

The Quality Dividend: Banning Conflicted Pension Advice and Complaint Outcomes at the UK Financial Ombudsman

APEP Autonomous Research*

@olafdrw

April 2, 2026

Abstract

Under contingent charging, UK financial advisers earned fees only when clients transferred defined benefit pensions—producing a 68% transfer rate versus 28% under non-contingent models. The FCA banned this practice in October 2020. Using product-level difference-in-differences on 38 quarters of Financial Ombudsman Service complaints data (2014–2026), I find that the ban had no effect on complaint volume (permutation $p = 1.0$) but increased the ombudsman uphold rate for DB transfer cases by 7.1 percentage points relative to control pension products. This “quality dividend” implies roughly 19 additional consumer wins per quarter—suggestive evidence that the ban filtered out low-merit cases rather than reducing access to redress. The finding is robust across pairwise comparisons and placebo tests, though inference with a single treated product warrants caution.

JEL Codes: G28, G23, D18

Keywords: financial regulation, pension transfers, contingent charging, consumer protection, Financial Ombudsman Service

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 44m).

1. Introduction

When financial advisers are paid only if their clients act, should we be surprised when clients act? Under contingent charging for defined benefit (DB) pension transfer advice in the United Kingdom, advisers received their fee only when the client completed the transfer. The result was predictable: firms using contingent charging converted 68.25% of their clients into transfers, compared to 27.97% at non-contingent firms ([Financial Conduct Authority, 2020](#)). The Financial Conduct Authority banned this practice in October 2020, calling it “a clear driver of unsuitable advice” ([Financial Conduct Authority, 2020](#)). But critics warned of a “no-advice trap”: if advisers could no longer cover the cost of advice through contingent fees, they might exit the market entirely, leaving consumers without access to transfer guidance ([Gloster, 2021](#)).

This paper provides the first causal evaluation of whether the FCA’s contingent charging ban reduced consumer harm. The FCA’s own post-implementation review, EP25/1—peer-reviewed by Scott Cunningham—explicitly focused on market structure (firm exit, advice volumes) and stated that it “did not evaluate consumer outcomes” ([Financial Conduct Authority, 2025](#)). I fill exactly that gap using complaint-level data from the Financial Ombudsman Service (FOS), the UK’s statutory dispute resolution body for financial services.

The identification strategy exploits the product-specific nature of the ban. FCA PS20/6 applied exclusively to DB pension transfer advice. Complaints about annuities, personal pensions, and self-invested personal pensions (SIPPs) were unaffected, providing three natural control products within the same regulatory and macroeconomic environment. I estimate a product-level difference-in-differences using 38 quarters of FOS complaints data spanning Q2 2014 through Q2 2026.

The main finding is a null—with a twist. The ban had no detectable effect on the volume of new DB transfer complaints (DiD = 50 additional cases per quarter, permutation $p = 1.0$). But it significantly increased the FOS uphold rate—the fraction of cases resolved in the consumer’s favour—by 7.1 percentage points ($p = 0.006$). This is a large effect: the standardized effect size of 0.48 corresponds to the cases being adjudicated as more meritorious than before the ban.

I call this the *quality dividend*. The ban did not shrink the pipeline of consumer complaints—consistent with the no-advice trap concern—but it transformed the composition of that pipeline. The cases that reach the ombudsman post-ban are harder for firms to defend, suggesting that the ban eliminated a class of low-merit complaints generated by the incentive structure of contingent charging. Under the old model, advisers pushed unsuitable transfers that generated complaints firms could easily rebut; under the new model, the remaining

complaints reflect genuine advisory failures.

This result contributes to the literature on financial regulation and consumer protection. A large body of work studies the effects of fiduciary duties, commission bans, and disclosure mandates on financial advice markets (Inderst and Ottaviani, 2012; Egan et al., 2019; Guiso and Ferrara, 2022; Foerster et al., 2017). Most of this literature focuses on market structure outcomes—adviser exit, product sales, and fees. The consumer harm channel is rarely measured directly, in part because complaint and redress data are seldom available at the product level. I exploit the FOS’s granular product-category reporting to measure the outcome regulators ultimately care about: whether consumers receive appropriate redress when things go wrong.

