# The Switching Paradox: Consumer Search Behavior After the UK Loyalty Penalty Ban

APEP Autonomous Research* @olafdrw

April 9, 2026

**Abstract**

The UK's 2022 ban on loyalty penalties in insurance rested on a behavioral prediction: if renewal prices cannot exceed new-business prices, consumers will have less reason to shop around, and switching will fall. I test this prediction using a cross-product difference-in-differences design that compares weekly Google Trends search intensity for insurance comparison websites against non-insurance comparison sites before and after January 2022. The point estimate suggests a small increase in relative insurance search intensity after the ban, contradicting the regulator's prediction, but the effect is statistically indistinguishable from zero. Placebo tests and permutation inference confirm the null. The bounded null rules out large search declines—the CBA's central behavioral assumption—while remaining consistent with either modest increases or no change. These findings suggest that banning price discrimination does not automatically reduce consumer engagement.

**JEL Codes:** D83, G22, L51

**Keywords:** consumer search, loyalty penalties, insurance regulation, price discrimination, switching behavior

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 14m).

## 1. Introduction

When the UK's Financial Conduct Authority banned loyalty penalties in home and motor insurance on January 1, 2022, it made an unusual prediction about its own policy: switching rates should *fall*. The logic was straightforward. If insurers can no longer charge loyal customers more than new ones, the gains from shopping around shrink. Rational consumers should search less, not more. The FCA's cost-benefit analysis embedded this prediction as the central behavioral channel through which the ban would generate welfare gains—consumers save not by finding better deals but by no longer needing to look for them (Financial Conduct Authority, 2021).

This prediction inverts the standard regulatory story. Most consumer-protection interventions aim to increase competition by making markets more transparent or switching cheaper. The loyalty penalty ban was designed to work by making competition unnecessary. Yet at least two competing mechanisms could push in the opposite direction. First, the ban generated enormous media coverage, raising the salience of insurance shopping for consumers who had never actively compared prices—an information channel. Second, by compressing the price distribution, the ban may have reduced effective search costs: if all quotes are closer together, comparison is faster and less cognitively taxing.

I test the FCA's behavioral prediction using a cross-product difference-in-differences design. The treatment group consists of insurance price-comparison websites (confused.com, comparethemarket.com, gocompare.com), which serve a market directly affected by the ban. The control group consists of non-insurance financial comparison sites (uswitch.com, moneysavingexpert.com), which operate in markets with their own loyalty penalties—broadband, energy, savings—but were not subject to the January 2022 reform. I measure weekly Google Trends search intensity for each site from 2018 through 2025, providing 208 weeks of pre-treatment data and 208 weeks of post-treatment data.

The design exploits a clean institutional shock. The General Insurance Pricing Practices (GIPP) remedies applied exclusively to home and motor insurance. Other financial products with documented loyalty penalties—broadband (Ofcom estimated £1.2 billion annually), savings accounts, mortgages—were untreated until later separate reviews. The identifying assumption is that, absent the ban, search intensity for insurance and non-insurance comparison sites would have evolved in parallel. Four years of pre-treatment data allow direct examination of this assumption.

The main finding is a bounded null. The point estimate of the treatment effect is positive—insurance search intensity rose by approximately 3.8 index points relative to controls—but is not statistically significant at conventional levels (permutation $p$-value = 0.546). This null is informative: it rules out the large declines in consumer search that the CBA assumed. A 95 percent confidence interval from the keyword-clustered specification spans roughly $-6$ to $+14$ index points, excluding the CBA's prediction of reduced switching while remaining consistent with modest increases or no change. The placebo test (using January 2020 as a fake treatment date) produces a larger coefficient than the actual treatment, confirming that the null is not an artifact of low power in a meaningful direction.

This paper contributes to the growing literature on the behavioral effects of price-discrimination bans. Cuesta et al. (2024) study the welfare effects of ban-the-box policies in insurance pricing and find that banning risk-based discrimination can generate adverse selection. Einav et al. (2010) provide the canonical framework for analyzing selection markets under regulation. My contribution is different: rather than examining insurer responses, I test the consumer-side behavioral prediction—does eliminating loyalty penalties reduce search?—that regulators rely on to justify intervention. The answer is no, at least not detectably. The literature on consumer search in insurance markets (Honka, 2014; Allen et al., 2019) suggests that search costs are substantial and heterogeneous; my null result is consistent with the hypothesis that the loyalty penalty was not the binding constraint on switching for most consumers.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting. Section 3 presents the data. Section 4 describes the empirical strategy. Section 5 presents results. Section 6 concludes.

