Fallible Firms:
The Market for Correcting Firm Mistakes*

Charlotte Haendler†
Southern Methodist University

Rawley Z. Heimer‡
Arizona State University

Abstract

Firms make mistakes. We study who bears the cost of firm fallibility using confidential records on nearly two million disputes over financial services filed with the Consumer Financial Protection Bureau, in which a successful claim paid an average of $2,295 in 2024. The complexity of the underlying contract and the claimant's own advocacy explain most of the variation in who obtains redress; firm identity adds little explanatory power. Higher-income consumers recover more restitution likely because they file better-supported claims, not because firms allocate resources differently. We identify the effect of enforcement pressure using staggered confidential report cards, in which the CFPB privately warns firms of deficiencies in their complaint handling. Regulatory scrutiny raises redress, but mainly for claimants who already engage effectively. A structural model of the firm's grant-or-deny decision, combined with population data on household product usage, values the errors that are never corrected. Foregone restitution among filers is modest, while the much larger loss falls on consumers who never complain, who are disproportionately lower-income. Redress in private markets thus mirrors the take-up problem in public programs: broadening awareness of the claims channel raises welfare far more than strengthening enforcement or exhorting self-advocacy.

*We are grateful to Lewis Kirvan who facilitated access to the confidential data and provided invaluable feedback. Heimer was previously an unpaid contract employee of the CFPB and the opinions of this article do not reflect the opinions of the CFPB or any governmental entity. Any errors or omissions are the authors' responsibility.

†Southern Methodist University, Edwin L. Cox School of Business. Email: chaendler@smu.edu

‡Arizona State University, W.P. Carey School of Business. Email: rawley.heimer@asu.edu

Introduction

Firms make mistakes. Charges are misapplied, contract terms are executed incorrectly, and obligations go unpaid. In most markets for physical goods, the resolution of firm errors is straightforward. For example, a defective toaster oven goes back to the shop's return counter, and the incidence of the firm's mistake falls, mostly, on the firm. In markets for services governed by bilateral contracts, however, verifying that an error occurred is itself costly. The harmed party must recognize the error, articulate why it merits redress, and marshal evidence; the firm, weighing its own costs, decides whether to pay. This paper studies that process at scale and asks a question that is fundamental to how markets allocate the costs of imperfection but that has received almost no direct empirical attention: when firms err, who ends up bearing the cost?[1]

The logic is straightforward. Suppose a bank client is charged an overdraft fee that her account agreement does not authorize. She may not notice the charge; if she notices, she may not know that the fee was unauthorized; if she knows, she may know channels to seek redress; and if she seeks redress, she must contest the error in ways that make denial costly for the firm. Each link in that chain is skill-intensive, and the firm's decision arrives only at the end of it. Redress therefore reaches the harmed party through a filter that selects on the ability to claim in addition to the fact of harm. This paper shows that who bears the cost of a firm's error depends less on the firm than on the claimant's ability to navigate the claims process. The loss is large and regressive, and is plausibly largest among consumers who never enter the process at all.

Our laboratory is the market for consumer financial services. We study novel confidential administrative records on nearly two million complaints filed with the Consumer Financial Protection Bureau (CFPB) between 2014 and 2024. Three properties of the setting make it suited to studying firm fallibility in general. First, financial services are the canonical complex contract: bilateral, long-lived, and difficult for one side to verify, so that actual (and perceived) errors are common and their correction is genuinely contested. Second, the CFPB channel is nonjudicial, free, and requires no legal representation, so the frictions we measure are informational and cognitive rather than artifacts of court costs.[2] Third, we observe the full anatomy of each dispute: the product, the narrative the consumer wrote, the documentation attached, the resources employed by the firm to address the complaint, the firm's decision, and the dollars paid. We combine these records with previously unexamined data on the regulator's confidential communications with firms. No comparable window into the ex post correction of firm errors exists.

[1] Economic models of financial services typically abstract from this margin entirely, treating transactions as frictionless and consumer costs as determined by interest rates and fees. See, e.g., Green (1973), Jaffee and Russell (1976), Stiglitz and Weiss (1981), and Jappelli and Pagano (1993).

[2] Collective mechanisms exist but reach a different margin: regulatory enforcement actions and class-action settlements target common issues at specific firms and do not resolve the individual disputes consumers must navigate on their own. On class settlements, see Schwemm and Taren (2010), Agarwal et al. (2014), Eisenberg and Miller (2015), and Hylton (2015).

This paper makes three contributions. First, we provide what is, to our knowledge, the most comprehensive measurement of firm error and its correction in the financial sector and show that redress varies overwhelmingly with contract complexity and claimant advocacy. Second, exploiting confidential regulatory report cards delivered to firms at staggered dates, we offer among the first causal evidence on how private enforcement pressure reallocates redress, and show that it complements rather than substitutes for claimant sophistication. Third, we develop a structural framework that converts claims-level administrative data into a measure of who bears the cost of firm mistakes, and we extend that measure to consumers who never file. The framework carries the take-up economics of public programs over to private transfers and delivers a policy ranking in which broadening awareness dominates both strengthening enforcement and teaching filers to argue their claims more effectively. The framework requires only claims-level data and is portable to other domains such as insurance and medical billing.

We begin by establishing that firm error and its correction are economically important. Consumers have filed nearly two million complaints since 2014, and the 7.5% resolved with monetary relief have returned over $220 million, with the average payment rising from $697 in 2014 to $2,295 in 2024.[3] Complaint volume has nearly tripled over the decade while the share paid has remained stable between 7 and 8%, which reflects growth in the use of the channel rather than a dilution of merit as awareness of the CFPB portal expands. These are the corrections that happen. The corrections that do not happen, whose victims are denied or never file claims, cannot be read off the complaint records. Valuing these errors requires a model of how firms decide to pay and of how consumers come to claim, which we develop in the second half of the paper.

We then study what determines which errors get corrected. We find that firm identity barely matters. If corporate misconduct was prevalent, with a subset of bad actors systematically short-changing customers, restitution rates should track firms.[4] They do not. In a variance decomposition, product type explains over 99% of the explained variation in restitution outcomes, dwarfing firm fixed effects, and the pattern holds within multi-product firms. The complexity of the financial product is the strongest predictor of redress. Roughly 17% of complaints about bank accounts and credit cards, simple products whose errors are easily verified, end in payment. Term loans, such as student loans and mortgages, have restitution rates between 2% and 5%. As such, fallibility is not a property of firms that perpetually engage in misconduct, but a baseline rate common across firms.[5]

[3] We exclude over two million credit reporting disputes because of their distinct nature and because credit reporting agencies do not provide in-kind financial service obligations.

[4] The literature on financial advisers finds that some firms persistently engage in misconduct (see Dimmock et al. (2018) and Egan et al. (2019)).

[5] The pattern is consistent with evidence that operational shortfalls are pervasive rather than confined to low-quality firms (Bloom and Van Reenen, 2007; DellaVigna and Gentzkow, 2019; Hortaçsu et al., 2019), and its natural theoretical antecedent is Carlin (2009).

On the other hand, we find substantial heterogeneity in consumer redress. Consumers from top-quintile-income census tracts obtain financial restitution in about 10% of cases against under 6% in the bottom quintile. This gradient survives product, time, and firm fixed effects. Specifically, higher-income claimants are more likely to engage in effective self-advocacy by demanding a refund or alleging fraud explicitly. They are also more likely to attach supporting documentation. These actions raise the restitution rate by approximately 46%. Firms, by contrast, show no evidence of favoritism in their resource allocation responding to complaints. Using internal identifiers, we track the experience and engagement of the employee assigned to each complaint and find no variation by claimant income. Also, employee fixed effects have little incremental explanatory power beyond firm fixed effects. These findings suggest that the income gradient in redress is similar to shopping-effort frictions (Woodward and Hall, 2012) rather than differential treatment by firms (Becker, 1971; Arrow, 1973).

Nonetheless, firms may deny valid low-income claims for reasons that we do not observe. We use previously unexamined regulatory data to explore this possibility. Since 2020 the CFPB has audited samples of individual complaint responses giving us an external benchmark for how each complaint should have been resolved. Low-income complaints are somewhat more likely to be judged mishandled, but the disparity is small relative to the restitution gap and disappears with firm fixed effects. Crucial for identification, the CFPB delivered reports about complaint-handling errors to more than fifty firms at staggered dates, which lets us estimate the causal effect of private enforcement pressure. After receiving a report, firms raise the share of complaints paid by 1.7 percentage points, and the increase accrues almost entirely to high-income and self-advocating claimants. Therefore, enforcement amplifies the returns to consumer sophistication rather than leveling the playing field.

The final part of the paper develops a structural model of the firm's grant-or-deny decision, both to explain the reduced-form findings and to estimate the incidence of firm error in the population that never files. The firm pays a claim when the expected enforcement cost of denying a meritorious complaint exceeds the restitution owed. The model features product-level intercepts that read as what the firm can deny with impunity, and they recover the complexity ordering documented nonparametrically above: around 1% baseline payment rate for student loans up to 17% for credit cards. The estimated policy function then decomposes the income gradient. The return to self-advocacy is large and precisely estimated, while the direct effect of income, conditional on advocacy, is small and non-significant. The model reproduces the dollars actually disbursed to within half a percent.

To value the errors that are never corrected, we extend the firm's policy function in two directions. We embed it in population data on household product usage, and we model how consumers learn that the complaint channel exists. The learning process follows an epidemiological diffusion in which awareness spreads through contact between households, disciplined by the observed

evolution of filing rates across income groups. Google searches for "CFPB complaint" have grown steadily since 2014 (Internet Appendix Figure I.2), consistent with a population learning about the CFPB channel gradually rather than responding discretely to its inception as a Federal agency. Among those who claim, foregone merited relief is modest: roughly $0.87 is left unpaid per dollar paid, and the loss tilts toward higher-income filers (accounting for 56.8% of foregone relief), who hold more claims. The loss on the participation margin is substantially larger. Consumers who experience an error but never claim are disproportionately lower income, and we estimate their awareness of the channel at roughly half that of higher-income households.

Finally, we use counterfactual experiments to rank the policy levers. On the firm side, consumer redress is limited by how cheaply firms can deny complaints about complex products. Universal self-advocacy and sharper monitoring and supervision raise total restitution ($42 million in 2024) by 7.54% and 14.04%, respectively, and mostly redistribute toward the already engaged. Every firm-side lever is dwarfed by the participation margin. Equalizing and accelerating awareness of the claims channel recovers an additional 38.49% per year, and does so progressively, with 91.62% of the gains accruing to below-median-income households. The incidence of firm mistakes is regressive not because firms discriminate, but because correcting mistakes is an ordeal for most consumers.

These results carry a general lesson about the limits of regulation as a consumer-protection technology. Enforcement changes how firms resolve the claims they receive, but it does nothing about the claims that are never made. Rules designed to protect consumers are therefore least effective for those who never enter the system. The welfare-maximizing intervention among the levers we consider is not more enforcement, but to increase awareness. More broadly, any market in which consumers must individually contest a firm's errors will allocate redress by the ability to contest them.

Related literature. Our paper connects four literatures that have developed largely in isolation: the public economics of take-up and ordeals, the law and economics of dispute resolution and enforcement, the economics of contracts and firm quality, and empirical household finance.

Ordeals, take-up, and the incidence of claiming costs. A central finding of public economics is that transfers requiring costly claiming go substantially unclaimed, with non-take-up concentrated among the least advantaged (Currie, 2006; Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019). Ordeal mechanisms can in principle improve targeting by screening on need (Nichols and Zeckhauser, 1982; Alatas et al., 2016), but in practice application costs often screen out the neediest (Deshpande and Li, 2019; Homonoff and Somerville, 2021), and administrative burden is regressive (Herd and Moynihan, 2018; Sunstein, 2019). We show that the same economics governs a large class of *private* transfers: restitution that firms already owe for their own errors. The process is simiar—a claims ordeal, sophistication-based screening, and regressive non-take-up—but the welfare accounting differs in an instructive way: unclaimed public benefits save

the government money, whereas unclaimed restitution is a pure transfer from harmed consumers to the firms that harmed them. Relatedly, even sophisticated agents leave money unclaimed when claiming is complex—firms themselves forgo tax refunds (Zwick, 2021)—underscoring that the friction is the claiming technology, not idiosyncrasies of household finance.

Dispute resolution, enforcement, and the selection of claims. The law-and-economics tradition analyzes which grievances become claims and which claims are paid as the outcome of private cost–benefit calculations that diverge from social ones (Priest and Klein, 1984; Kaplow and Shavell, 1994; Shavell, 1997; Spier, 2007), and studies the design of public enforcement (Becker, 1968; Polinsky and Shavell, 2000; Glaeser and Shleifer, 2003). Empirically, this literature has focused on courts and formal settlements (Eisenberg and Miller, 2015; Shavell, 1997); in consumer finance, Egan et al. (2025) analyzes arbitration outcomes involving uninformed consumers, Cheng et al. (2021) and LaVoice and Vamossy (2024) study debt-collection settlements, and Argyle et al. (2025) and Lee (2024) study bankruptcy outcomes. We contribute large-scale evidence on a *nonjudicial* redress channel—free, lawyer-less, and spanning the everyday products held by most households—where the selection frictions are informational rather than pecuniary. Furthermore, we provide plausibly causal evidence that enforcement pressure complements claimant sophistication, so that stricter enforcement can regressively reallocate redress even as it raises its level. Health economics has documented an analogous phenomenon on the provider side, where insurer payment denials impose large frictional costs on physicians (Dunn et al., 2024); we study the household side, where the claimant's capacity to contest denial varies systematically with income.

