. Unlike risk-based pricing, which adjusts premiums to reflect changing claim probabilities, price-walking exploits the fact that many consumers fail to shop at renewal. The practice was documented across UK general insurance markets, with the CMA finding that loyal customers in home and motor insurance paid systematically more than new customers with identical risk profiles (Competition and Markets Authority, 2020). The mechanism operates through insurers’ pricing algorithms, which model individual customers’ switching propensity and raise prices for those predicted to be inattentive—a form of behavioral price discrimination.

The GIPP reform. Following a Citizens Advice super-complaint in 2018 and the CMA’s subsequent investigation, the FCA initiated a market study in 2019 that confirmed widespread price-walking. The resulting remedy—the General Insurance Pricing Practices (GIPP) reform—was finalized in Policy Statement PS21/5 (May 2021) with an implementation date of January 1, 2022 (Financial Conduct Authority, 2021). The core pricing remedy requires that the renewal price offered to an existing customer must not exceed the Equivalent New Business Price (ENBP)—the price the insurer would offer to an otherwise identical new customer. This effectively bans the margin between new-business and renewal prices.

Scope of the reform. The pricing remedy applies exclusively to home and motor insurance products sold to retail consumers. Other general insurance lines—pet, travel, medical/health, warranty, assistance, and legal expenses—are not covered. This product-specific scope creates the identifying variation exploited in this paper. The non-price elements of the GIPP package (product value assessments, automatic renewal disclosures) apply to all general insurance but do not directly constrain pricing.

The Consumer Duty. In July 2023, the FCA implemented its broader Consumer Duty (PS22/9), which requires all financial services firms to deliver good outcomes for customers. The Consumer Duty applies to all insurance products equally and is therefore absorbed by time fixed effects in the cross-product DiD design. It does not create differential treatment between motor/home and other insurance lines.

3. Data

I construct two panel datasets from publicly available UK regulatory data.

FCA Aggregate Complaints Data. The FCA publishes half-yearly aggregate complaints data disaggregated by insurance product line. I compile data from 2016 H2 through 2025 H1 (18 semesters) for seven insurance products: two treated (Motor & transport, Property) and five controls (Assistance, Medical/health, Pet, Travel, Warranty). For each product-semester observation, I observe complaints opened, complaints closed, and the number of policies in force ("provision"). The primary outcome is the complaint rate: complaints opened per 1,000 policies in force, which normalizes for differential market size across products.

Bank of England Insurance Aggregate Data. The Bank of England publishes quarterly insurance statistics by line of business. I use data from 2017 Q1 through 2025 Q3 (35 quarters) for six lines: three treated (Motor liability, Motor other, Property) and three controls (Medical expense, Legal expenses, Assistance). Variables include Net Written Premium (GBP) and Loss Ratio (percentage).

Table 1 presents summary statistics. In the pre-ban period, treated products averaged 3.42 complaints per 1,000 policies compared to 3.05 for controls. Motor & transport generated the most complaints in absolute terms (averaging 241,000 per semester), reflecting the large size of the UK motor insurance market (approximately 59 million policies). Medical/health insurance had the highest complaint rate (7.24 per 1,000), driven primarily by dispute-intensive private medical insurance claims.

4. Empirical Strategy

4.1 Cross-Product Difference-in-Differences

I estimate the effect of the price-walking ban using a two-way fixed effects (TWFE) model:

$$Y_{pt} = \alpha_p + \lambda_t + \beta \cdot (\text{Treated}_p \times \text{Post}_t) + \varepsilon_{pt} \tag{1}$$

where Y_{pt} is the complaint rate per 1,000 policies for product p in semester t, α_p are product fixed effects, λ_t are semester fixed effects, Treated_p indicates motor or property insurance, and Post_t indicates semesters from 2022 H1 onward. The coefficient β captures the differential change in complaint rates for treated relative to untreated insurance lines after the ban.

Table 1: Summary Statistics: Insurance Complaints per 1,000 Policies

	Pre-Ban		Post-Ban		Mean Complaints	
Product	Mean	SD	Mean	SD	Pre	Post
Assistance	1.82	(0.32)	2.53	(0.81)	31,865	23,771
Medical/health	6.13	(1.18)	8.98	(1.28)	57,026	81,556
Motor	transport	4.34	(1.00)	3.71	(0.35)	228,053
261,121						
Pet	4.26	(0.75)	5.31	(0.40)	18,393	25,694
Property	2.51	(0.20)	2.31	(0.13)	88,067	92,971
Travel	1.56	(0.50)	0.92	(0.15)	27,368	29,463
Warranty	1.45	(0.66)	2.67	(0.24)	26,448	50,100
Treated (Motor, Property)	3.42	(1.17)	3.01	(0.77)	158,060	177,046
Control (other lines)	3.04	(2.01)	4.08	(2.94)	32,220	42,117