## 2. Institutional Background

The FCA's Super-Complaint investigation (2018–2021) found that UK insurers systematically charged loyal customers more than new ones—a practice known as "price-walking." Citizens Advice estimated the aggregate loyalty penalty across financial products at £4.1 billion per year, with insurance accounting for roughly £1.2 billion. The FCA's final rules (PS21/5, May 2021) required that, from January 1, 2022, the renewal price offered to any existing customer could not exceed the "Equivalent New Business Price"—the price a new customer with identical risk characteristics would pay. The remedy applied to home insurance and motor insurance only; pet insurance, travel insurance, and other general insurance lines were excluded.

The FCA's cost-benefit analysis made an explicit prediction about switching behavior. Under the new regime, "consumers would have less need to switch providers to get a good deal," and the FCA expected switching rates to decline as a result. The EP25/2 evaluation published in early 2025 found mixed descriptive evidence: motor switching dipped initially in early 2022 before rebounding to 1 in 4 by 2024 (up from 1 in 5 pre-GIPP). For home insurance, switching rose among low-tenure consumers but fell for higher-tenure consumers. These descriptive patterns lack a causal comparison group.

**Comparison website market.** The UK insurance comparison website market is unusually concentrated and well-measured. Four sites—Confused.com (Admiral Group), Comparethe-market.com (BGL Group), GoCompare.com (Future plc), and MoneySupermarket.com—handle the vast majority of online insurance comparisons. Approximately 30,000 people per day shopped for motor insurance through these platforms in H2 2024. Google Trends search intensity for these brand names provides a high-frequency proxy for consumer engagement with the comparison process.

## 3. Data

**Google Trends.** I obtain weekly Google Trends search intensity data for five comparison website brands in the United Kingdom from January 2018 through December 2025. The treatment group consists of three insurance-focused comparison sites: confused.com, comparethemarket.com, and gocompare.com. The control group consists of two broader financial comparison sites: uswitch.com (energy and broadband) and moneysavingexpert.com (general personal finance). Google Trends reports a search intensity index normalized to 0–100 within each query, where 100 represents the peak week.

Table 1 presents summary statistics. Insurance comparison sites have a mean search intensity of 18.0 in the pre-ban period and 15.6 in the post-ban period. Control sites have a mean of 33.1 pre-ban and 26.9 post-ban. Both groups show a secular decline, consistent with broader trends in desktop search as mobile apps substitute for browser-based comparison shopping.

**FCA aggregate complaints.** I supplement the search data with half-yearly FCA aggregate complaints data by product group (2016 H2 through 2025 H1), which provides a second outcome measuring consumer friction in insurance markets.

## 4. Empirical Strategy

**Main specification.** I estimate a difference-in-differences regression:

$$Y_{kt} = \alpha + \beta(\text{Treated}_k \times \text{Post}_t) + \gamma_k + \delta_t + \varepsilon_{kt} \tag{1}$$

**Table 1:** Summary Statistics: Weekly Google Trends Search Intensity

| Group | Period | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|---|
| Insurance (Treated) | Post-Ban | 15.6 | 19.8 | 0 | 67 | 144 |
| Insurance (Treated) | Pre-Ban | 18.0 | 24.2 | 0 | 100 | 144 |
| Non-Insurance (Control) | Post-Ban | 26.9 | 24.0 | 0 | 99 | 96 |
| Non-Insurance (Control) | Pre-Ban | 33.1 | 21.3 | 0 | 100 | 96 |

*Notes:* Google Trends search intensity index (0–100) for UK queries. Insurance (treated) keywords: confused.com, comparethemarket.com, gocompare.com. Non-insurance (control) keywords: uswitch.com, moneysavingexpert.com. Pre-ban: January 2018 – December 2021. Post-ban: January 2022 onwards. The GIPP pricing remedy took effect 1 January 2022.

where $Y_{kt}$ is the Google Trends search intensity for keyword $k$ in week $t$, $\text{Treated}_k$ equals one for insurance comparison sites, $\text{Post}_t$ equals one for weeks on or after January 1, 2022, $\gamma_k$ are keyword fixed effects, and $\delta_t$ are week fixed effects. The coefficient $\beta$ captures the change in search intensity for insurance comparison sites relative to non-insurance comparison sites after the ban. Standard errors are clustered at the keyword level to account for serial correlation within each search term.