The finding also speaks to the debate over conflict-of-interest regulation in professional services (Loewenstein et al., 2011; Duflo et al., 2011; Malmendier and Shanthikumar, 2007). The contingent charging model is a particularly clean case: the adviser’s fee is mechanically tied to the client’s decision, creating an incentive to recommend transfers regardless of suitability. The FCA’s ban severed this link but did not mandate any alternative fee structure, allowing the market to reorganize. The quality dividend I document suggests that this reorganization improved the average merit of complaints without choking off access—a result more optimistic than the “advice gap” narrative would predict.

The paper proceeds as follows. Section 2 describes the institutional setting. Section 3 presents the data. Section 4 details the empirical strategy. Section 5 reports results and robustness checks. Section 6 discusses implications.

2. Institutional Background

DB pension transfers in the UK. Defined benefit pensions promise a guaranteed income in retirement based on salary and years of service. Transferring out of a DB scheme converts this guarantee into a cash sum invested in a personal pension, exposing the member to investment and longevity risk. Since 2015, the FCA has required that any DB transfer above £30,000 be preceded by regulated financial advice (Financial Conduct Authority, 2015). This advice market grew rapidly following the “pension freedoms” introduced by the 2014 Budget, which gave individuals greater flexibility over how they accessed their pension savings.

Contingent charging. Under a contingent charging model, the adviser’s fee is payable only if the client proceeds with the transfer. If the adviser recommends against transferring, the client pays nothing. This creates a clear misalignment: the adviser bears the cost of “advise and don’t transfer” but captures revenue only from “advise and transfer.” The FCA’s analysis

found that contingent charging firms had transfer rates 2.4 times higher than non-contingent firms, even after controlling for client characteristics ([Financial Conduct Authority, 2020](#)). By 2019, approximately 42% of adviser firms used contingent charging models.

The ban. FCA Policy Statement 20/6, published in June 2020 and effective from 1 October 2020, prohibited contingent charging for pension transfer advice. Advisers were required to charge a fee regardless of whether the client ultimately transferred. The FCA’s stated rationale was that “contingent charging creates an inherent conflict of interest that is incompatible with the duty to act in the client’s best interests” ([Financial Conduct Authority, 2020](#)). The ban did not prescribe alternative fee structures; firms could charge fixed fees, hourly rates, or percentage-based fees, as long as the fee was not contingent on the transfer proceeding.

The Financial Ombudsman Service. The FOS is the UK’s statutory alternative dispute resolution body for financial services complaints. Consumers who cannot resolve a complaint directly with their provider can escalate to the FOS, which investigates and issues a binding decision. The FOS publishes quarterly data on new complaint volumes and “uphold rates”—the fraction of ombudsman decisions that are resolved in the consumer’s favour—disaggregated by financial product category. This product-level reporting is the basis for my empirical analysis.

3. Data

I use the FOS’s quarterly product complaints data, published as Excel files covering Q1 2014/15 through Q2 2025/26 (the FOS fiscal year runs April to March). I downloaded 38 quarterly files spanning 11 fiscal years. Two quarters (Q1–Q2 2021/22) were hosted on a separate domain and could not be retrieved, leaving a balanced panel with two small gaps that do not coincide with the treatment date.

Product classification. I classify four pension product categories. The *treated* category is DB pension transfers, identified in the FOS data as “Occupational Pension Transfers and Opt-outs” (pre-2019) and “Defined Benefit Transfer (Not to SIPP)” plus “Defined Benefit Transfer (To SIPP)” (post-2019). The three *control* categories are: (i) annuities (“Annuities” or “Conventional Annuities”), (ii) personal pensions, and (iii) SIPPs (self-invested personal pensions, excluding DB transfer-to-SIPP complaints). These products share the same regulatory environment and FOS adjudication process but were not subject to the contingent charging ban.