Contracts, firm quality, and the economics of mistakes. That complex, long-lived contracts are necessarily incomplete, leaving surplus to be allocated ex post, is a foundation of organizational economics (Williamson, 1975; Grossman and Hart, 1986; Hart and Moore, 1990; Tirole, 2009). Asymmetric information about quality and the role of reputation in disciplining sellers are crucial to such markets (Akerlof, 1970; Klein and Leffler, 1981; Shapiro, 1983; Board and Meyer-ter Vehn, 2013). Our contribution is to study the *error margin* of this problem empirically: not strategic quality choice, but routine operational failure and its ex post correction. Our finding that firm identity explains almost none of the variation in redress—while contract complexity explains nearly all of it—indicates that fallibility is a pervasive feature of production rather than the signature of low-quality firms, consistent with broader evidence that firms systematically deviate from frictionless optimization (Bloom and Van Reenen, 2007; Goldfarb and Xiao, 2011; DellaVigna and Gentzkow, 2019; Hortaçsu et al., 2019). The complaint channel we study provides consumers with a formal mechanism for seeking redress—an application of Hirschman (1970)'s "voice," whose local determinants are documented in Hayes et al. (2021). We provide, to our knowledge, the first large-scale accounting of what voice is worth, who exercises it, and who is silent.

Consumer financial markets. Within household finance, consumer harm has been attributed to behavioral mistakes by consumers (Gabaix and Laibson, 2006; Lusardi and Mitchell, 2014; Keys

et al., 2016; Gathergood et al., 2019; Jørring, 2024) or to exploitation by firms with informational or market power (Carlin, 2009; Schenone, 2010; Agarwal et al., 2015; Heidhues and Kőszegi, 2017; Flanagan and Purnanandam, 2024; Bhutta et al., 2026), with a prominent strand on adviser misconduct affecting wealthier households (Gurun et al., 2018; Dimmock et al., 2018; Egan et al., 2019). We introduce a third margin: even when consumers choose reasonable products and firms do not act strategically, services fail, and the correction of failures is itself a market—one whose outcomes depend on self-efficacy and engagement (Kuhnen and Melzer, 2018; Parise and Peijnenburg, 2019). Finally, a growing literature uses CFPB complaint data as a proxy for service quality or studies the disclosure of complaint narratives (Sedunov, 2020; Begley and Purnanandam, 2021; Fuster et al., 2023; Butler et al., 2023; Bian et al., 2023; Li, 2023; Dou and Roh, 2024; Dou et al., 2024; Huang et al., 2026; Kim et al., 2024; Mazur, 2024; Jou et al., 2024). In contrast to these papers, we observe complaint-level *outcomes*—decisions, dollars, handler identity, and the regulator's confidential interventions—and hence, to our knowledge, provide the first systematic measurement of the resolution of firm error, its incidence, and the welfare at stake for broader populations, claimants and non-claimants alike.

I Consumer Complaint Data from the CFPB

Since its inception in 2011, the Consumer Financial Protection Bureau (CFPB) has maintained a portal through which consumers can submit complaints against financial service providers. Approximately 80% of complaints are submitted via this online portal, with the remaining 20% arriving through email, fax, phone, postal mail, or referral. The submission process requires consumers to: (1) identify the relevant product or service, (2) describe the issue using both structured fields and free-form narrative, and (3) specify the company involved, the desired resolution, and provide contact information, as documented in Figure I.1. The CFPB then forwards the complaint to the company. The company is expected to respond to the complaint within 15 days.

The CFPB publishes a redacted version of these complaints in a publicly available data set. The publicly available data include the consumer's zip code (sometimes truncated to the first three digits), and indicators for whether the filer is elderly or a service member or veteran. The public data include the company name, product category, issue type, and response type (closed with explanation, closed with non-monetary relief, or closed with monetary relief).

We rely mostly on the CFPB's internal complaint database to enable a more comprehensive analysis of the complaint and dispute resolution process. Compared to the public dataset, the internal data offers granular filer information, more details about the complaint, and insights into how the company addresses the complaint. Key differences include, for example, (1) the dollar amount of any monetary relief provided, (2) whether the filer attached supporting documentation, (3) unique IDs for company employees handling each complaint, and (4) login records for both

consumers and company representatives. In contrast to the public database, full addresses, complaint narratives, and desired outcomes are always included in the internal data. Access to filer name and address allows us to merge the complaint data with Census tract-level socioeconomic indicators, in particular household median income.

We also leverage confidential information about the CFPB's internal review of company complaint responses. Since 2020, the CFPB has been assessing the quality of companies' responses to complaints, specifically, whether companies respond promptly and accurately. When conducting its reviews, the CFPB randomly selects a set of past complaints to the company. When the reviewers identify significant deficiencies, the CFPB issues confidential monitoring report cards to the relevant companies. Between 2020 and 2024, the CFPB sent approximately 80 such report cards to companies, which were delivered staggered over time. Some companies were sent report cards multiple times, and these too were delivered at different times. Overall, while the set of companies that received report cards is certainly not random, the timing with which they were sent the cards was. This setup allows us to use a difference-in-differences framework to test the effect of private regulatory pressure on company behavior.

Our analysis includes all non-credit-reporting complaints submitted between January 2014 and December 2024.[6] Since we primarily aim to understand the direct financial costs to consumers, we exclude the "credit reporting" category of complaints. These complaints are primarily directed at the three major credit reporting agencies and typically seek to resolve errors in people's credit reports. Though these credit reporting errors can have significant downstream costs—such as difficulties getting a loan—they rarely result in financial redress.

Figure 1 presents the different ways in which companies resolve complaints. Overall, a sizable number of complaints result in the consumer receiving financial compensation: 7.48% of complaints are resolved with monetary relief. The remainder are either closed with an explanation (80.55%) or closed with non-monetary relief (10.72%).

Figure 2 presents trends in the volume of complaints overall and across different product categories. Panel A shows a steady rise in total complaints, from 123,769 in 2014 to 368,969 in 2024, with an average annual growth rate of around 10%. Furthermore, as the overall number of complaints grew, the likelihood of a complaint resulting in monetary restitution remained relatively stable over time (around 7-8%). Panel B shows that the largest product categories in the CFPB complaints are *debt collection* (representing 33.55% of all complaints), *cards* (debit, credit, prepaid cards) (18.77% of all complaints), *mortgage* (16.16% of all complaints), and *accounts* (checking or savings) (15.48% of all complaints). They are also the fastest-growing categories over time, with debt collection, cards, and accounts witnessing significant increases over the last 5 years.[7]

[6] We exclude the complaints filed before 2014 because the complaint system had not yet incorporated all financial products.

[7] Debt collection complaints doubled in volume in 2024. The increase was driven largely by consumers using this product category to complain about erroneous information on their credit reports related to debt. Since credit

Smaller but still significant categories include complaints about money transfers (money transfer, virtual currency, or money service), student loans, car loans, and other loans (payday, title, personal loans).

II The Scope for Financial Restitution

Over the course of our sample (2014–2024), companies returned over $220 million in monetary relief to the 7.48% of complainants whose disputes ended in a payment. We describe the trends in that restitution and use them to gauge the scope for consumer losses from mistakes in financial services.

Table 1, Panel A, shows that the volume of complaints nearly tripled between 2014 and 2024, growing from 123,769 to 368,969 annually. The share resolved with monetary relief nonetheless remained stable between 7% and 8%. Restitution rates fell to 6.22% in 2024, but that decline follows a disproportionate increase in debt collection complaints, which rarely deliver financial restitution. The average relief amount rose over the same period from $697 to $2,295. Hence, most total monetary relief was distributed in the second half of the sample.

Panel B reports the distribution of monetary relief across products. Complaints about accounts and credit cards are the most likely to result in restitution, at 17.52% and 17.25%. Student loan and debt collection complaints are the least likely, at 2.07% and 0.67%. Other loans, such as mortgages and car loans, fall between these extremes.

These patterns describe a growing but underutilized channel for consumer redress. Complaint volume and average relief have both risen steadily while the share of cases paid has not, so the growth in the use of the CFPB system does not reflect a rise in frivolous or unsubstantiated claims. A free complaint channel might have attracted weaker claims as it became better known; the stable payment rate suggests it did not.

We return to the scale of these uncorrected losses in Section V, where a structural model of the firm's resolution decision measures the consumer welfare at stake, both for filers whose complaints are denied and for the far larger population who never file.

III Who Receives Financial Restitution?

Next, we examine what determines which complaints result in financial restitution. Two explanations compete. Some companies could be systematic bad actors whose misconduct disproportionately targets vulnerable consumers. Alternatively, errors could be a routine feature of administering

reporting is not a source of monetary relief, this dynamic mostly explains the decrease in the percentage of complaints receiving monetary relief in 2024 compared to all previous years.

complex contracts, in which case correction depends on whether the consumer recognizes the error and presses the claim. Which explanation holds determines whether consumer losses trace to company misconduct or to the difficulty of navigating the dispute process.

A Does Firm Identity Predict Restitution?

We begin with a variance decomposition to evaluate whether firm-specific misconduct drives restitution outcomes. Table 2, Panel A, regresses *solved wMR*, an indicator equal to one if the complaint was resolved with monetary relief, on combinations of quarter, company, product, and county fixed effects.[8]

In Column 1, quarter and company fixed effects together explain roughly 10% of the variation in restitution outcomes, with company fixed effects accounting for about two-thirds of that explanatory power. Once product fixed effects are included in Column 2, however, the model's explanatory power rises to 12.18%, and product type accounts for over 99% of the explained variation. Adding county fixed effects in Column 3 makes little difference. The dominance of product fixed effects holds when we restrict the sample to companies offering more than one financial product (Columns 4 to 6), so it does not reflect specialized firms operating in a single product market.

Restitution outcomes are therefore driven far more by the type of financial product involved than by which company is named in the complaint. This aligns with Table 1: straightforward products such as checking accounts and credit cards carry higher restitution rates, while term loans and other complex products are much less likely to result in compensation. Product complexity, not systematic firm misconduct, is the dominant source of variation in restitution. If firms do not systematically differ in how they resolve complaints, then consumer-level factors determine who receives redress, and we turn to those next.

B Differences in Consumers' Ability to Secure Restitution

Contract complexity is the dominant predictor of whether a firm's error is corrected, which suggests that redress turns on the consumer's capacity to recognize a valid claim and file it effectively. We therefore examine which consumers fare better in the complaint process. We focus on filer income because financial literacy varies systematically across the income distribution (Lusardi and Mitchell, 2014) and because the distributional stakes are largest along that margin.

Figure 3, Panel A, presents raw restitution rates by income quintile. The gradient is steep. Consumers from the top income quintile receive restitution in nearly 10% of cases, while those from the bottom quintile do so in fewer than 6%. Panel B plots the difference in restitution rates

[8] Decomposition percentages are based on Mean Squares, which account for the degrees of freedom associated with each explanatory variable.

between the highest and lowest income quintiles by product category. High-income consumers are more likely to receive monetary relief across almost all products, and the gap is largest for cards and money transfers. The disparity therefore persists within narrowly defined complaint types.

We formalize these patterns in Panel A of Table 3, which estimates complaint-level regressions of the probability that a complaint results in monetary relief:

$$solved\ w/MR_{i,c,q,r} = \gamma_q + \gamma_c + \beta_1 \times X_{i,q} + \varepsilon \tag{1}$$

The dependent variable equals one if complaint i against company c in quarter q was resolved with monetary relief by company employee r. The independent variable $X_{i,q}$ represents complaint and filer characteristics: an indicator for above-median census tract income in Columns 1 to 3, or the natural logarithm of median tract income in Columns 4 to 6. Specifications progressively load fixed effects. Columns 1 and 4 include quarter and company fixed effects, Columns 2 and 5 replace the latter with company-product fixed effects, and Columns 3 and 6 add the filer's county fixed effects.

The coefficient on income is stable across all six specifications, so the income gradient does not reflect compositional differences in geography, product, or provider. In our most saturated specification (Column 6), a one-standard-deviation increase in the filer's tract household median income raises the probability of receiving monetary compensation by 0.62%. This is economically meaningful against an average restitution probability of 7.48%.

Panel B of Table 3 extends the analysis to restitution amounts. We estimate fixed-effects Poisson models, since the relief amount is continuous and truncated at zero. Under the strictest specification, a complaint from an above-median-income tract receives 31.1% higher relief than one from a below-median-income tract. Income shapes not only whether relief is granted but also how much is awarded.

C Why Do Some Consumers Fare Better?

The previous section documented a substantial income gradient in financial restitution, with higher-income consumers more likely to receive compensation. Two mechanisms could generate these findings. Firms may treat complaints from high-income consumers more seriously, assigning more experienced personnel or investing greater effort in resolving their disputes. Alternatively, higher-income consumers may be more effective in how they file. We test these explanations using detailed information on both consumer behavior and firm engagement with each complaint.