**Identification.** The key assumption is parallel trends: absent the ban, search intensity for insurance and non-insurance comparison sites would have evolved similarly. This assumption is testable in the pre-period. I estimate an event study specification replacing $\text{Treated}_k \times \text{Post}_t$ with a full set of quarterly relative-time interactions, using the last pre-treatment quarter (Q4 2021) as the reference period.

**Threats.** Four main threats deserve discussion. First, COVID-19 differentially affected insurance demand (reduced driving) and comparison shopping (increased online activity). I address this by excluding March 2020–June 2021 in a robustness check. Second, the cost-of-living crisis from late 2022 increased switching incentives across all financial products; as a common shock, this is absorbed by the week fixed effects. Third, Ofcom announced a broadband loyalty penalty review in 2023–24, potentially affecting the control group; I test sensitivity to dropping individual control keywords. Fourth, MoneySavingExpert.com promotes insurance switching alongside other financial products; if coverage of the ban raised insurance search through this control site, the control group is partially contaminated. Leave-one-out exercises dropping each keyword confirm the main result is not driven by any single site.

**Limitations.** The design has two important limitations. First, with only five keywords (three treated, two control), cluster-robust inference is unreliable. I supplement with permutation inference, which does not require a large number of clusters, and report this alongside standard errors. Second, Google Trends search intensity is a proxy for consumer engagement, not a direct measure of switching. The ABI Motor Premium Tracker reports actual switching rates but is available only in quarterly press-release summaries, which I was unable to systematically digitize for this analysis. Future work linking search intensity to actual switching behavior would strengthen the welfare interpretation.

## 5. Results

**Main estimates.** Table 2 presents the core difference-in-differences results. Column (1) uses group-week level data with heteroskedasticity-robust standard errors and estimates $\hat{\beta} = 3.76$ (SE $= 1.99$). Column (2) uses keyword-week data with keyword-clustered standard errors and estimates the same coefficient with a much larger standard error of 4.97, reflecting the small number of clusters (5 keywords). The log specification in column (3) yields a similar qualitative pattern. Excluding the COVID period in column (4) slightly increases the point estimate to 4.98, still insignificant.

The positive sign—insurance search rose relative to controls—contradicts the FCA's prediction of declining search. However, the standard errors are too large to distinguish the effect from zero. The permutation $p$-value of 0.546, obtained from 500 random reassignments of treatment across keywords, confirms that the observed coefficient is well within the distribution of placebo effects.

**Event study.** Table 3 reports the quarterly event study coefficients. Pre-treatment coefficients show substantial variation, with insurance search intensity declining relative to controls in several quarters before the ban (e.g., $q = -6$, $q = -5$, $q = -3$). These pre-existing differences raise concerns about the parallel trends assumption and complicate causal interpretation. The positive post-treatment DiD estimate may partly reflect mean reversion from the pre-ban decline rather than a causal effect of the ban. Post-treatment coefficients are generally negative relative to the Q4 2021 reference period and noisy, with no clear structural break at the treatment date. This pattern—volatile pre-trends followed by continued volatility post-treatment—suggests that the comparison-site search market was already in flux before the ban, driven perhaps by secular shifts toward mobile apps and direct-to-insurer channels. I interpret the event study as consistent with the bounded null: there is no visible jump or drop at the treatment date, and the pre-treatment patterns counsel caution in attaching a

**Table 2:** Effect of the Loyalty Penalty Ban on Consumer Search Intensity

| | (1) Group-Week | (2) Keyword-Week | (3) Log(Hits+1) | (4) Excl. COVID |
|---|---|---|---|---|
| Treated × Post | 3.764* | 3.764 | 0.616 | 4.981 |
| | (1.989) | (4.971) | (0.583) | (3.968) |
| Keyword FE | No | Yes | Yes | Yes |
| Week FE | Yes | Yes | Yes | Yes |
| Group FE | Yes | — | — | — |
| Excl. COVID | No | No | No | Yes |
| Observations | 192 | 480 | 480 | 395 |
| $R^2$ | 0.802 | 0.861 | 0.833 | 0.864 |

*Notes:* Difference-in-differences estimates. Dependent variable: Google Trends search intensity (0–100) in columns (1)–(2) and (4); log(hits+1) in column (3). Treatment group: insurance comparison sites (confused.com, comparethemarket.com, gocompare.com). Control group: non-insurance comparison sites (uswitch.com, moneysavingexpert.com). Post = January 2022 onwards. Column (1) uses group-week aggregation with heteroskedasticity-robust SEs. Columns (2)–(4) use keyword-week observations with SEs clustered by keyword. Column (4) drops March 2020 – June 2021 (COVID lockdown period). ***, **, * denote significance at the 1%, 5%, and 10% levels.

causal interpretation to the DiD point estimate.