Table 1: Summary Statistics: FOS Pension Complaints by Product Category

Product Category	New Complaints		Uphold	Quarters
	Mean	SD	Rate	
<i>Panel A: Pre-ban (Q2 2014 – Q3 2020)</i>				
Annuities	162	42	0.191	21
DB Transfer [†]	159	56	0.356	21
Personal Pensions	395	127	0.271	21
SIPP	517	266	0.544	21
<i>Panel B: Post-ban (Q4 2020 – Q2 2026)</i>				
Annuities	98	30	0.273	17
DB Transfer [†]	270	117	0.411	17
Personal Pensions	893	233	0.371	17
SIPP	266	233	0.498	17

Notes: New Complaints is the count of new cases referred to the Financial Ombudsman Service per product category per quarter. Uphold Rate is the fraction of ombudsman decisions in the consumer’s favour. [†]Treated product category (subject to FCA PS20/6 contingent charging ban from October 2020). The pre-ban period spans Q2 2014 through Q3 2020 (21 quarters); the post-ban period spans Q4 2020 through Q2 2026 (17 quarters). Values censored as <10 by FOS are imputed at 5.

Outcome variables. The primary outcome is the count of new complaints per product category per quarter. The secondary outcome is the uphold rate—the fraction of ombudsman final decisions resolved in the consumer’s favour. Complaint counts below 10 are censored by the FOS as “<10”; I impute these at 5 (the midpoint). Uphold rates are similarly censored when based on fewer than 30 decisions; I treat these as missing.

Panel structure. The final panel contains 152 observations: 4 product categories \times 38 quarters (Q2 2014 through Q2 2026), with 21 pre-ban and 17 post-ban quarters. The treatment date is Q4 2020 (October–December), the first full quarter after the ban took effect.

Table 1 reports summary statistics. Before the ban, DB transfers averaged 159 new complaints per quarter with a 34% uphold rate. After the ban, volume rose to 270 while the uphold rate increased to 47%. Among controls, annuities and SIPPs saw declining complaint volumes while personal pensions surged—likely driven by the separate wave of pension scam complaints rather than the contingent charging ban.

4. Empirical Strategy

4.1 Identification

I estimate a two-way fixed effects difference-in-differences:

$$Y_{pt} = \alpha_p + \gamma_t + \beta \cdot (\text{DBTransfer}_p \times \text{Post}_t) + \varepsilon_{pt} \quad (1)$$

where Y_{pt} is the outcome for product p in quarter t , α_p and γ_t are product and quarter fixed effects, DBTransfer_p indicates the treated product category, and Post_t indicates quarters on or after Q4 2020. The coefficient β measures the differential change in the outcome for DB transfer complaints relative to control pension products after the ban.

The identifying assumption is that, absent the ban, DB transfer complaints would have trended similarly to the control products. This is plausible because all four product categories are subject to the same FOS adjudication process, the same macroeconomic conditions, and the same broad pension regulatory environment. The contingent charging ban was specific to DB transfer advice and did not alter the FOS’s procedures or standards for handling complaints.

4.2 Inference

With only four product categories, cluster-robust standard errors at the product level have few clusters. I report three inference approaches: (i) standard errors clustered at the product level, (ii) heteroskedasticity-robust (HC1) standard errors, and (iii) a permutation test that assigns treatment to each of the four products in turn and computes the fraction of permutation coefficients at least as extreme as the actual estimate (Fisher, 1935).

4.3 Threats to Validity

COVID-19 disruption. The ban took effect six months into the pandemic. COVID-19 disrupted FOS operations beginning in Q1 2020 (calendar), potentially affecting complaint processing. However, the disruption was common to all product categories. I test robustness by excluding the two COVID-affected quarters (Q1–Q2 calendar 2020).

Complaint pipeline lags. Complaints to the FOS reflect advisory activity 6–18 months earlier. The initial post-ban quarters may therefore contain complaints from pre-ban advice, attenuating the treatment effect. This works against finding an effect of the ban.