Starting with consumer behavior, the CFPB complaint form includes two free-text narrative sections and an option to upload supporting documentation. We code the narratives for indicators of self-advocacy, including whether the filer explicitly requests a refund or cites potential fraud.[9]

[9] Please refer to Table II.b for more details.

Table 4, Panel A, shows that these strategies increase strongly with income. Filers from above-median-income tracts are 2.33 percentage points more likely to self-advocate in the complaint description and 1.34 percentage points more likely to do so in the desired-resolution field, and they are 0.37 percentage points more likely to attach supporting documentation. All three differences are precisely estimated under quarter, company-product, and county fixed effects.

We next investigate whether firms engage differently with complaints from high- and low-income consumers. The internal CFPB data identify the specific company employee who handled each complaint. For each handler, we construct two measures: employee experience, the natural logarithm of the number of complaints previously addressed, and employee login time, the percentage of days from complaint submission to resolution on which the handler logged into the complaint system. Panel B of Table 4 tests whether these characteristics vary with filer income.[10] Neither does. The estimated differences are economically negligible and statistically indistinguishable from zero, so firms do not allocate more experienced or more engaged staff to higher-income complainants.

Individual employees could still drive the gradient even if firms do not allocate them differentially. In Internet Appendix Table III.a, we add employee fixed effects to Equation 1. The income coefficient is unchanged, so the restitution gap does not reflect the particular employees who handle complaints.

Table 5 places these determinants in a single specification, augmenting Equation 1 with interactions between filer income and filer complaint actions, and between income and company employee actions. The table uses an indicator for filers in above-median-income tracts, while Internet Appendix Table III.b reports results with the continuous measure. Self-advocacy and documentation are the dominant channels. Relative to the 7.48% unconditional restitution rate, self-advocacy raises the likelihood of relief by 29% and attaching documentation raises it by 16% (column 1). Both returns are larger for high-income filers: self-advocacy is worth 34% to an above-median-income filer against 24% to a below-median-income filer, and documentation is worth 20% against 13% (column 2).

The interactions absorb the income gradient almost entirely. The income coefficient falls from 11% of the unconditional mean when advocacy enters additively to 4.6% once the interactions are included, and to 2.6% in the fully saturated specification, where it is statistically indistinguishable from zero. Among filers who neither self-advocate nor document their claim, in other words, income does not predict whether a firm pays. What income predicts is the propensity to do those things and the return to doing them.

[10] The number of observations is lower in Panel B because the respondent identifier is occasionally missing from the data.

Employee involvement affects outcomes even though firms do not allocate it differentially.[11] Employee experience bears no relation to restitution for the average complaint, but it enters positively for high-income complaints. Login time runs the other way: it is negatively related to relief, and the negative relation is concentrated among high-income complaints, while for low-income complaints it is indistinguishable from zero.

Income-based differences in complaint resolution therefore trace to consumer behavior. Higher-income consumers frame and support their complaints in ways that raise the likelihood of restitution, and we find no evidence of differential firm effort or triage based on consumer characteristics. Financial restitution is not automatically triggered by company error; it depends on how the consumer engages the dispute process.

IV Regulatory Scrutiny and Financial Restitution

Next, we study the effects of regulatory scrutiny on financial restitution. We leverage confidential data on the CFPB's internal monitoring reviews of the complaint process. Since 2020, the CFPB has monitored and reviewed companies' processing of complaints. Specifically, the CFPB assesses whether companies manage complaint volume effectively and whether they properly respond to complaints. In conducting these reviews, the CFPB analyzes a random selection of companies' complaints over the previous few years. Between 2020 and 2024, nearly 550 companies were monitored through reviews, amounting to a total of over 194,000 screened complaints. These numbers represent around 10% of both the number of all companies that have ever received a CFPB complaint and the number of total complaints, highlighting the broad scope of this CFPB monitoring effort.

Once the CFPB has screened a random complaint as part of the monitoring review, it assigns metrics that capture how well the company processed the complaint. In particular, we focus on the two internal metrics that value the company's response to the filer. $I(\textit{company not address compl.})_{i,c,q,r}$ is a dummy variable equal to 1 if the CFPB deems that company c did not properly address complaint i following its review. $I(\textit{response not substantive})_{i,c,q,r}$ is a dummy variable equal to 1 if the CFPB deemed that company c did not provide a substantive response for complaint i following its review. Overall, companies did not properly address around 5.25% of the reviewed complaints and did not provide a substantive response to 3.4% of them. If the CFPB finds that the company is processing complaints poorly, it then issues the company a confidential monitoring report card highlighting the deficiencies. The CFPB sent 85 such cards between 2020-2024 and the complaints were delivered staggered over time.

[11] The number of observations in these columns is lower because the respondent identifier is occasionally missing from the data.

In the following section, we use these confidential monitoring report cards to establish an objective benchmark for income differences in financial restitution. Then, we assess whether regulatory pressure affects company behavior and financial restitution outcomes.

A CFPB Company Monitoring Reviews

By looking at the reviewed complaints that the CFPB flags (or not) as properly processed, we can gather information on the optimal benchmark for complaint processing. Further, having found differences in restitution across income, we can use this information to assess whether those differences are driven by companies not correctly processing complaints from lower-income filers.

We find that complaints from low-income filers are relatively more likely to be poorly addressed or receive an unsatisfactory response, especially at companies that receive a monitoring report from the CFPB. Figure 4 plots $I(\textit{company not address compl.})_{i,c,q,r}$ (Panel A) and $I(\textit{response not substantive})_{i,c,q,r}$ (Panel B) by quintiles of the reviewed complaints' census tract household median income. Within each Panel, the graphs on the left pool together all reviewed complaints, while the graphs on the right distinguish between reviewed complaints that resulted in a confidential monitoring report card being sent to the company (highlighting the deficiencies) and reviewed complaints that did not result in a confidential monitoring report card being sent to the company. The graphs depict a clear monotonic trend that is opposite the patterns of financial restitution: low-income filers are more likely to have their complaints improperly handled, especially at companies that end up receiving a monitoring report.

However, the disparities in complaint processing are much smaller than the income differences for filers receiving financial restitution. Comparing Figure 4 with Figure 3, complaints in the top quintile of filer tract income are $\sim$20% less likely to be flagged as poorly handled in investigations than complaints in the bottom quintile, while they are $\sim$65% more likely to receive financial restitution.

Table V.a uses regression analysis to test for differences in complaint handling across income groups. These regressions allow us to control for quarter, company, and product fixed effects. We find that the inclusion of company or product fixed effects eliminates differences in complaint handling across incomes. The difference in incomes becomes small and indistinguishable from zero with company fixed effects or negative with product fixed effects. These findings suggest that companies are not treating income groups differently, but rather, companies with more complaints from certain income groups are more likely to have deficiencies in their handling of complaints.

Overall, there is only weak evidence that differences in financial restitution are driven by companies failing to process low-income complaints properly, and the primary driver of the differences remains how consumers themselves engage with the complaint process (self-advocating, attaching supporting documentation, etc.).

B Effects of CFPB Monitoring and Enforcement

Next, we analyze how companies respond to receiving a confidential monitoring report card from the CFPB. The increased regulatory scrutiny by the CFPB potentially affects the propensity for companies to deliver financial restitution, and their response may differ across filers from different income groups.

Internet Appendix Figure IV.1 provides information on the companies that received a confidential monitoring report card from the CFPB between 2020 and 2024. Panel A highlights how 85 companies out of the over 550 companies reviewed between 2020 and 2024 received a confidential monitoring report card, as the CFPB sends reports only to the reviewed companies that display poor complaint-handling standards. Panel B shows how these companies usually account for around 0.5% of all complaints.

We use the staggered distribution of these confidential monitoring report cards to conduct a difference-in-differences analysis of the effects of increased regulatory scrutiny on companies' financial restitution patterns. We use the companies that receive confidential monitoring report cards as our treated sample. We restrict our control sample to the companies that were reviewed but never received a report, as long as they had at least 100 yearly complaints in 2018 (prior to the review process). This restriction on the control sample accounts for the fact that companies that receive the negative cards tend to be larger. Panel A of Internet Appendix Table V.b presents the balance t-tests for the treated and the control groups prior to the start of monitoring reviews in 2020. Even imposing a size threshold, treated companies tend to receive, on average, significantly more complaints. Treated companies also tend to attract more deposit-related complaints than control companies (suggesting they are more likely to be banks than specialized entities), and witness a higher rate of filer attachments. However, there is no difference between treated and controls in how many complaints these companies get from low-income filers, on average. Panel B presents balance t-tests for the early-treated companies (receiving confidential monitoring report cards in 2020-2022) and the late-treated companies (receiving confidential monitoring report cards in 2023-2024). The CFPB first targeted larger banks, then expanded to smaller banks and other financial institutions, which happen to get a higher percentage of complaints from low-income filers, on average.

We apply a difference-in-differences methodology to the complaint-level data of these companies:

$$(2)\quad solved\ w/MR_{i,c,q,r} = \gamma_c + \gamma_q + \beta_1 \times I(post\text{-}report)_{i,c,q} + \beta_2 \times X_{i,t} + \beta_3 \times I(post\text{-}report)_{i,c,q} \times X_{i,q} + \varepsilon$$

In this specification, $solved\ w/MR_{i,c,q,r}$ is a dummy equal to 1 if complaint i against company c in quarter q has been solved with monetary relief by company employee r. $I(post\text{-}report)_{i,c,q}$

is a dummy equal to 1 if complaint i is filed against company c in quarter q, after the company has received a confidential monitoring report card from the CFPB.[12] $X_{i,q}$ is a high-income dummy based on complaint i's filer census tract and quarter q.

We find that after receiving a confidential monitoring report card, companies increase the share of complaints resolved with monetary relief by several percentage points. However, they do so mostly for complaints filed by high-income filers. In Table 6, we observe non-significant coefficients on $I(post\text{-}report)_{i,c,q}$, but positive and significant coefficients on the interaction between $I(post\text{-}report)_{i,c,q}$ and the income variable. This result suggests that instead of narrowing the income gap in restitution, heightened regulatory scrutiny widens it.

We find that self-advocacy increases the likelihood of receiving financial restitution by 23.80% in normal times but by 50.09% after the company receives a CFPB monitoring report card.[13] We find that attaching supporting documentation increases the likelihood of receiving financial restitution by 15.64% in normal times but by 37.30% after the company receives a CFPB monitoring report card.[14] Further, the post-report financial restitution likelihood with attachments gets even higher if the complaint is filed from an above-median-income tract (column 7).

Since our treatment effects are heterogeneous over time and potentially across groups, we also employ the Callaway and Sant'Anna (2021) estimator in place of the two-way fixed effects diff-in-diffs model. Figure 5 reports the corresponding coefficient plots. Panel A distinguishes between the share of complaints solved with monetary relief coming from below-median-income tracts (left graph) and the share of complaints solved with monetary relief coming from above-median-income tracts (right graph). Looking at the complaints from above-median-income tracts, there is a marked increase in the likelihood of financial restitution after the delivery of the confidential monitoring report card. However, for the complaints filed from below-median-income tracts, the uptick is delayed and much smaller, validating the diff-in-diffs results.

We further replicate the Callaway and Sant'Anna analysis using a tighter control sample. While so far we have kept all investigated companies with at least 100 complaints as controls, now we further restrict control companies to the three nearest neighbors of the treated ones based on overall number of complaints, share of complaints regarding loan-related products, and share of complaints with filer attachments.[15] Under this more balanced sample, the post-monitoring report increase in restitution likelihood for high-income complaints but not low-income ones is even more apparent.

[12] $I(post\text{-}report)_{i,c,q}$ is the equivalent of *Treated x Post* in the traditional difference-in-differences setup, while *Treated* and *Post* are being absorbed by company and quarter fixed effects.

[13] From columns 4 and 5: (0.0178 + 0.0203) / 0.0748 = 50.09% and (0.0178 + 0.0203) / 0.0748 = 50.09%).

[14] From columns 6 and 7: 0.0117 / 0.0748 = 15.64% and (0.0117 + 0.0162) / 0.0748 = 37.30%.).

[15] These were the three dimensions on which the previous treated and control samples differed the most according to the balance tests in Table V.b.

Lastly, we explore whether these findings are due to a change in the composition of complaints received by firms that receive monitoring report cards by the CFPB. To do so, Internet Appendix Table V.c replicates the difference-in-differences analysis but uses a low-income complaint dummy as the dependent variable. We find that receiving a monitoring report does not meaningfully change the income composition of people who complain about a company's financial products.

Overall, under increased regulatory scrutiny, companies seem to improve their responsiveness primarily to those consumers who were already more likely to secure restitution. This finding further supports the role of financial aptitude and self-advocacy in determining restitution outcomes—consumers who can effectively navigate the dispute resolution process are best positioned to benefit from regulatory oversight.

V A Structural Model of Complaint Resolution

Sections III and IV establish three facts: restitution falls sharply with product complexity; among filers, self-advocacy rather than income affects the firm's restitution decision; and CFPB scrutiny raises resolution rates, but disproportionately for engaged filers. This section develops a structural model of the firm's complaint-resolution policy to ask how much consumer welfare is foregone when meritorious complaints go unredressed, and how that loss responds to broader self-advocacy, wider monitoring coverage, stronger baseline enforcement, and wider awareness of the complaint process.