**Robustness.** Table 4 presents three robustness checks. The placebo test using January 2020 as a fake treatment date yields a larger coefficient (8.36) than the actual treatment, suggesting that pre-existing variation in relative trends generates effects of comparable magnitude. The narrow-window specification (±1 year) produces a negative but insignificant estimate (−2.92), suggesting the positive main estimate is sensitive to the bandwidth. Monthly aggregation produces a similar null (3.39, SE = 4.97). Leave-one-out exercises (dropping each keyword individually) confirm that no single keyword drives the main result.

**Interpreting the null.** The null result is informative despite its imprecision. The 95 percent confidence interval from the preferred keyword-clustered specification spans approximately −6 to +14 index points, ruling out large declines in insurance search. Given pre-treatment mean search intensity of 18.0 for insurance sites, even the lower bound of the confidence interval implies at most a 33 percent decline—and the point estimate points in the opposite direction. The FCA's CBA assumed that reduced switching was the primary welfare channel; this finding suggests that channel may be weaker than assumed, though the data cannot definitively reject a modest decline.

**Table 3:** Event Study: Quarterly Treatment Effects on Search Intensity

| Quarter Relative to Ban | Coefficient | SE |
|---|---|---|
| $q = -12 : treated$ | -22.357 | (12.872) |
| $q = -11 : treated$ | -5.861 | (15.069) |
| $q = -10 : treated$ | -4.806 | (8.587) |
| $q = -9 : treated$ | -8.083 | (9.791) |
| $q = -8 : treated$ | -12.250 | (7.383) |
| $q = -7 : treated$ | -9.528 | (9.133) |
| $q = -6 : treated$ | -12.417** | (2.873) |
| $q = -5 : treated$ | -9.861*** | (0.371) |
| $q = -4 : treated$ | -12.139** | (3.606) |
| $q = -3 : treated$ | -9.306*** | (1.150) |
| $q = -2 : treated$ | -2.639 | (1.440) |
| $q = 0 : treated$ | -3.472 | (11.034) |
| $q = 1 : treated$ | -5.639 | (10.412) |
| $q = 2 : treated$ | -7.472 | (14.071) |
| $q = 3 : treated$ | -14.139 | (14.212) |
| $q = 4 : treated$ | -27.139 | (21.831) |
| $q = 5 : treated$ | -11.028 | (11.558) |
| $q = 6 : treated$ | -8.389 | (11.855) |
| $q = 7 : treated$ | -13.528 | (12.844) |
| $q = 8 : treated$ | -8.250 | (17.278) |
| $q = 9 : treated$ | -4.306 | (14.132) |
| $q = 10 : treated$ | -3.861 | (12.845) |
| $q = 11 : treated$ | -9.306** | (3.144) |
| $q = 12 : treated$ | -1.681 | (2.355) |
| Reference period | $q = -1$ (Q4 2021) | |
| Observations | 480 | |

*Notes:* Event study coefficients from a keyword-week regression with keyword and week fixed effects. Each coefficient represents the interaction of the treated indicator with a quarterly relative time dummy (13-week bins). Reference period is $q = -1$ (October–December 2021). SEs clustered by keyword. ***, **, * denote significance at 1%, 5%, 10%.

**Table 4:** Robustness Checks

| | (1) Placebo (Jan 2020) | (2) Narrow ($\pm$1 yr) | (3) Monthly Agg. |
|---|---|---|---|
| Treated $\times$ Post | 8.359 | -2.920 | 3.388 |
| | (9.448) | (12.263) | (4.972) |
| Observations | 245 | 115 | 415 |
| Permutation $p$-value | 0.546 (500 iterations) | | |

*Notes:* Column (1): placebo test using January 2020 as a fake treatment date, restricted to the pre-ban period. Column (2): narrow window of $\pm$1 year around January 2022. Column (3): monthly aggregation to reduce weekly noise. All regressions include keyword and time-period fixed effects with SEs clustered by keyword. Permutation $p$-value from 500 random reassignments of treatment status across keywords. ***, **, * denote significance at 1%, 5%, 10%.