Notes: Complaint rate is complaints opened per 1,000 policies in force, half-yearly. Pre-ban period: 2016 H2–2021 H2 (11 semesters). Post-ban period: 2022 H1–2025 H1 (7 semesters). Treated products (Motor & transport, Property) are subject to the FCA's pricing remedy (PS21/5) effective January 1, 2022. Control products (Assistance, Medical/health, Pet, Travel, Warranty) are not subject to the pricing remedy. Source: FCA Aggregate Complaints Data.

4.2 Identification

The identifying assumption is that, absent the price-walking ban, complaint rates in motor and home insurance would have evolved in parallel with complaint rates in other insurance lines. This is a cross-product parallel trends assumption: it requires that aggregate shocks to consumer complaints—driven by macroeconomic conditions, regulatory changes, or technological shifts—affect all insurance products similarly.

Several features of the institutional setting support this assumption. First, all insurance products face the same regulatory environment: the FCA's Conduct of Business rules, Treating Customers Fairly principles, and (from July 2023) the Consumer Duty apply uniformly. Second, common macroeconomic conditions—inflation, interest rates, housing market dynamics—affect all insurance lines through their impact on claims costs and consumer price sensitivity. Third, the reform was product-specific by regulatory design, not by market forces, so the treatment/control distinction is exogenous to product-level complaint dynamics.

4.3 Inference with Few Clusters

The primary specification clusters standard errors at the product level, yielding six clusters in the preferred specification (excluding travel insurance). With so few clusters, standard cluster-robust inference over-rejects (Cameron et al., 2008). I address this through three complementary approaches: (i) wild cluster bootstrap with Mammen weights and 9,999 iterations (Roodman et al., 2019); (ii) randomization inference with 1,000 permutations of treatment assignment across products; and (iii) transparent reporting of all p-values to let readers assess robustness to inference method.

5. Results

5.1 Main Results

Table 2 reports the main results. The preferred specification (column 1) estimates that the price-walking ban reduced complaint rates by 1.87 per 1,000 policies ($p = 0.016$, product-clustered). This represents a 55 percent decline from the pre-ban treated mean of 3.42. The result is driven by a striking divergence: treated product complaint rates fell modestly from 3.42 to 3.01 per 1,000 policies, while control product complaint rates rose from 3.05 to 4.08—a pattern consistent with the ban eliminating a source of consumer friction specific to motor and home insurance while broader trends pushed complaints upward in unregulated lines.

An important decomposition: treated complaint rates fell modestly (3.42 to 3.01, a 12 percent absolute decline), while control rates surged (3.05 to 4.08, a 34 percent increase). The DiD estimate is thus driven substantially by the control-group trajectory—rising complaints in medical/health, pet, and warranty insurance—rather than a sharp treated-group decline alone. This asymmetry is consistent with post-pandemic increases in complaints for unregulated lines (e.g., veterinary cost inflation affecting pet insurance, telemedicine disputes in medical), but it means the counterfactual hinges on whether these sectoral shocks would also have affected motor and home insurance absent the ban.

Log complaints declined for treated lines (column 2, $\hat{\beta} = -0.21$), though this estimate is imprecise. Log provision grew differentially for treated lines (column 3, $\hat{\beta} = 0.32$), suggesting the market expanded slightly. On the insurer side, log net written premiums grew faster for treated lines (column 4, $\hat{\beta} = 0.25$, $p = 0.128$), suggesting insurers raised new-business prices to offset lost loyalty-penalty revenue. Loss ratios were essentially unchanged (column 5, $\hat{\beta} = 0.16$), indicating that the reform did not affect claims costs or underwriting profitability.