Our exercise starts with a relatively standard static random-utility model of the firm's grant-or-deny decision. This prices the welfare lost by consumers who file and are denied, the filer margin (intensive margin). Next, we augment the model with a calibrated awareness-diffusion process that recovers the size of the non-filing population, which we combine with the estimated policy function to value their foregone relief.

A The Economic Environment

Setting and observables. Time is static on the firm's side: each complaint is a single grant-or-deny decision, and a household files once per error. There is a finite set of financial products $p \in \mathscr{P}$. A complaint i is described by its product p_i, the filer's income type $\theta_i \in \{0,1\}$ (an indicator for an above-median census-tract household income), the complaint's quality $q_i \in \{0,1\}$ (equal to one for a self-advocated complaint, meaning one that cites fraud, requests a refund, or attaches documentation), the firm's CFPB monitoring report card state $\omega_i \in \{0,1\}$ (equal to one for the

company having previously received a report card), and the firm's size class $\text{FS}_i \in \{0,1\}$ (equal to one for larger firms).[16]

Underlying each complaint is a latent grievance g_i, the dollar magnitude of the financial error. The grievance is a property of the underlying transaction, fixed before the complaint is filed and independent of how the complaint is presented. A complaint is *meritorious* if $g_i \geq \bar{g}_p$ for a product-specific threshold $\bar{g}_p$. We do not estimate g_i, because our confidential data record realized restitution only on resolved complaints, which is a noisy signal of g_i. The model works with the grievance only through the *merit probability*

$$\phi(q,\theta,p) \;\equiv\; \Pr(g_i \geq \bar{g}_p \,\big|\, q,\theta,p)\,, \tag{3}$$

the probability that a denied complaint was in fact owed redress, which is the object that enters welfare considerations in Subsection C.

The household's problem. Conditional on being aware of the CFPB channel, a household makes two nested choices: whether to file, and, if filing, at what quality q (with or without self-advocacy). The expected utility of filing at quality q is

$$U_i(q) \;=\; \sigma(q,\theta_i,p_i;\omega_i)\,\mu(g_i,p_i) \;-\; c(q,\theta_i,\kappa_{p_i}) \;+\; \varepsilon_i^q, \tag{4}$$

where $\sigma(\cdot)$ is the firm's resolution policy, taken as given by the household and derived below; $\mu(g_i,p_i) = \mathbb{E}[\,r_i \mid g_i,p_i,\text{grant}\,]$ is the expected restitution amount conditional on a grant (the prize, a property of the underlying grievance and product and independent of q); and $c(q,\theta,\kappa_p) = c_q\,q + c_{q\theta}\,q\theta + c_{q\kappa}\,q\kappa_p$ is the cost of assembling a complaint of quality q, with $c_q > 0$ (effort rises with quality), $c_{q\theta} < 0$ (higher income and literacy lower the marginal cost of self-advocacy), and $c_{q\kappa} > 0$ (κ_p capturing product complexity, with complex products being costlier to document). The taste shock ε_i^q is i.i.d. Type-I extreme value, and the outside option of not filing carries a normalized payoff of zero.

The household upgrades from a basic to a self-advocated complaint when the marginal expected benefit of doing so weakly exceeds its marginal cost; with extreme-value shocks this quality choice is a binary logit in the net benefit of upgrading. The model follows three sign predictions: the probability of self-advocacy rises with income ($c_{q\theta} < 0$), falls with product complexity ($c_{q\kappa} > 0$), and rises with the size of the prize. All three are visible directly in the data—self-advocacy increases with income and decreases with complexity (Table 4).

[16] The variable ω_i equals one if the company has received a monitoring report card from the CFPB over the previous year. FS_i is a time-invariant firm attribute equal to one for companies spanning at least three products *and* clearing at least 250 complaints in every year they appear. It proxies for firm scale itself, and the supervisory intensity that comes with scale, since these are the firms the CFPB is most likely to examine. We write "large" and "small" for $\text{FS}_i = 1$ and $\text{FS}_i = 0$ throughout.

We do not estimate the household problem in its own right, but previous empirical analysis validates it, and it provides the economic foundation for the CFPB awareness and firm problems. First, the filing decision—the participation margin—cannot be recovered from complaint records, which contain only complaints that were actually filed; households that experience an error but never file never appear. We therefore delegate participation to the awareness-diffusion model of Subsection D, whose product of awareness and filing-given-awareness is the participation probability of equation (4) marginalized over the grievance. Second, the quality margin supplies the economic reading of q and the conditional exogeneity that licenses treating q as a covariate in the firm's policy function: self-advocacy reflects the filer's skill and literacy and is conditionally independent of the underlying grievance.

The firm's problem. For each complaint of type (q, θ, p) at a firm in monitoring report card state ω and size class $c \in \{\text{large}, \text{small}\}$, the firm chooses to grant or deny so as to minimize its expected cost. Granting costs the restitution paid, R_p, which is product-specific and observed in the data. Denying is costly through regulatory exposure: denying a complaint that is in fact meritorious raises the firm's poorly-handled rate and hence the probability that the CFPB escalates. Crucially, that exposure is not uniform across firms—the Bureau's supervisory attention is greater for firms of scale, so a firm's size class shifts both how sensitively a poorly-handled rate translates into a report and how large an escalation is. The firm's expected dollar cost of denying is

$$D(q, \theta, p; \omega, c) \;=\; \underbrace{\phi_0(p)\, e^{\varphi_q q + \varphi_\theta \theta}}_{\phi(q,\theta,p)} \quad \underbrace{\rho_0'(c)\, \Phi(c)}_{\text{enforcement exposure}} \quad \cdot \underbrace{e^{(\delta_{q\omega} q + \delta_{\theta\omega}\theta)\,\omega}}_{s(q,\theta,\omega)}, \tag{5}$$

where $\rho_0'(c)$ is the baseline marginal probability of receiving a monitoring report card per unit of poorly-handled complaint, $\Phi(c)$ is the dollar penalty of an escalation, $\phi_0(p)$ is the baseline merit probability, and $\varphi_q, \varphi_\theta \geq 0$ are the merit log-gradients in quality and income. For a small firm, which the regulator rarely examines, ρ_0' is close to zero: a denial generates almost no regulatory exposure, whatever the complaint looks like. The monitoring scale $s(q, \theta, \omega)$ equals one whenever $\omega = 0$ or $q = \theta = 0$: stricter monitoring sharpens the penalty for denying an *engaged* complaint but leaves a baseline complaint's deny cost untouched. This regulatory-complementarity restriction is taken from our finding of no post-report effect for low-income, low-advocacy complaints (Section IV).

The firm grants whenever the deny cost weakly exceeds the grant cost, $D/R_p \geq 1$. Taking logs, this threshold rule is "grant iff $\eta_i \geq 0$," where the index

$$\eta_i \;=\; \alpha_{p(i)} + \gamma \mathrm{FS}_i + \big(\varphi_q + \delta_{q\omega}\,\omega_i\big) q_i + \big(\varphi_\theta + \delta_{\theta\omega}\,\omega_i\big)\,\theta_i. \tag{6}$$

The deterministic threshold would make every cell all-grant or all-deny, whereas observed within-cell resolution rates lie strictly inside $(0, 1)$. We therefore add an idiosyncratic, econometrician-

invisible shock to the firm's perceived *log cost ratio*—the case officer's private read, internal documentation, processing idiosyncrasies—distributed standard logistic. Integrating over it, the resolution probability is a logit on the index,

$$\sigma_i \;=\; \Pr(\text{solved w/MR} \mid q_i, \theta_i, p_i; \omega_i, \text{FS}_i) \;=\; \Lambda(\eta_i)\,, \qquad \Lambda(x) = \frac{e^x}{1+e^x}, \tag{7}$$

which is the policy function we estimate.[17] The baseline log cost ratio separates additively into a product component and a firm-size component,

$$\alpha_p \;=\; \log\frac{\phi_0(p)\,\rho_0'\,\Phi}{R_p}, \qquad \gamma \;=\; \log\frac{\rho_0'(\text{large})\,\Phi(\text{large})}{\rho_0'(\text{small})\,\Phi(\text{small})} \;-\; \log\frac{R_p(\text{large})}{R_p(\text{small})}, \tag{8}$$

with α_p evaluated at a small firm and γ the log ratio by which enforcement exposure rises at a large one.

Equation (8) is the model's central reading of product complexity. The intercept α_p is the log ratio of the firm's baseline expected deny cost to grant cost, capturing its enforcement exposure to the restitution it would pay on a reference complaint ($q = \theta = \omega = \text{FS} = 0$): $\Lambda(\alpha_p)$ is the baseline resolution probability, and its complement $1 - \Lambda(\alpha_p)$ is the baseline denial rate—how much the firm can deny with impunity. A very negative α_p means the expected enforcement cost of denying is only a fraction of the restitution owed ($\phi_0(p)\rho_0'\Phi \ll R_p$), so the firm denies even meritorious complaints.[18] We therefore summarize the baselines by a complexity index normalized to the simplest product,

$$\kappa_p \;=\; 1 - \frac{\Lambda(\alpha_p)}{\max_{p' \in \mathscr{P}} \Lambda(\alpha_{p'})}, \tag{9}$$

so that "α_p," "product complexity κ_p," and "what the firm gets away with" are three readings of the same estimated object. Because the four dollar primitives $\phi_0(p), \rho_0', \Phi, R_p$ enter only through the single ratio in (8), the data identify α_p but none of the four separately.

The regulator. The CFPB samples resolved complaints, flags those that were meritorious but denied, and escalates firms whose poorly-handled rate is high—the mechanism that makes regulatory exposure costly to the firm. In the estimated model the regulator enters in two reduced-form ways: through the firm's size class FS_i, which sets the standing level of that exposure, and through the firm's monitoring report card state ω_i, which steepens the returns to engagement. We do not separately estimate the regulator's sampling primitives ρ_0' and Φ; they enter only through the ratios α_p and γ.

[17] Because the shock lives on a dimensionless log ratio, its scale is a free normalization and the model carries no separate dispersion parameter; this is also why the dollar stake R_p drops out of σ entirely and reappears, in levels, only in welfare.

[18] This is the complex, low-resolution end of the product space (student loans, auto loans, mortgages).

Summary. The environment delivers a six-parameter policy function, $\{\alpha_p, \gamma, \varphi_q, \varphi_\theta, \delta_{q\omega}, \delta_{\theta\omega}\}$: a full set of product baselines α_p that read as complexity, a firm-size baseline shifter γ, merit gradients φ_q and φ_θ in self-advocacy and income, and two monitoring interactions $\delta_{q\omega}, \delta_{\theta\omega}$ that capture regulatory complementarity. Subsection B takes equation (7) to the complaint-level data.

B Estimating the Policy Function

Data and sample. The estimation uses the CFPB Consumer Complaint Database at the complaint level over 2015–2024, linked to household median income (which defines θ_i), to the Bureau's monitoring report card timing (which defines ω_i), and to a firm-level size indicator (which defines FS_i, equal to one for companies spanning at least three products and clearing at least 250 complaints in every year they appear in the complaint database). The outcome is whether the complaint was closed with monetary relief, and the dollar stake R_p is the mean realized relief amount among resolved complaints of product p. Credit report and debt-collection complaints are dropped before any estimation: they rarely involve an in-kind service obligation—their relief rate is under one percent—and are largely mis-filed or credit-reporting issues, so reading their resolution rate as a firm deny-cost parameter would be misleading. Monetary transfer complaints are also dropped due to their more mechanical resolution dynamics. Dropping these complaints takes the raw panel to 1,034,613 complaints across six products and 4,618 firms, which is the population used for welfare calculations.

Estimation itself runs on the intersection of two further restrictions the data require. Complaint quality q_i is coded from the complaint narrative, so estimation is confined to the publicly available narrative subsample (roughly 47% of the panel); and the monitoring report card state ω_i varies meaningfully only after the CFPB's systematic review campaign begins in 2020, so we restrict to the post-2020 window. This leaves 266,287 complaints at 2,443 firms. Standard errors are clustered at the company level, since complaints are correlated within firm and both monitoring and size class are attributes of the firm. Table 7 reports the estimates and Table 8 the implied product complexity.

Estimator. We estimate equation (7) by maximum likelihood, treating each complaint as an independent grant/deny draw with success probability $\sigma_i = \Lambda(\eta_i)$. Maximum likelihood is the natural estimator for a discrete-choice model of this kind, in which the logistic form is part of the economic specification—the distribution of the firm's idiosyncratic grant/deny payoff. Every parameter block is estimated freely, and the model is over-identified in the sense that it fits far more cell-level moments than it has parameters, and we assess that fit directly, on moments the estimator does not target, below.