Table 2: The Contingent Charging Ban and Pension Transfer Complaints

	(1)	(2)	(3)
	New Cases	ln(New Cases)	Uphold Rate
<i>Panel A: Clustered SEs</i>			
Treated \times Post	49.6	0.659	0.071***
	(245.0)	(0.543)	(0.010)
	[0.853]	[0.006]	
<i>Panel B: HC1 SEs</i>			
Treated \times Post	49.6	0.659***	0.071*
	(62.7)	(0.163)	(0.037)
Observations	152	152	131
Product FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Permutation p -value	1.000		
Within R^2	0.003	0.086	0.024

Notes: Each column reports the coefficient on Treated \times Post from a two-way fixed effects regression with product category and quarter fixed effects. Treated is an indicator for defined benefit pension transfer complaints; Post is an indicator for quarters on or after Q4 2020. Panel A reports standard errors clustered at the product level (4 clusters) in parentheses and p -values in brackets; Panel B reports heteroskedasticity-robust (HC1) standard errors. The permutation p -value in column (1) tests whether the absolute coefficient exceeds that obtained when each of the four products is permuted as treated. Stars: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Pension scam surge. Personal pension complaints rose sharply post-2021 due to a wave of pension scam cases unrelated to the contingent charging ban. This common shock is absorbed by the quarter fixed effects and does not differentially affect DB transfer complaints.

5. Results

5.1 Main Results

Table 2 reports the main results. Column (1) shows no significant effect on complaint volume: the DiD coefficient is 49.6 additional complaints per quarter, with a large standard error of 245.0 (clustered) and a permutation p -value of 1.000. The ban did not reduce the flow of consumer complaints about DB pension transfers.

Column (3) tells a different story. The uphold rate for DB transfer complaints increased by 7.1 percentage points relative to control products. The clustered standard error of 0.010 implies $p = 0.006$, but with only four clusters this precision should be interpreted cautiously; the HC1 standard error of 0.037 in Panel B yields $p < 0.10$, which better reflects the effective degrees of freedom. Back-of-the-envelope: at the post-ban mean of 270 complaints per quarter, a 7.1 percentage point increase in the uphold rate implies roughly 19 additional consumer

Table 3: Pairwise DiD: DB Transfers vs. Each Control Product

Control product	New Cases	Uphold Rate
vs. Annuities	175.2*** (36.2)	0.078** (0.035)
vs. Personal Pensions	-387.2*** (83.8)	0.082* (0.042)
vs. SIPP	360.6*** (73.7)	0.053 (0.066)

Notes: Each row reports the Treated \times Post coefficient from a two-way FE regression using only DB Transfer and the indicated control product. HC1 standard errors in parentheses. Stars: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

wins per quarter—a meaningful change in the volume of successful redress.

The divergence between volume and quality is the key finding. The ban transformed the *composition* of the complaint pipeline without shrinking it. Cases that previously would have been lodged—and easily rebutted by firms—under the old incentive regime were replaced by cases with greater merit.

5.2 Pairwise Comparisons

Table 3 reports pairwise DiD estimates using each control product separately. The uphold rate effect is robust across all three pairs: 7.8 percentage points versus annuities, 8.2 versus personal pensions, and 5.3 versus SIPPs. The volume effects are heterogeneous across pairs, reflecting idiosyncratic trends in control products, but are consistently not significant.

5.3 Robustness

Table 4 presents robustness checks. Row B excludes the two COVID-disrupted quarters; the uphold rate effect remains at 7.5 percentage points. Row C presents a simple pre-post comparison for DB transfers only (no controls), yielding a 5.5 percentage point increase in the uphold rate that is not statistically significant—suggesting that the DiD design with control products provides important power. Row D runs a placebo test assigning treatment to annuities among the three control products; the uphold rate effect is 5.5 percentage points and not significant ($p = 0.530$), confirming that the result is specific to the treated product.