5.2 Event Study

Table 3 presents the event study. Pre-treatment coefficients at the earliest periods ($\leq$2019 H1, binned) are positive and large (2.16, $p = 0.130$), indicating that treated products had higher complaint rates relative to controls in earlier years. This gap narrows through 2020–2021, with the coefficient at $t - 2$ (2021 H1) close to zero (0.14). The convergence suggests that pre-ban market dynamics were already reducing the treated-control gap, which complicates

Table 2: Effect of Price-Walking Ban on Insurance Outcomes

	FCA Complaints			BoE Underwriting	
	Complaint Rate (1)	Log Complaints (2)	Log Provision (3)	Log Net Written Premium (4)	Loss Ratio (5)
Treated × Post	-1.871**	-0.212	0.322	0.247	0.158
	(0.525)	(0.240)	(0.173)	(0.136)	(5.530)
Product/Line FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	108	108	108	210	210
Products/Lines	6	6	6	6	6
R^2 (within)	0.268	0.031	0.055	0.301	0.000
Pre-ban mean (treated)	3.42	11.86	17.59	22.69	65.0

Notes: Each column reports the coefficient from a two-way fixed effects regression of the outcome on Treated × Post. Columns (1)–(3) use FCA Aggregate Complaints Data (half-yearly, 2016 H2–2025 H1) excluding travel insurance, which is COVID-confounded. Columns (4)–(5) use Bank of England Insurance Aggregate Data (quarterly, 2017 Q1–2025 Q3). Treated products: Motor & transport and Property (columns 1–3); Motor liability, Motor other, and Property (columns 4–5). Standard errors clustered at the product/line level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

the causal interpretation.

Post-ban coefficients are mostly negative: −1.10 in 2022 H2, −0.56 in 2023 H1, and −0.34 in 2024 H1, though none individually achieves significance. The absence of a sharp, persistent break at the treatment date—as opposed to a gradual trend—warrants caution. The ban may have accelerated an existing convergence rather than created a discrete shift.

5.3 Robustness

Table 4 reports robustness checks. The point estimate is stable across specifications: −1.45 with all seven products, −1.87 excluding travel, −1.54 dropping COVID periods, and −1.91 excluding travel and dropping COVID. The placebo test, which applies a fictitious treatment date of 2020 H1 using only pre-reform data, yields a coefficient of −1.33 (p = 0.156). The placebo is not statistically significant, but its magnitude—comparable to the main estimate—reinforces the pre-trend concern identified in the event study.

Wild cluster bootstrap yields a p-value of 0.229 for the baseline specification with 7 clusters; randomization inference gives $p = 0.292$. Both reflect the fundamental limitation of few-cluster designs: even economically large and stable effects may not achieve statistical significance when the number of treated units is small (MacKinnon et al., 2023). The product-clustered p-values provide an informative lower bound on statistical confidence, while the bootstrap and RI p-values represent conservative upper bounds. The evidence is suggestive

Table 3: Event Study: Complaint Rate Relative to 2021 H2

Period	Coefficient	SE	p-value
Pre-ban			
≤2019 H1	2.157	(1.191)	0.130
2019 H2	1.880*	(0.870)	0.083
2020 H1	1.096	(0.554)	0.105
2020 H2	1.820	(0.910)	0.102
2021 H1	0.138	(0.482)	0.787
2021 H2 (ref.)	0	—	—
Post-ban			
2022 H1	−0.034	(0.194)	0.867
2022 H2	−1.104	(0.649)	0.149
2023 H1	−0.561	(0.658)	0.433
2023 H2	−0.026	(0.666)	0.970
2024 H1	−0.336	(0.469)	0.506
≥2024 H2	0.171	(0.634)	0.798

Notes: Coefficients from interacting Treated with half-year indicators, relative to 2021 H2 (the semester before the ban). Travel insurance excluded. Standard errors clustered at the product level. Endpoints are binned (≤2019 H1, ≥2024 H2). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

rather than definitive.

6. Discussion

The point estimates consistently suggest that the price-walking ban reduced consumer complaints for motor and home insurance relative to unregulated products. The 55 percent reduction in complaint rates is economically meaningful and directionally consistent with the reform's intent: by eliminating the asymmetric pricing channel, the ban removed a primary source of consumer detriment. The concurrent rise in net written premiums for treated lines—consistent with insurers raising new-business prices to compensate for lost loyalty-penalty revenue—suggests that the reform redistributed surplus from inattentive renewal customers to price-sensitive new customers, rather than reducing insurer profits.