The model reproduces the reduced form. The product baselines are uniformly negative and precisely estimated ($|t|$ between 6 and 18), and they recover the complexity ordering the reduced form establishes, from cards and accounts at the simple end to mortgages, auto loans, and student

loans at the complex end (Table 8). Read through equation (8), the implied baseline resolution rate—roughly 11% for cards against under 1% for student loans—measures what a firm can deny with impunity. Complex products are not ones where errors are rarer; they are ones where denial carries little enforcement cost relative to the restitution at stake, so even meritorious complaints are denied. The self-advocacy gradient is large and precise, $\widehat{\varphi}_q = 0.257$ (standard error 0.072), raising the resolution probability by about three percentage points at the average complaint, in line with Table 5. A substantiated complaint raises the firm's posterior that a denial would be judged wrongful, and so raises the deny cost. The income gradient, by contrast, is small and statistically indistinguishable from zero, $\widehat{\varphi}_\theta = 0.072$ (standard error 0.062). The firm attaches no additional weight to a complaint because it comes from a higher-income tract, once quality is held fixed, which is what one would expect of a firm that responds to the content of a complaint and does not observe the filer's income. This does not contradict the income gradient of Section III; it decomposes it. The structural model results confirm that income raises restitution because higher-income filers self-advocate more, not because firms grant their complaints a premium.

The firm-size shifter and the two monitoring interactions are identified off variation that is not as good as random, and we read them accordingly. Size class is a between-firm contrast, estimated at $\widehat{\gamma} = +0.519$ (standard error 0.332): denying is materially cheaper at small firms at every product—baseline resolution of 0.010 against 0.016 for student loans, 0.114 against 0.178 for cards (Table 8)—but the magnitude is imprecise, and we retain it on fit grounds rather than treating it as sharply measured.[19] Which firms are under report cards at a point in time is selectively determined, only 47 firms are ever observed switching state, and neither interaction is sharply estimated: $\widehat{\delta}_{q\omega} = +0.0719$ with a 95% confidence interval of $[-0.152, +0.296]$, and $\widehat{\delta}_{\theta\omega} = +0.035$ with $[-0.132, +0.201]$. We therefore place the fitted channel beside the staggered-timing difference-in-differences of Section IV, whose identifying variation is the timing of a confidential report rather than the fact of being reviewed. The model's implied difference-in-differences runs above that benchmark on the advocacy channel (+0.0249 against +0.0203) and below it on the income channel (+0.0096 against +0.0174). The gaps bound how much weight the monitoring channel can carry rather than correcting it.

Non-targeted moments. We validate the model on moments it was not estimated to match. The most relevant test is the income-by-advocacy interaction: the index carries no $q \times \theta$ term, yet the curvature of Λ alone generates a model difference-in-differences of +0.0026 against an empirical +0.0017 on this sample (Table VI.a, panel C), so the specification reproduces the com-

[19] Among large firms, 24 of 34 are ever monitored and $\omega = 1$ some 60% of the time; among small firms, 25 of 2,409 and 6.5% of the time, so 92.6% of all monitored complaints are filed against large firms. There is effectively no FS $\times\, \omega$ variation from which to learn how a small firm responds to a report card. What the data can speak to is the level difference between the classes, which is why size class enters the baseline through γ and monitoring enters the slopes through the two δ's. Re-estimating on the $\omega = 0$ subsample alone, where no firm is currently carded, gives $\widehat{\gamma} = +0.547$ against +0.527 in the full fit: what separates the classes survives with monitoring switched off everywhere, which is what makes it a baseline rather than a monitoring effect wearing its clothes.

plementarity (while just falling short of the larger +0.008 benchmark of Table 5, estimated on the full complaints database). Regarding other derived moments, every moment involving ω, is hit to within a percentage point (Table VI.a, panels A and B). Because there is no standalone ω term, the model under-delivers the observed monitoring lift there; and because there is no product-by-ω term, the common log-odds lift under-delivers for exactly the low-baseline products where the data show the largest relative monitoring responses (Table VI.a, panel D). We view this as the cost of parsimony, not as a misspecification, since reduced-form results showed no significance of the standalone ω channel (Section IV, Table 6, panel B). [20]

C Foregone Consumer Welfare

The policy function helps value complaints that the firm denies. Among denied complaints, some were meritorious—the firm should have paid—and we consider the model's fitted resolution probability $\widehat{\sigma}_i$ the weight the model attaches to this event, since a firm that grants meritorious complaints resolves a type-$(q,\theta,p;\omega,\mathrm{FS})$ complaint with probability $\widehat{\sigma}_i$. We therefore measure the expected foregone restitution on a denied complaint as a function of the structural model's $\widehat{\sigma}_i$, and sum over all denied complaints to estimate foregone consumer welfare as follows:

$$W^{\text{denied}} = \sum_{i:\,\text{denied}} m \cdot \widehat{\sigma}_i \cdot R_{p,t,\omega}(i). \tag{10}$$

where m is a multiplier indexing how much more likely a denied complaint is to have been meritorious than the firm's grant rate would suggest,[21] and $R_{p,t,\omega}$ is the mean restitution among resolved complaints of the relevant product-year-monitoring bucket. [22]

Because the estimation sample is the publicly available narrative subsample restricted to the post-2020 time frame but welfare is a statement about all complaints, we scale to the full complaint data by imputing complaint quality deterministically. Each complaint inherits the self-advocacy rate $\Pr(q=1 \mid p,t,\omega)$ of its product-year-monitoring report card bucket estimated on the publicly available narrative subsample, and the firm's policy function is evaluated at that continuous imputed value. Foregone welfare is then summed over the complaints that were *actually* denied, which is observed for every complaint, so the denied set is known exactly and no policy-function fit error enters the shares. A back-of-the-envelope check supports the calculation's credibility: the model reproduces the dollars actually paid. Expected relief $\sum_i \widehat{\sigma}_i R_i$ on the full complaint data is

[20] For a more detailed discussion of model fit, please refer to Appendix VI Subsection VI.

[21] $m=1$ is the conservative anchor, in which the merit rate equals the resolution rate, and we report $m \in \{1, 1.5, 2\}$ throughout.

[22] We let average restitution amounts vary with the monitoring report card state ω because monitoring report cards raise realized relief amounts: a Poisson pseudo-maximum-likelihood regression of relief on complaint characteristics with product fixed effects puts the monitoring multiplier on expected restitution at $e^{0.539} \approx 1.71$, and the granted-cell ratios exceed one for every product (from 1.15 on auto loans to 1.94 on student loans).

$176,391,001 against $176,433,143 actually disbursed, matching nearly perfectly, and the fit holds separately for large and small firms (ratios 0.994 and 1.025).

Among filers, the complaint system leaves roughly $148 million of merited relief unpaid over 2015–2024 at $m = 1$ (bottom of panel A, Table 9)—about $37 million on the 2024 cohort alone (bottom of panel B, Table 9), reflecting today's higher complaint volumes—and proportionally more at higher m. Measured against the relief it actually delivers, the system leaves about $0.84 of merited relief unpaid for every dollar it pays. The loss is not evenly distributed: roughly 56% accrues to higher-income tracts and 44% to lower-income tracts. Among filers, in other words, the foregone welfare tilts toward the better-off—precisely because they file and self-advocate more, so more merited-but-denied dollars are theirs to lose.[23] The larger and more regressive loss, however, is borne by the consumers who never file at all, which we turn to next.

D Awareness Diffusion and Non-Filer Welfare

The welfare in equation (10) is conditional on filing. But the participation margin may be the more important one: the majority of consumers who experience a financial error never file a complaint, in part because they are unaware that the channel exists. Complaint records cannot speak to that population by construction, so we build the universe of potential filers separately, with an awareness-diffusion model that supplies the object the household's participation first-order condition requires.

The filing probability. The model factors the probability that a single problem episode becomes a filed complaint into three interpretable pieces,

$$\Pr(\text{file} \mid \text{problem with product } p,\ \text{income tier } \theta,\ \text{month } t) \;=\; a_\theta(t) \cdot \zeta_p \cdot \lambda, \tag{11}$$

where $a_\theta(t)$ is the share of income tier-θ households that know the CFPB channel exists, ζ_p is a product complaint-relevance weight normalized to mean one—how readily a problem with product p is treated as a CFPB matter—and λ is a common filing-given-awareness rate. This factorization is the model's substantive content. Knowing that the CFPB exists is a fact a household learns once, never unlearns, and carries to every product it holds; whether a mortgage servicing error reads as "a CFPB matter" is a property of the product, not of the household's information. Loading both onto a single per-cell number, as a product-specific awareness state would, leaves neither interpretable.

Awareness diffusion. The model keeps three populations distinct: the fixed stock of households in each income tier, the fixed stock of holders of each product within a tier, and the annual *flow* of problem episodes, which is re-drawn every year. Awareness is a permanent stock carried on the fixed stock of households in each income tier; filings are read off the annual flow of prob-

[23] This is the within-filer mirror of the income gradient: the same engagement that secures relief also concentrates the residual losses among the engaged.

lem episodes. In the first layer, awareness diffuses logistically by word of mouth over the whole income tier population, at a tier-specific rate,

$$a_\theta(t+1) \;=\; a_\theta(t) \;+\; \beta_\theta\, a_\theta(t)\big(1-a_\theta(t)\big), \tag{12}$$

one curve per income tier, shared by every product in that tier. In the second layer, each year's $N^{\text{prob}}_{p,\theta}$ problem episodes split into unaware and aware households in proportion to the population share: $S_{p,\theta}(t) = N^{\text{prob}}_{p,\theta}\big(1-a_\theta(t)\big)$ for unaware ("susceptible") households and $I_{p,\theta}(t) = N^{\text{prob}}_{p,\theta}\, a_\theta(t)$ for aware ("infected") households, with aware households filing at rate $\lambda\zeta_p$. Written in headcounts these are the standard epidemiological laws of motion of a Susceptible, Infected, and Recovered model, with the contagion rate common to all products within a tier and cumulative filings a read-only accumulator (keeping tally of the "recovered" households), so the awareness system is closed and converges as $a_\theta \to 1$.

Data sources. The diffusion model draws on four inputs outside the complaint data. Household counts by income tier come from public national statistics;[24] product penetration by tier—what share of households hold each product—comes from the 2022 Survey of Consumer Finances; per-product annual problem incidence comes from consumer-problem surveys and related sources (Internet Appendix Table VII.a); and the income tier diffusion rates β_θ are calibrated from income tier-aggregate complaint growth over 2019–2023, deliberately excluding 2024 so that the first projection year can be validated against data the calibration did not see.

Identification. Under stationary problem flows and filing rates, total tier complaints move one for one with awareness, so complaint growth identifies awareness growth: this gives lower-income tier $\beta_0 = 0.01194$ and higher-income tier $\beta_1 = 0.01156$ per month, or 14.34% and 13.88% growth per year as per complaint data. The remaining objects are identified from the terminal complaint cross-section alone, which supplies one equation per product-income tier cell. Their overall level is not separately identified—$a_\theta(0)$, ζ_p and λ enter only as a product—so the λ grid absorbs it and we run simulations across it. What *is* identified is the structure. Complaint intensity per problem-household, taken within a product and across income tiers, cancels both λ and ζ_p and returns the relative tier awareness $\rho = a_0(0)/a_1(0)$ directly; because each product yields an independent estimate, the spread across products is a built-in specification test of the universal-awareness premise, and it is reasonably tight, with a pooled estimate of $\rho = 0.771$, as reported in Internet Appendix Table VII.b. The cross-product pattern of intensity then identifies the weights ζ_p up to scale, and they span an order of magnitude: mortgage problems are taken to the Bureau some 2.5 times more readily than the average product, auto-loan problems less than a quarter as readily. Finally, each income tier's seed is backed out from its own complaint total against a ζ_p-

[24] 85.9 million below-median and 45.4 million above-median households, roughly 131 million in total, since household median income is defined within the universe of complaints, and sits higher than nationally.

weighted problem base, so the model reproduces the below-median and above-median complaint counts exactly.

The awareness gap is large, and organic diffusion closes it slowly. At a reference filing-given-awareness rate of $\lambda = 0.5$, the model puts base-year awareness for below-median income households at 0.75%, against 1.22% for above-median income households—a near two-to-one gap—rising over ten years to 3.06% and 4.69% (Internet Appendix Table VII.c). Because the two income tiers start apart and diffuse at nearly the same rate, the logistic *widens* the absolute gap before saturation eventually closes it. An unaware household forgoes the relief it would have obtained in expectation had it filed, $m \cdot \widehat{\sigma}_{p,\theta} \cdot R_p$, with $\widehat{\sigma}_{p,\theta}$ the policy function's cell-level resolution probability held pinned at its estimated value—so the diffusion changes *how many* households are unaware, while the dollar value of each unaware household's foregone complaint stays anchored to the structural model. Evaluated on the base-year stock of unaware problem-households, that quantity runs to roughly \$8.5 billion a year at $m = 1$ and $\lambda = 0.5$ (bottom of Table 9). We treat this strictly as a bounding calculation rather than a point estimate: it scales one-for-one with the per-product problem-incidence shares—the most assumption-laden inputs to the model, drawn from heterogeneous external surveys—as well as with λ and m. Its role is qualitative but important, in that it establishes that the participation margin, not wrongful denial among those who show up, is where the recoverable consumer welfare lies.