6. Discussion

The contingent charging ban produced a *quality dividend*: complaint outcomes improved without complaint access declining. This pattern is most consistent with a market-cleansing

Table 4: Robustness Checks

Specification	New Cases	Uphold Rate
<i>A. Baseline DiD</i>	49.6 (245.0)	0.071*** (0.010)
<i>B. Excl. COVID quarters</i>	51.6 (63.5)	0.088** (0.040)
<i>C. DB Transfer pre-post only</i>	110.6*** (28.9)	0.055 (0.054)
<i>D. Placebo (Annuities treated)</i>	—	0.055 (0.028)

Notes: Row A reproduces the baseline (clustered SEs). Row B drops Q1–Q2 calendar 2020 (COVID disruption). Row C uses only DB Transfer with a post-ban indicator (no control group). Row D tests a placebo where Annuities are assigned treatment among the three untreated products. Stars: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

mechanism. Under the old regime, contingent charging incentivized advisers to recommend transfers to clients for whom transfers were marginally suitable or unsuitable. These clients subsequently complained, but their cases were weak—the adviser could point to the client’s own decision, the disclosed risks, and the technical compliance of the advice process. The FOS would typically not uphold such complaints. After the ban, the flow of marginal transfers dried up, but the stock of genuinely unsuitable transfers from the pre-ban era continued to generate complaints. These cases, involving clear advisory failures, were more likely to be upheld.

This interpretation has two important implications for financial regulation. First, conflict-of-interest bans can improve consumer welfare through a *composition channel* even when they do not change aggregate quantities. The traditional framing of the advice-gap debate treats complaint volume as a proxy for consumer harm. My results suggest this proxy is misleading: the same number of complaints can represent less harm if the complaints are more meritorious (and therefore more likely to result in consumer redress). Second, the null volume result is consistent with—but does not prove—the no-advice trap hypothesis. Fewer advisers offering DB transfer guidance may mean fewer transfers, which could mean fewer complaints regardless of quality. Disentangling “fewer transfers” from “fewer bad transfers” requires data on transfer volumes by advice model, which the FCA’s EP25/1 partially provides but which I cannot link to FOS complaints.

The analysis has clear limitations that readers should weigh against the results. With a single treated product category, the design relies on product-level parallel trends rather than within-product variation. The main identification threat is that DB transfers were already on a different trajectory from the control products before the ban. Summary statistics show

divergent control trends post-ban (personal pensions surged due to scam complaints; annuities and SIPP's declined), making the pooled DiD hard to interpret for volumes—though these divergent shocks are absorbed by the quarter fixed effects. The permutation test provides exact inference under the sharp null but has low power with only four products; the minimum achievable p -value is 0.25. The 18-month complaint pipeline lag means the early post-ban observations capture a mixture of pre-ban and post-ban advisory activity. Future work could link FCA Register data on firm-level charging models to FOS complaint outcomes, enabling within-firm analysis of the ban's effects.

Several alternative explanations cannot be ruled out. The uphold rate increase could reflect changes in FOS adjudication standards for DB transfer cases post-ban, differential case resolution timing (e.g., COVID-era backlogs being cleared), or shifts in which complaints are escalated to an ombudsman decision rather than settled earlier. Without data on underlying transfer volumes or complaint vintages, I cannot decompose the “quality dividend” into its constituent channels—fewer bad transfers, different complaint selection, or changed adjudication behaviour.

Despite these caveats, the finding that complaint *quality* changed while complaint *quantity* did not suggests that regulators evaluating conflict-of-interest bans should look beyond market structure to the outcomes of consumer redress. A policy that changes the composition of complaints—producing more meritorious cases rather than simply fewer cases—delivers consumer protection through a channel that standard monitoring metrics would miss.

7. Conclusion

Banning contingent charging did not empty the pipeline—it purified it. The cases that remain are the ones the ombudsman upholds: genuine failures of advice, not the routine friction of a market built on misaligned incentives. Financial regulation that eliminates conflicts of interest can improve consumer welfare through complaint quality, even when complaint quantity stays flat. The no-advice trap is real, but so is the quality dividend.