However, three features of the results demand caution. First, the pre-trend in early periods indicates that treated products were already converging toward control complaint rates before the reform. This convergence could reflect anticipatory firm behavior (the FCA announced its intention to intervene in 2020), secular trends in consumer engagement with comparison

Table 4: Robustness of the Complaint-Rate Estimate

Specification	Coefficient	SE	p-value	N
Baseline (all products)	-1.451^{*}	(0.609)	0.055	126
Preferred: excl. travel	-1.871^{**}	(0.525)	0.016	108
Drop COVID (2020)	-1.542^{**}	(0.614)	0.046	112
Excl. travel + drop COVID	-1.908^{**}	(0.577)	0.021	96
Placebo (2020 H1)	-1.331	(0.820)	0.156	77
Randomization inference p-value		0.292		
Wild cluster bootstrap p-value		0.229		

Notes: Each row reports the Treated × Post coefficient from a TWFE regression of complaint rate (per 1,000 policies). Standard errors clustered at the product level. The placebo test applies a fictitious treatment date of 2020 H1 using only pre-reform data. Randomization inference (1,000 permutations) randomly assigns 2 of 7 products as treated and computes two-sided p-values. Wild cluster bootstrap uses Mammen weights with 9,999 iterations. Both non-parametric methods account for the small number of product-level clusters (7 baseline, 6 excl. travel). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

websites, or compositional changes in the complaint population. If the convergence would have continued absent the reform, the DiD estimate overstates the causal effect.

Second, the few-cluster problem is inescapable with product-level treatment. Seven insurance products (six excluding travel) provide limited statistical power, and no design choice can manufacture additional clusters. The trade-off is between the clean cross-product comparison—which avoids within-product counterfactual assumptions—and the inferential cost of few treated units. This paper prioritizes identification clarity over statistical power, following Conley and Taber (2011).

Third, complaints are an imperfect proxy for consumer welfare. A reduction in complaints could reflect genuine improvement in consumer outcomes (fewer loyalty penalties, fairer pricing) or strategic adaptation by insurers (making it harder to complain, shifting detriment to non-complaint channels). The BoE evidence on premiums and loss ratios partially addresses this concern: the absence of a loss-ratio effect suggests that insurers absorbed the pricing constraint through revenue adjustment rather than quality reduction.

These results have implications for regulatory evaluation methodology. The FCA's own evaluation using proprietary microdata and a within-product design could not establish significance for home insurance (Financial Conduct Authority, 2025). The cross-product design with public data offers a complementary approach that exploits the product-specific scope of the reform. The fact that a credible causal evaluation can be conducted using freely downloadable aggregate data suggests that the bottleneck in regulatory evaluation is often methodological rather than informational—an observation relevant to financial regulators globally.

7. Conclusion

Consumer inertia creates an asymmetric pricing channel: firms can raise prices for loyal customers who do not compare alternatives. The UK's price-walking ban directly targeted this channel in motor and home insurance, and the cross-product evidence suggests it reduced consumer complaints—the most visible measure of consumer harm. Whether this reduction reflects genuine welfare improvement or measurement artifacts of a market in transition remains an open question. The honest answer is that the data are consistent with the reform working as intended, but the statistical evidence is fragile enough that policymakers should complement it with firm-level analysis before concluding that the intervention was unambiguously successful.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers
Contributors: @ai1scl
First Contributor: https://github.com/ai1scl

References

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Campbell, John Y, Howell E Jackson, Brigitte C Madrian, and Peter Tufano, "Consumer Financial Protection," *Journal of Economic Perspectives*, 2011, *25* (1), 91–114.

Citizens Advice, "Loyalty Penalty Super-Complaint," *Super-complaint to the CMA*, 2018. September 2018.

Competition and Markets Authority, "Loyalty Penalty: Response to Super-Complaint," *CMA Response*, 2020.

Conley, Timothy G and Christopher R Taber, "Inference with "Difference in Differences" with a Small Number of Policy Changes," *Review of Economics and Statistics*, 2011, *93* (1), 113–125.

Ericson, Keith Marzilli, "Market Design When Firms Interact with Inertia Consumers: Evidence from Medicare Part D," *American Economic Review*, 2014, *104* (1), 1–37.

Financial Conduct Authority, "General Insurance Pricing Practices: Final Rules," *Policy Statement PS21/5*, 2021. May 2021. Available at https://www.fca.org.uk/publication/policy/ps21-5.pdf.

_ , "The Impact of the General Insurance Pricing Practices Remedies," *Evaluation Paper EP25/2*, 2025. Available at https://www.fca.org.uk/publication/evaluation-papers/ep25-2.pdf.

Gabaix, Xavier and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

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

Heidhues, Paul and Botond Kőszegi, "Behavioral Industrial Organization," *Handbook of Behavioral Economics*, 2018, *1*, 517–612.

Honka, Elisabeth, "Quantifying Search and Switching Costs in the US Auto Insurance Industry," *RAND Journal of Economics*, 2014, *45* (4), 847–884.

Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L Puller, "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market," *American Economic Journal: Economic Policy*, 2017, *9* (4), 192–226.