The quantity we emphasize instead is the flow this margin actually delivers, which is disciplined by the complaint cross-section rather than by the incidence shares. Organic diffusion alone recovers about \$44.2 million of restitution in the first projection year and \$887.6 million cumulatively over ten years at $\lambda = 0.5$ (bottom of panel A, Table 9)—the latter roughly five times the relief the entire complaint system paid out over 2015–2024. Ten years of organic diffusion nonetheless reduces the stock of unaware households' foregone relief by only a few percentage points depending on λ, making awareness a policy margin rather than a problem that resolves itself (Appendix Table VII.c). The distribution of this margin runs opposite to the filer-side losses: the unaware population is concentrated among lower-income households, and so is the unclaimed relief. Putting the two margins together completes the distributional picture. Among the consumers who file, the system favors the engaged and better-off; the consumers who never enter the process at all are disproportionately lower-income. Figure VII.1 plots the awareness paths and Tables VII.c and VII.d report the resulting restitution.

E Counterfactual Policy Experiments

The two models support two families of counterfactuals: experiments that perturb the firm's policy function, holding the filing population fixed, and experiments that broaden awareness, holding the policy function fixed. Table 9 reports both, each expressed as a percentage change against its own model baseline, at two horizons—the full 10 years behind the two models (2015–2024 for the

structural model, and 2024-2033, 120-month path, for the awareness diffusion model) in the first two columns, and the one overlap year (2024) in the last two columns. The resolution experiments are numbered CF1–CF5 and the awareness experiments CF6–CF8. Awareness experiments are reported at $\lambda = 0.50$.

The horizon is not a neutral rescaling, because the two channels are different objects in time. The resolution counterfactuals are static: each perturbs a primitive of the policy function and re-evaluates σ on complaints actually filed, so widening the window multiplies dollars in proportion to complaint volume while leaving percentage effects almost unchanged. The awareness counterfactuals are dynamic: each modifies a diffusion path, and the difference between two paths compounds, so doubling the diffusion rate is worth about $+7\%$ over twelve months and $+146\%$ over 120. A single-year snapshot therefore understates the awareness channel. A second asymmetry runs the other way. The resolution effects are evaluated on complaints that exist in the data, while the awareness effects inherit whatever error the diffusion model carries in β_θ, in the problem-household stock, and in the assumption that the estimated $\widehat{\sigma}_{p,\theta}$ applies to households who have never filed.

Firm-side experiments. The first panel of Table 9 reports five perturbations of the resolution index. Making every complaint self-advocated (CF1: $q = 1$ for all) raises expected relief by 7.67%—self-advocacy helps, but universal engagement is a second-order lever. Strengthening the monitoring report card complementarity by half (CF2: $\delta_{q\omega}, \delta_{\theta\omega} \times 1.5$) adds 1.17%, and we report it as an illustration of the channel rather than as a signed policy effect: it scales the two least sharply identified parameters in the index. Being an interaction rather than a level, it also redistributes resolution toward already-engaged complainants without materially raising the aggregate. Holding every firm to the large-firm standard (CF3: $\text{FS}_i \to 1$) raises expected relief by 11.77%; only the 37% of complaints filed against small firms move, since large firms are untouched by construction. [25]

By far the largest firm-side lever is the product baseline itself: raising baseline resolution odds by half (CF4: $\alpha_p + \ln 1.5$), the structural analog of reducing product complexity or raising the enforcement cost of denial, adds 39.07% to expected relief—roughly \$69 million over the sample period.[26] Combining it with the previous three levers (CF5) adds 67.91%. The ranking is the model's sharpest firm-side statement: the binding constraint on filer redress is what firms can

[25] Because FS is a size proxy and γ a between-firm contrast, counterfactual CF3 cannot separate the supervisory intensity that comes with scale from everything else that scales with it—compliance staffing, legal expertise, standardized remediation—so it delivers its full effect only under the supervisory reading and is otherwise an upper bound; and because the merit weight values each denial at the firm's own resolution rate, which is low for small firms precisely because enforcement there is weak, its foregone-welfare gain is measured conservatively relative to its effect on relief paid.

[26] The lever is an additive log-odds shift rather than a probability multiplier because α_p is negative whenever baseline resolution is below one half, which it always is here; a uniform shift $\alpha_p + \ln k$ multiplies every product's resolution *odds* by exactly k and preserves cross-product odds ratios. Because resolution is rare, $k = 1.5$ is almost exactly a $+50\%$ shift in the resolution probability.

deny with impunity on complex products—not the engagement of individual consumers and not the sharpness of monitoring, both of which help only at the margin.

It is worth reading the foregone-welfare column alongside the restitution column, because the two behave very differently. Even CF5, which raises relief paid by more than two thirds, reduces foregone filer welfare by only 10.92%. This is not an inconsistency but the arithmetic of a low-resolution system: a policy that raises σ moves complaints out of the denied set, but the complaints that remain denied are still valued at their own merit weight, and the vast majority of complaints remain denied under every scenario considered. The filer margin is simply not where a large fraction of the loss can be recovered.

Awareness experiments. On the participation margin, the levers are substantially larger in magnitude, and unlike the firm-side levers, they are progressive in incidence. A one-time information campaign that lifts below-median income awareness to the above-median income seed (CF6) raises cumulative ten-year restitution by 29.70%—from $887.6 million to $1.15 billion—and does so with by-construction progressive incidence, since the intervention moves only the below-median seed (the below-median income share of cumulative restitution rises from 48% to 60.1%). Sustained outreach that doubles the diffusion rate (CF7) is worth +135.89% over a decade, though it is nearly distributionally neutral because it scales both income tiers. Combining these two levers (CF8) reaches +203.21%—roughly $2.7 billion of cumulative restitution against $0.9 billion under organic diffusion—with 60.8% of the total flowing to below-median income filers. The same policies also make visible inroads on the stock of unaware households: CF8 cuts terminal non-filer foregone welfare by 13.77%.

The one-year panel is the like-for-like comparison, albeit being a lower bound for awareness.[27] Over twelve months CF6 delivers +29.74% of restitution and CF8 +38.49%, against +38.82% and +68.84% for CF4 and CF5—the channels are comparable in magnitude within a single year. It is the compounding over the decade, not any within-year advantage, that makes the awareness channel dominate at the longer horizon.

Taken together, the counterfactuals point to two distinct, complementary interventions. Among filers, the gains come from lowering the baseline denial rate on complex products and from extending large-firm enforcement exposure to small firms—the enforcement margin—rather than from exhorting consumers to advocate or from sharpening monitoring alone. But the largest welfare gains come from the participation channel: broadening awareness of the complaint process, especially among lower-income households, addresses the largest component of the welfare loss and is the only lever considered here whose incidence is progressive. A consumer-protection policy that sharpens firm oversight while leaving the awareness gap and the complexity of redress untouched would reach neither margin where it matters most.

[27] 2024 represents only the first 12 months of the simulated 10-year diffusion process.

VI Conclusion

When a firm makes a mistake in a complex contract, correcting it requires the harmed consumer to file a claim and the firm to agree to pay. Using confidential regulatory data on nearly two million complaints filed with the CFPB between 2014 and 2024, we ask who bears the cost when firms err. The answer depends far less on which firm made the mistake than on who is able to operate the machinery of redress.

We present several findings about the correction of firm errors. First, firm identity barely predicts whether consumers receive restitution; the complexity of the contract in which the error occurred predicts restitution. Second, among claimants, redress rises with the vigor of self-advocacy, while firms show no evidence of favoritism in the resources they devote to a complaint. Third, confidential monitoring reports raise the share of complaints firms pay, showing the effects of regulatory enforcement, and that increase accrues to the claimants already best equipped to press their case. Enforcement amplifies the returns to sophistication rather than leveling the playing field.

A structural model of the firm's grant-or-deny decision, extended with a diffusion process for how consumers learn that the complaint channel exists, converts these patterns into a measure of who bears the cost of firm error. Among filers, foregone relief is modest. Losses are largest among consumers who likely experience an error and never claim, which falls disproportionately on lower-income households. Increasing awareness of the claims channel recovers more than improving consumers' self-advocacy or increasing regulatory scrutiny, with the benefits .

The dollars at stake are large for the typical household. The average payment in our sample reached $2,295 by 2024, more than five times the $400 unexpected expense that 37% of U.S. adults report they could not cover with cash or its equivalent.[28] Uncorrected errors are a transfer from those households to firms, and the transfer is regressive: the consumers least able to absorb the loss are the ones least likely to contest it. Enforcement changes how firms resolve the claims they receive, but it does nothing about the claims that are never made. Our findings show that markets in which consumers must individually contest a firm's errors will allocate redress by the ability to contest them.

[28] https://www.federalreserve.gov/newsevents/pressreleases/files/other20260513a1.pdf

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Yan Zhang, "Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide," *American Economic Review*, 2014, *104* (9), 2830–2857.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, "Regulating consumer financial products: Evidence from credit cards," *The Quarterly Journal of Economics*, 2015, *130* (1), 111–164.

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

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A. Olken, and Rema Hanna, "Self-Targeting: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 2016, *124* (2), 371–427.

Argyle, Bronson, Sasha Indarte, Benjamin Iverson, and Christopher Palmer, "Racial Disparities and Bias in Consumer Bankruptcy," Working Paper 33575, National Bureau of Economic Research 2025.

Arrow, Kenneth J., "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, 1973, pp. 3–33.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169–217.

_ , *The Economics of Discrimination*, 2 ed., Chicago, IL: University of Chicago Press, 1971.

Begley, Taylor A and Amiyatosh Purnanandam, "Color and credit: Race, regulation, and the quality of financial services," *Journal of Financial Economics*, 2021, *141* (1), 48–65.

Bhargava, Saurabh and Dayanand Manoli, "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review*, 2015, *105* (11), 3489–3529.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, "Paying Too Much? Borrower Sophistication and Overpayment in the US Mortgage Market," *Journal of Finance*, 2026, *81* (1), 49–90.

Bian, Bo, Michaela Pagel, Huan Tang, and Devesh Raval, "Consumer Surveillance and Financial Fraud," *Working Paper*, 2023.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

Board, Simon and Moritz Meyer ter Vehn, "Reputation for Quality," *Econometrica*, 2013, *81* (6), 2381–2462.

Butler, Alexander W, Erik J Mayer, and James P Weston, "Racial disparities in the auto loan market," *The Review of Financial Studies*, 2023, *36* (1), 1–41.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Carlin, Bruce I, "Strategic price complexity in retail financial markets," *Journal of Financial Economics*, 2009, *91* (3), 278–287.

Cheng, Ing-Haw, Felipe Severino, and Richard R Townsend, "How do consumers fare when dealing with debt collectors? Evidence from out-of-court settlements," *The Review of Financial Studies*, 2021, *34* (4), 1617–1660.

Currie, Janet, "The Take-Up of Social Benefits," in Alan J. Auerbach, David Card, and John M. Quigley, eds., *Public Policy and the Distribution of Income*, New York: Russell Sage Foundation, 2006, pp. 80–148.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in U.S. Retail Chains," *Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Deshpande, Manasi and Yue Li, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–248.

Dimmock, Stephen G, William C Gerken, and Nathaniel P Graham, "Is fraud contagious? Coworker influence on misconduct by financial advisors," *The Journal of Finance*, 2018, *73* (3), 1417–1450.

Dou, Yiwei and Yongoh Roh, "Public disclosure and consumer financial protection," *Journal of Financial and Quantitative Analysis*, 2024, *59* (5), 2164–2198.

Dou, Yiwei, Mingyi Hung, Guoman She, and Lynn Linghuan Wang, "Learning from peers: Evidence from disclosure of consumer complaints," *Journal of Accounting and Economics*, 2024, *77* (2-3), 101620.

Dunn, Abe, Joshua D. Gottlieb, Adam Hale Shapiro, Daniel J. Sonnenstuhl, and Pietro Tebaldi, "A Denial a Day Keeps the Doctor Away," *Quarterly Journal of Economics*, 2024, *139* (1), 187–233.

Egan, Mark, Gregor Matvos, and Amit Seru, "The market for financial adviser misconduct," *Journal of Political Economy*, 2019, *127* (1), 233–295.

_ , _ , and _ , "Arbitration with Uninformed Consumers," *Review of Economic Studies*, 2025, *92* (6), 3888–3923.

Eisenberg, Theodore and Geoffrey P. Miller, "An Empirical Look at Compensation in Consumer Class Actions," *NYU Journal of Law & Business*, 2015, *11* (4), 767–802.

Finkelstein, Amy and Matthew J. Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP," *Quarterly Journal of Economics*, 2019, *134* (3), 1505–1556.

Flanagan, Thomas and Amiyatosh Purnanandam, "Did banks pay fair returns to taxpayers on TARP?," *The Journal of Finance*, 2024, *79* (5), 2909–2941.

Fuster, Andreas, Matthew C. Plosser, and James I. Vickery, "Does CFPB Oversight Crimp Credit?," Staff Report 857, Federal Reserve Bank of New York 2023. Originally issued June 2018; revised July 2023.

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

Gathergood, John, Neale Mahoney, Neil Stewart, and Jörg Weber, "How do individuals repay their debt? The balance-matching heuristic," *American Economic Review*, 2019, *109* (3), 844–875.

Glaeser, Edward L. and Andrei Shleifer, "The Rise of the Regulatory State," *Journal of Economic Literature*, 2003, *41* (2), 401–425.

Goldfarb, Avi and Mo Xiao, "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets," *American Economic Review*, 2011, *101* (7), 3130–3161.

Green, Jerry R., *Information, Efficiency, and Equilibrium* Harvard Economic Studies, Harvard University Press, 1973.

Grossman, Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 1986, *94* (4), 691–719.