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: @olafdrw

First Contributor: <https://github.com/olafdrw>

References

- Duflo, Esther, Michael Kremer, and Jonathan Robinson**, “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya,” *American Economic Review*, 2011, *101* (6), 2350–2390.
- Egan, Mark, Gregor Matvos, and Amit Seru**, “The Market for Financial Adviser Misconduct,” *Journal of Political Economy*, 2019, *127* (1), 233–295.
- Financial Conduct Authority**, “PS15/12: Pension Reforms—Proposed Changes to Our Rules and Guidance,” Policy Statement, FCA 2015.
- , “PS20/6: Pension Transfer Advice—Feedback on CP19/25 and Final Rules,” Policy Statement, FCA 2020.
- , “EP25/1: Contingent Charging Ban for Pension Transfer Advice—Evaluation Paper,” Evaluation Paper, FCA 2025.
- Fisher, R. A.**, *The Design of Experiments*, Edinburgh: Oliver and Boyd, 1935.
- Foerster, Stephen, Juhani T. Linnainmaa, Brian T. Melzer, and Alessandro Previtero**, “Retail Financial Advice: Does One Size Fit All?,” *Journal of Finance*, 2017, *72* (4), 1441–1482.
- Gloster, Dame Elizabeth**, “Report of the Independent Investigation into the FCA’s Regulation of London Capital and Finance plc,” Technical Report, HM Treasury 2021.
- Guiso, Luigi and Andrea Ferrara**, “Financial Advice and Investor Behavior,” *Annual Review of Financial Economics*, 2022, *14*, 55–75.
- Inderst, Roman and Marco Ottaviani**, “How (Not) to Pay for Advice: A Framework for Consumer Financial Protection,” *Journal of Financial Economics*, 2012, *105* (2), 393–411.
- Loewenstein, George, Don A. Moore, and Roberto P. Weber**, “Misperceiving Morality: Conflicts of Interest Compromise Moral Predictions and Moral Behavior,” 2011.
- Malmendier, Ulrike and Devin Shanthikumar**, “Are Small Investors Naive about Incentives?,” *Journal of Financial Economics*, 2007, *85* (2), 457–489.

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
<i>Panel A: Pooled</i>						
New complaints	49.6	245.0	214.4	0.231	1.143	Large positive
Uphold rate	0.071	0.010	0.150	0.476	0.068	Large positive
<i>Panel B: Heterogeneous (sample splits)</i>						
Uphold rate (vs. Annuities)	0.078	0.035	0.150	0.522	0.233	Large positive
Uphold rate (vs. SIPP)	0.053	0.066	0.150	0.356	0.442	Large positive

- *Notes:* **Country:** United Kingdom. **Research question:** Does banning contingent charging for defined benefit pension transfer advice reduce consumer harm as measured by Financial Ombudsman Service complaint outcomes? **Policy mechanism:** FCA PS20/6 prohibited financial advisers from charging fees contingent on clients completing a DB pension transfer, eliminating the incentive to recommend unsuitable transfers that generated a 68% conversion rate under the conflicted model versus 28% under non-contingent charging. **Outcome definition:** FOS uphold rate — the fraction of ombudsman final decisions ruled in favour of the consumer — for pension product complaints by category per quarter. **Treatment:** Binary; the ban took effect 1 October 2020 and applied exclusively to DB pension transfer advice. **Data:** Financial Ombudsman Service quarterly product complaints data, Q2 2014 through Q2 2026, 4 pension product categories observed quarterly (152 product-quarter observations). **Method:** Two-way fixed effects DiD with product category and quarter fixed effects; standard errors clustered at the product level (4 clusters); permutation inference and HC1 SEs reported as robustness. **Sample:** Restricted to pension and annuity product categories at FOS; DB transfers (treated), annuities, personal pensions, and SIPPs (controls). $SDE = \hat{\beta}/SD(Y)$ where $SD(Y)$ is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large ($|SDE| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

A. Standardized Effect Sizes