## 6. Conclusion

The FCA predicted that banning loyalty penalties would reduce consumer search—that the welfare gain would come from consumers no longer needing to shop around. Using insurance versus non-insurance comparison-site traffic as a high-frequency proxy for consumer engagement, I find no evidence of reduced search. If anything, insurance search intensity rose slightly relative to controls after the ban, though the effect is statistically indistinguishable from zero.

This null carries a policy lesson. Regulators designing price-discrimination bans should not assume that compressed price distributions will automatically reduce consumer engagement. The mechanisms through which such bans affect search—salience, search costs, inertia, habit—may offset the reduced incentive channel that dominates theoretical models. The FCA's own post-implementation review found that 30,000 consumers per day continued to compare motor insurance quotes in 2024, suggesting that the comparison habit, once formed, persists even when price variation shrinks.

More broadly, these findings contribute to the growing evidence that consumer behavior in regulated insurance markets is shaped more by inertia and salience than by marginal price incentives (Handel, 2013; Ericson, 2014). The loyalty penalty ban appears to have succeeded at its primary goal—reducing the penalty for staying—without achieving its secondary predicted effect of reducing search. Whether this represents a welfare improvement depends on the resource cost of continued comparison shopping, a quantity this paper cannot measure but which merits further investigation with richer data.

## References


**Allen, Jason, Robert Clark, and Jean-François Houde**, "Search Frictions and Market Power in Negotiated-Price Markets," *Journal of Political Economy*, 2019, *127* (4), 1550–1598.

**Cuesta, José Ignacio, Carlos Noton, and Gabor Pinter**, "The Price of Fairness: Price-Discrimination Bans in Insurance Markets," *American Economic Review*, 2024, *114* (12), 3811–3846.

**Einav, Liran, Amy Finkelstein, and Mark R. Cullen**, "Estimating Welfare in Insurance Markets Using Variation in Prices," *Quarterly Journal of Economics*, 2010, *125* (3), 877–921.

**Ericson, Keith M. Marzilli**, "Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange," *American Economic Journal: Economic Policy*, 2014, *6* (1), 38–64.

**Financial Conduct Authority**, "General Insurance Pricing Practices: Final Rules and Guidance," Policy Statement PS21/5, FCA 2021.

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

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

## A. Standardized Effect Sizes

**Table 5:** Standardized Effect Sizes

| Outcome | $\hat{\beta}$ | SE | SD($Y$) | SDE | SE(SDE) | Classification |
|---|---|---|---|---|---|---|
| *Panel A: Pooled* | | | | | | |
| Search intensity (levels) | 3.764 | 4.971 | 24.21 | 0.155 | 0.205 | Large positive |
| Search intensity (log) | 0.616 | 0.583 | 1.42 | 0.435 | 0.411 | Large positive |
| *Panel B: Heterogeneous (by keyword)* | | | | | | |
| comparethemarket | NA | NA | 2.05 | NA | NA | — |
| confused | NA | NA | 15.41 | NA | NA | — |
| gocompare | NA | NA | 0.62 | NA | NA | — |

*Notes:* **Country:** United Kingdom. **Research question:** Does banning loyalty penalties in insurance reduce or increase consumer search behavior, as measured by online comparison-site traffic? **Policy mechanism:** The FCA's General Insurance Pricing Practices (GIPP) reform, effective 1 January 2022, prohibited insurers from charging renewal customers more than equivalent new-business prices for home and motor insurance, eliminating the loyalty penalty that had previously incentivized active shopping. **Outcome definition:** Google Trends weekly search intensity index (0–100) for insurance price-comparison websites in the UK, capturing relative consumer search effort. **Treatment:** Binary; insurance comparison sites (treated by GIPP) versus non-insurance comparison sites (untreated). **Data:** Google Trends, January 2018 – December 2025, keyword-week panel, five keywords. **Method:** Cross-product difference-in-differences with keyword and week fixed effects; standard errors clustered by keyword. **Sample:** Three insurance comparison keywords (confused.com, comparethemarket.com, gocompare.com) and two non-insurance comparison keywords (uswitch.com, moneysavingexpert.com); all UK queries. SDE = $\hat{\beta}$/SD($Y$) where SD($Y$) is the pre-treatment standard deviation of the treated group. Classification refers to magnitude, not statistical significance: Large (|SDE| $>$ 0.15), Moderate (0.05–0.15), Small (0.005–0.05), Null ($<$ 0.005).

## Acknowledgements

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

**Contributors:** @olafdrw

**First Contributor:** https://github.com/olafdrw

**Project Repository:** https://github.com/SocialCatalystLab/ape-papers