Gurun, Umit G, Noah Stoffman, and Scott E Yonker, "Trust busting: The effect of fraud on investor behavior," *The Review of Financial Studies*, 2018, *31* (4), 1341–1376.

Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 1990, *98* (6), 1119–1158.

Hayes, Rachel M, Feng Jiang, and Yihui Pan, "Voice of the customers: Local trust culture and consumer complaints to the CFPB," *Journal of Accounting Research*, 2021, *59* (3), 1077–1121.

Heidhues, Paul and Botond Kőszegi, "Naivete-based discrimination," *The Quarterly Journal of Economics*, 2017, *132* (2), 1019–1054.

Herd, Pamela and Donald P. Moynihan, *Administrative Burden: Policymaking by Other Means*, New York: Russell Sage Foundation, 2018.

Hirschman, Albert O., *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Cambridge, MA: Harvard University Press, 1970.

Homonoff, Tatiana and Jason Somerville, "Program Recertification Costs: Evidence from SNAP," *American Economic Journal: Economic Policy*, 2021, *13* (4), 271–298.

Hortaçsu, Ali, Fernando Luco, Steven L. Puller, and Dongni Zhu, "Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets," *American Economic Review*, 2019, *109* (12), 4302–4342.

Huang, Ruidi, James S. Linck, Erik J. Mayer, and Christopher A. Parsons, "Can Human Capital Explain Income-Based Disparities in Financial Services?," *Review of Financial Studies*, 2026, *39* (6), 1785–1822.

Hylton, Keith N., "The Economics of Class Actions and Class Action Waivers," *Supreme Court Economic Review*, 2015, *23*, 305–339.

Jaffee, Dwight M. and Thomas Russell, "Imperfect Information, Uncertainty, and Credit Rationing," *Quarterly Journal of Economics*, 1976, *90* (4), 651–666.

Jappelli, Tullio and Marco Pagano, "Information Sharing in Credit Markets," *Journal of Finance*, 1993, *48* (5), 1693–1718.

Jørring, Adam Tejs, "Financial sophistication and consumer spending," *The Journal of Finance*, 2024, *79* (6), 3773–3820.

Jou, Jeffrey, Anya Kleymenova, Andrea Passalacqua, László Sándor, and Rajesh Vijayaraghavan, "Disciplining Banks through Disclosure: Evidence from CFPB Consumer Complaints," *Available at SSRN*, 2024.

Kaplow, Louis and Steven Shavell, "Accuracy in the Determination of Liability," *Journal of Law and Economics*, 1994, *37* (1), 1–15.

Keys, Benjamin J, Devin G Pope, and Jaren C Pope, "Failure to refinance," *Journal of Financial Economics*, 2016, *122* (3), 482–499.

Kim, You Suk, Donghoon Lee, Tess Scharlemann, and James Vickery, "Intermediation frictions in debt relief: evidence from cares act forbearance," *Journal of Financial Economics*, 2024, *158*, 103873.

Klein, Benjamin and Keith B. Leffler, "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 1981, *89* (4), 615–641.

Kuhnen, Camelia M and Brian T Melzer, "Noncognitive abilities and financial delinquency: The role of self-efficacy in avoiding financial distress," *The Journal of Finance*, 2018, *73* (6), 2837–2869.

LaVoice, Jessica and Domonkos F Vamossy, "Racial disparities in debt collection," *Journal of Banking & Finance*, 2024, *164*, 107208.

Lee, Brian Jonghwan, "Bankruptcy lawyers and credit recovery," *Working Paper*, 2024.

Li, Xiang, "Does the disclosure of consumer complaints reduce racial disparities in the mortgage lending market," *Working Paper*, 2023.

Lusardi, Annamaria and Olivia S Mitchell, "The economic importance of financial literacy: Theory and evidence," *Journal of Economic Literature*, 2014, *52* (1), 5–44.

Mazur, Laurel, "The impact of the CFPB complaint database disclosure on banksâ€™ provision of mortgage credit to low-and moderate-income communities," *Working Paper*, 2024.

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *American Economic Review*, 1982, *72* (2), 372–377.

Parise, Gianpaolo and Kim Peijnenburg, "Noncognitive abilities and financial distress: evidence from a representative household panel," *The Review of Financial Studies*, 2019, *32* (10), 3884–3919.

Polinsky, A. Mitchell and Steven Shavell, "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, 2000, *38* (1), 45–76.

Priest, George L. and Benjamin Klein, "The Selection of Disputes for Litigation," *Journal of Legal Studies*, 1984, *13* (1), 1–55.

Schenone, Carola, "Lending relationships and information rents: Do banks exploit their information advantages?," *The Review of Financial Studies*, 2010, *23* (3), 1149–1199.

Schwemm, Robert G and Jeffrey L Taren, "Discretionary pricing, mortgage discrimination, and the Fair Housing Act," *Harv. CR-CLL Rev.*, 2010, *45*, 375.

Sedunov, John, "Small banks and consumer satisfaction," *Journal of Corporate Finance*, 2020, *60*, 101517.

Shapiro, Carl, "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics*, 1983, *98* (4), 659–679.

Shavell, Steven, "The Fundamental Divergence between the Private and the Social Motive to Use the Legal System," *Journal of Legal Studies*, 1997, *26* (S2), 575–612.

Spier, Kathryn E., "Litigation," in A. Mitchell Polinsky and Steven Shavell, eds., *Handbook of Law and Economics*, Vol. 1, Amsterdam: Elsevier, 2007, pp. 259–342.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, *71* (3), 393–410.

Sunstein, Cass R., "Sludge and Ordeals," *Duke Law Journal*, 2019, *68* (8), 1843–1883.

Tirole, Jean, "Cognition and Incomplete Contracts," *American Economic Review*, 2009, *99* (1), 265–294.

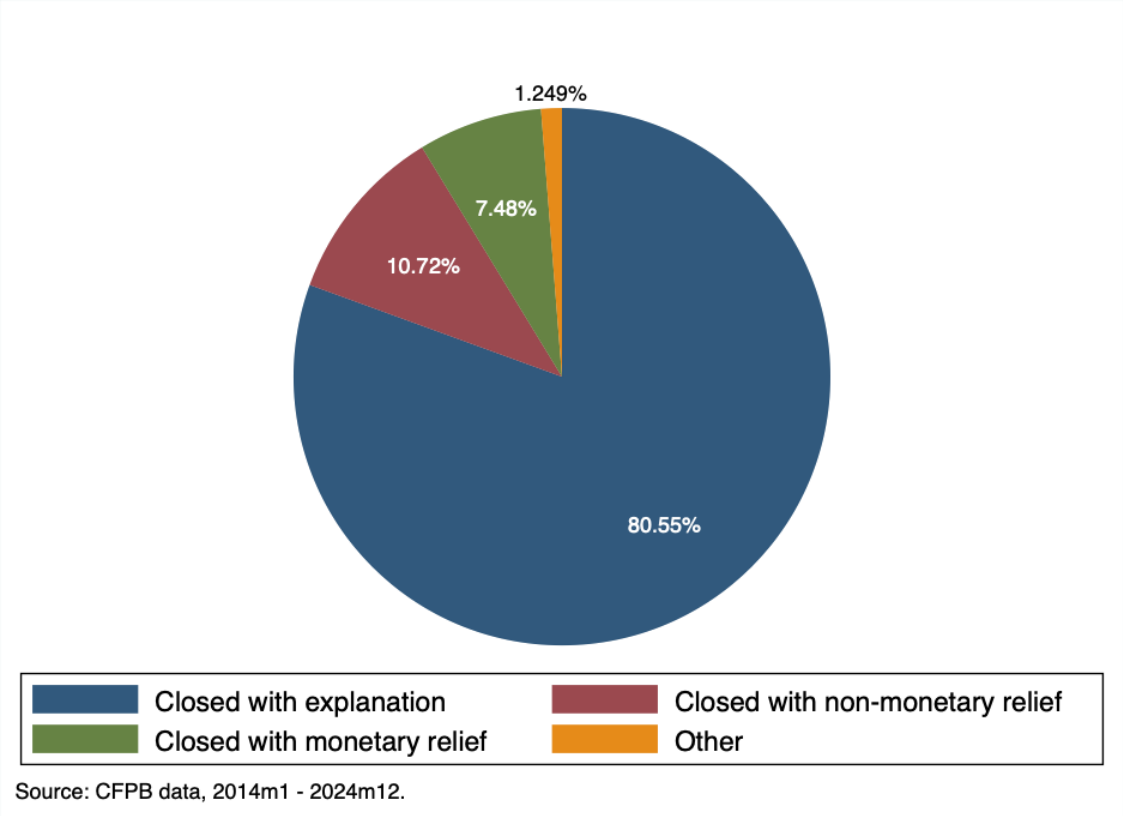
Williamson, Oliver E., *Markets and Hierarchies: Analysis and Antitrust Implications: A Study in the Economics of Internal Organization*, New York: Free Press, 1975.

Woodward, Susan E and Robert E Hall, "Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence," *American Economic Review*, 2012, *102* (7), 3249–76.

Zwick, Eric, "The Costs of Corporate Tax Complexity," *American Economic Journal: Economic Policy*, 2021, *13* (2), 467–500.

Figure 1. Company Answers to Complaints

The figure illustrates the ways in which companies answer CFPB consumer complaints. The vast majority of complaints is closed with an explanation (81.6%), followed by 10.8% closed with an action that does not entail monetary relief, and 7.6% of complaints are solved with monetary relief.

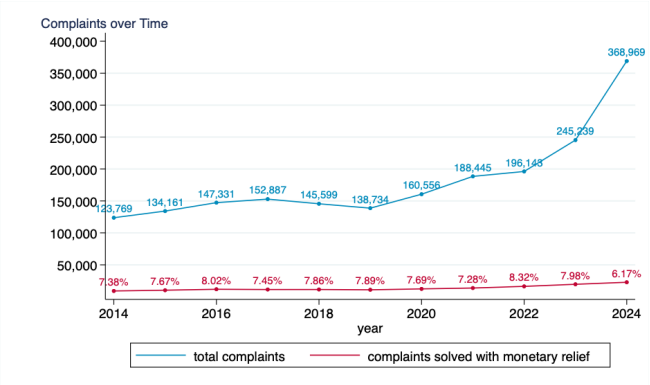


Source: CFPB data, 2014m1 - 2024m12.

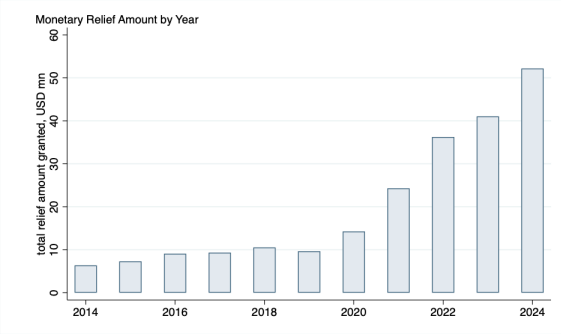
Figure 2. Complaints Dynamics

The figure investigates trends in CFPB complaints over time. The graph on the left-hand side of Panel A plots the overall number of complaints filed with the CFPB (blue line, with corresponding labels) and the overall number of complaints that received monetary relief (red line, with labels in terms of percentage of overall complaints) over time. The graph on the right-hand side of Panel A plots the total amount of monetary relief granted each year in millions. Panel B reports the total number of complaints within each product over time—distinguishing across debt collection, cards (debit, credit, prepaid), accounts (checking or savings), mortgage, money transfer (money transfer, virtual currency, or money service), student loan, car loan, other loan (payday, title, personal loan), and other (residual category).