MacKinnon, James G, Morten Ørregaard Nielsen, and Matthew D Webb, "Cluster-Robust Inference: A Guide to Empirical Practice," *Journal of Econometrics*, 2023, *232* (2), 272–299.

Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb, "Fast and Wild: Bootstrap Inference in Stata Using Boottest," *Stata Journal*, 2019, *19* (1), 4–60.

Shleifer, Andrei, "Can Regulators Improve Markets? A Survey," *Journal of Economic Literature*, 2012, *50* (2), 421–432.

A. Data Appendix

FCA Aggregate Complaints Data. Data are downloaded from https://www.fca.org.uk/data/aggregate-complaints-data. Each half-yearly publication is an Excel file containing a "Product Specific" sheet (or "Products" for 2020 files) with complaints opened, complaints closed, and policies in force by insurance product line. I use two primary files: the 2022 H1 publication (covering 2016 H2–2022 H1) and the 2025 H1 publication (covering 2021 H1–2025 H1). For overlapping periods (2021 H1–2022 H1), I use the 2025 H1 data. The complaint rate is computed as (complaints opened / provision) $\times$ 1,000.

Bank of England Insurance Aggregate Data. The BoE Insurance Aggregate Data CSV is downloaded from https://www.bankofengland.co.uk/statistics/insurance. I extract Chart 4.1 data, which contains Net Written Premium (GBP) and Loss Ratio (%) by line of business, quarterly from 2017 Q1 to 2025 Q3. Lines are mapped to treated/control groups: Motor liability, Motor other, and Property are treated; Medical expense, Legal expenses, and Assistance are controls.

Product	Classification	Rationale
Motor & transport	Treated	Subject to GIPP pricing remedy
Property	Treated	Subject to GIPP pricing remedy
Assistance	Control	Not covered by pricing remedy
Medical/health	Control	Not covered by pricing remedy
Pet	Control	Not covered by pricing remedy
Travel	Excluded	COVID-confounded
Warranty	Control	Not covered by pricing remedy

Product classification.

B. Robustness Appendix

Travel insurance exclusion. Travel insurance is excluded from the preferred specification because COVID-19 created an extreme asymmetric shock to this product line. Travel insurance claims surged during lockdowns (trip cancellations) and then collapsed as coverage restrictions tightened. Including travel weakens the estimate from -1.87 to -1.45 because travel's volatile complaint trajectory adds noise to the control group. Excluding travel is conservative: it removes the control product most likely to violate parallel trends.

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Complaint rate	-1.871	0.525	1.812	-1.032	0.290	Large negative
Log complaints	-0.212	0.240	0.935	-0.227	0.257	Large negative
Log net written premium	0.247	0.136	1.188	0.208	0.114	Large positive
Loss ratio	0.158	5.530	13.852	0.011	0.399	Small positive
Panel B: Heterogeneous						
Complaint rate (Motor)	-2.084	—	1.004	-2.076	—	Large negative
Complaint rate (Property)	-1.658	—	0.200	-8.310	—	Large negative

Notes: **Country:** United Kingdom. **Research question:** Does banning price-walking (loyalty penalties) in motor and home insurance reduce consumer complaints? **Policy mechanism:** The FCA's General Insurance Pricing Practices reform (PS21/5) prohibits insurers from charging renewal customers more than the equivalent new-business price, eliminating the practice of gradually increasing premiums for loyal customers who do not shop around. **Outcome definition:** Complaints opened per 1,000 policies in force, measured half-yearly from FCA Aggregate Complaints Data. **Treatment:** Binary; motor and home insurance subject to pricing remedy from January 1, 2022; other insurance lines are not. **Data:** FCA Aggregate Complaints Data (half-yearly, 2016 H2–2025 H1, 6 products $\times$ 18 semesters = 108 observations, excluding travel) and Bank of England Insurance Aggregate Data (quarterly, 2017 Q1–2025 Q3, 6 lines $\times$ 35 quarters = 210 observations). **Method:** Two-way fixed effects with product and time fixed effects; standard errors clustered at product level; wild cluster bootstrap and randomization inference for few-cluster robustness. **Sample:** UK-regulated general insurers; all firms reporting to FCA/BoE. Travel insurance excluded due to COVID confounding. SDE = $\hat{\beta}$/SD(Y) where SD(Y) is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large ($|\text{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

COVID period sensitivity. Dropping 2020 H1 and 2020 H2 entirely (the acute COVID period) yields a similar estimate (-1.54) with slightly improved precision. The combined exclusion of travel and COVID periods produces the largest estimate (-1.91), suggesting that pandemic-related disruption attenuated the measured treatment effect.

C. Standardized Effect Sizes