Panel A: Complaints and Monetary Relief over time



Complaints over time



Relief Amount over time

Panel B: Complaints by product over time

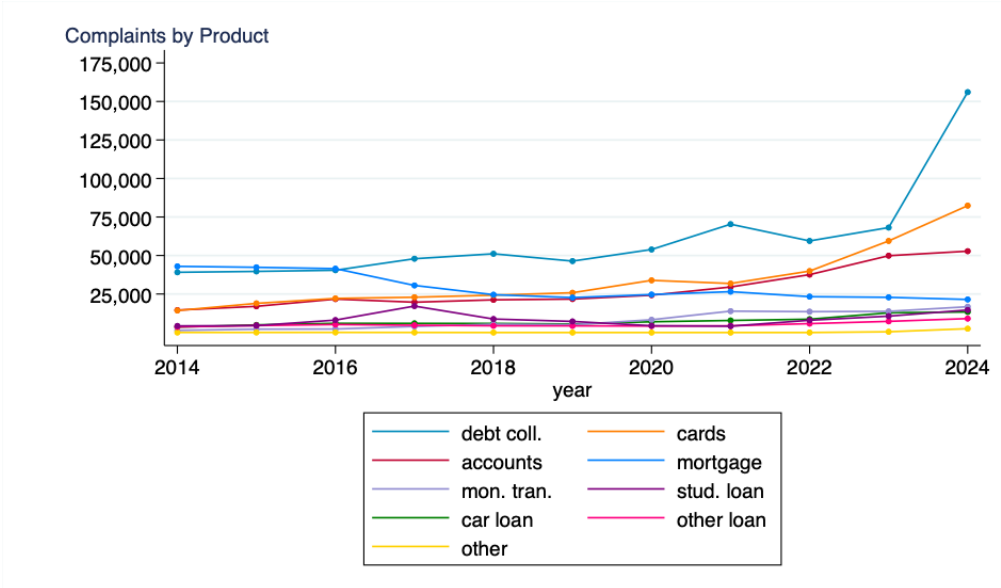
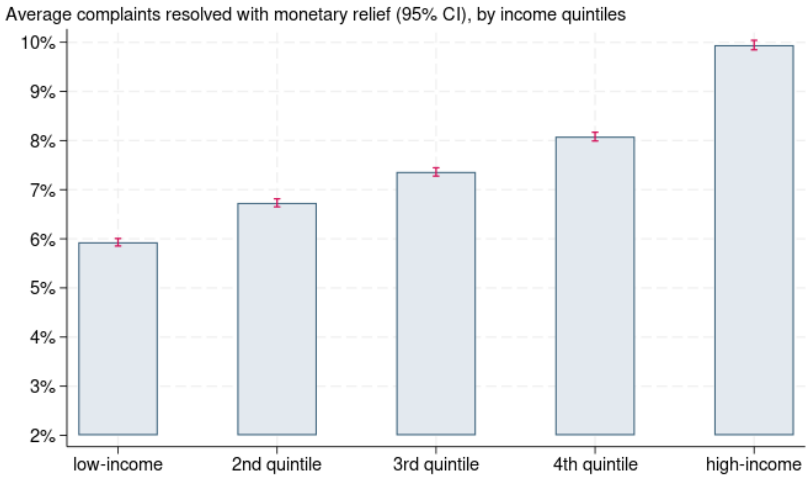


Figure 3. Financial Restitution Across Income Groups

The figure documents the relationship between financial restitution likelihood and income. Panel A reports the percentage of complaints solved with monetary relief by quintiles of the filer's Census tract household median income. Panel B reports for each product the difference between the share of complaints solved with monetary relief in the highest quintile of filer's tract household median income and the share of complaints solved with monetary relief in the lowest quintile of filer's tract household median income.

Panel A: Monetary relief likelihood on Census tract income quintiles



Panel B: Monetary relief differential (highest - lowest) across products:

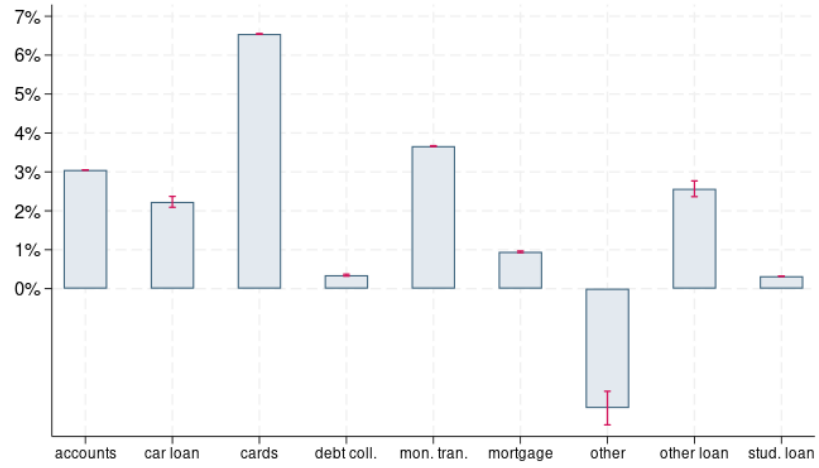
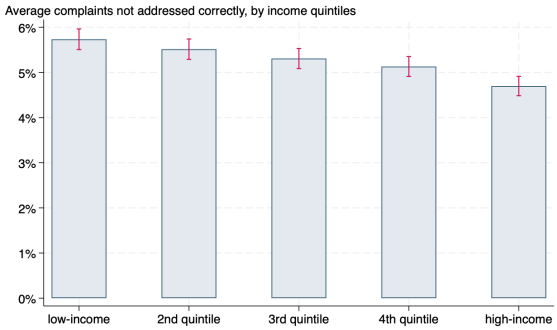
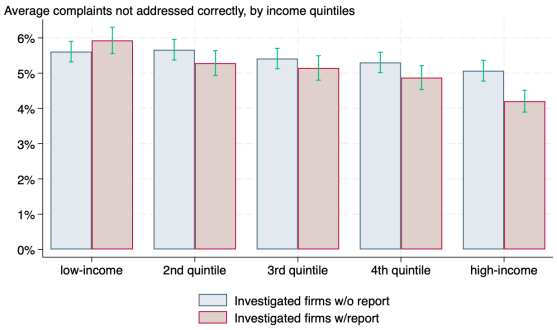


Figure 4. CFPB Company Monitoring Reviews by filer's census tract income

The figure reports the percentage of complaints flagged by the CFPB as having a problematic response (amongst the ones screened as part of the company monitoring reviews) by filer's Census tract household median income quintiles. Panel A reports the share of reviewed complaints flagged as not properly addressed by the company. Panel B reports the share of reviewed complaints flagged as lacking a substantive response. Across panels, plots on the left-hand side report results for all reviewed complaints, plots on the right-hand side distinguish between complaints at reviewed firms that never received a monitoring report (in blue) and complaints at reviewed firms that received a monitoring report at least once (negative feedback from CFPB, in red).

Panel A: Company did not address complaint





Panel B: Company response not substantive

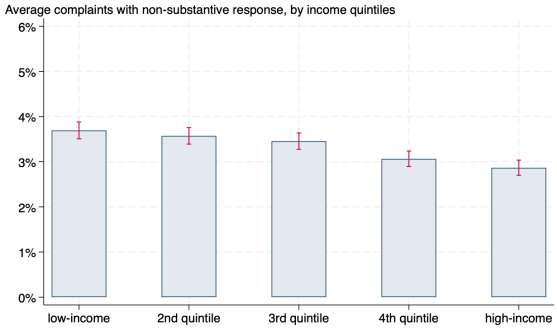
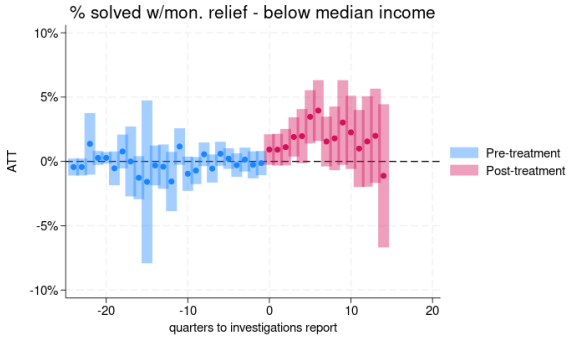




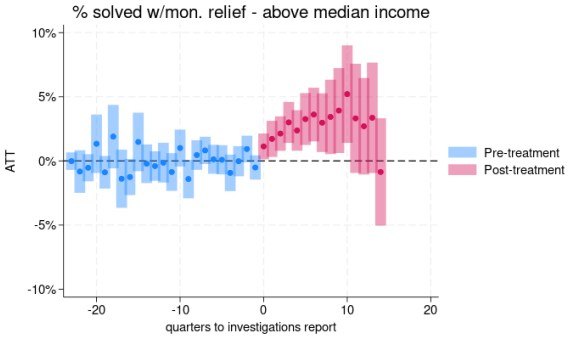
Figure 5. CFPB Monitoring Report Cards - Dynamic Difference in Differences

The figure reports event-study estimates of the effect of the company receiving a monitoring report card (negative feedback from the CFPB) on the likelihood of granting financial restitution, using the Callaway and Sant'Anna (2021) difference-in-differences estimator. Across panels, the left-hand side graph is estimated on complaints coming from below median household median income Census tracts, while the right-hand side graph is estimated on complaints coming from above median household median income Census tracts. In Panel A, the treated companies are companies that were reviewed and received a monitoring report card (negative feedback from CFPB, treatment), and the controls are companies that were reviewed but did not receive a report card and had at least 100 complaints in 2018, prior to the start of the company monitoring reviews. In Panel B, the controls are companies that were reviewed and did not receive a report card, and were the nearest neighbors of the treated firms in terms of overall number of complaints, share of loan-related complaints, and share of complaints with filer attachments prior the start of the company review period (matched controls).

Panel A:CFPB reports by low and high income

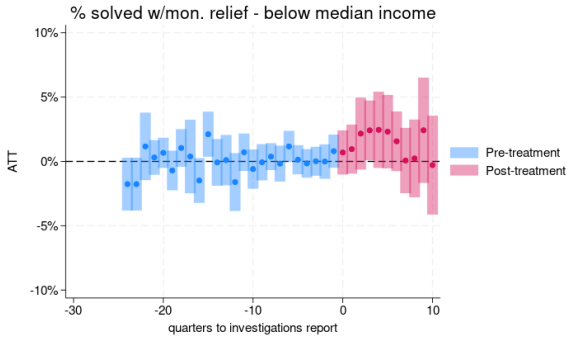


Low-income complaints

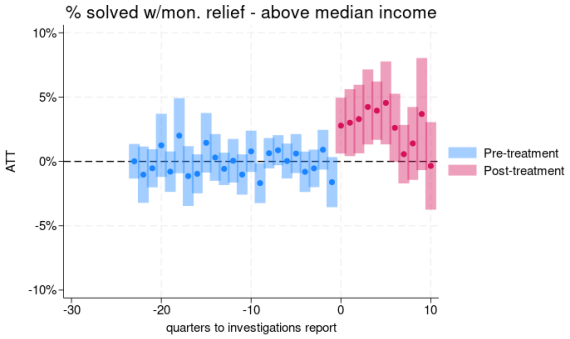


High-income complaints

Panel B: CFPB reports by low and high income, matched sample



Low-income complaints



High-income complaints

Table 1. Monetary Relief Dynamics

Description: The table shows financial restitution trends over time and across products. Panel A reports the evolution over time of the following metrics: the number of total complaints, the number of complaints solved with monetary relief (in absolute terms and in percentage), the average relief amount across complaints solved with financial restitution and its standard deviation. Panel B reports the same metrics by product type instead of year. Relief amounts are winsorized at the 2.5th and 97.5th percentiles.

Panel A: Monetary relief likelihood and amount over time

year	complaints	solved wMR	% solved wMR	relief mean	relief st. dev.
2014	123,769	9,134	7.38%	697.10	755.54
2015	134,161	10,294	7.67%	710.69	840.53
2016	147,331	11,817	8.02%	769.42	856.36
2017	152,887	11,391	7.45%	819.84	910.29
2018	145,599	11,442	7.86%	921.08	933.94
2019	138,734	10,945	7.89%	882.45	817.60
2020	160,556	12,346	7.69%	1,154.93	1,091.27
2021	188,445	13,725	7.28%	1,769.49	1,933.86
2022	196,143	16,317	8.32%	2,221.50	2,204.92
2023	245,239	19,568	7.98%	2,097.89	1,901.45
2024	368,969	22,749	6.17%	2,294.54	2,049.21
total	2,001,833	149,728	7.48%	1,472.07	1,682.25

Panel B: Monetary relief likelihood and amount across products

product	complaints	solved wMR	% solved wMR	relief mean	relief st. dev.
accounts	309,960	54,303	17.52%	1,907.41	1,626.78
car loan	81,392	3,101	3.81%	1,430.17	2,732.84
cards	375,756	64,828	17.25%	993.35	681.41
debt coll.	672,548	4,520	0.67%	1,238.36	2,094.72
mon. tran.	86,542	7,955	9.19%	2,264.92	3,042.08
mortgage	323,391	9,903	3.06%	1,542.52	1,841.46
other	3,155	145	4.60%	1,330.54	1,426.76
other loan	57,339	3,073	5.36%	1,202.11	2,275.84
stud. loan	91,750	1,900	2.07%	2,778.17	5,146.88

Table 2. Determinants of Restitution

Description: The table analyzes the determinants of the likelihood of receiving financial restitution. It reports variance decompositions of regressing *solved wMR*—a dummy variable equal to 1 if the complaint is resolved with monetary relief—on different combinations of quarter fixed effects, company fixed effects, product fixed effects, and filer's county fixed effects. Percentages are based on Mean Squares, which account for the degrees of freedom of the explanatory variables. The first three columns report results from regressions run on the entire sample; Columns 4, 5, and 6 report results from regressions run on the subsample of companies that provide more than one financial product.

	all companies			**companies > 1 product**		
dep. variable	**% model variation explained (MS)**			**% model variation explained (MS)**		
= solved wMR	(1)	(2)	(3)	(4)	(5)	(6)
Quarter FE	33.33%	0.5%	0.5%	19.42%	0.50%	0.48%
Company FE	66.67%	0.24%	0.24%	80.58%	0.52%	0.51%
product FE		99.26%	99.23%		98.98%	98.98%
County FE			0.03%			0.30%
model variation (R-squared)	10.29%	12.18%	12.56%	10.03%	11.94%	12.32%
obs	1,996,702	1,996,702	1,900,172	1,944,294	1,944,294	1,850,105

Table 3. Income and Financial Restitution

Description: The table reports estimates of the relationship between filer income and financial restitution. In Panel A the outcome is an indicator equal to one if the complaint was resolved with monetary relief, estimated by OLS. In Panel B the outcome is the dollar relief awarded, winsorized at the 2.5% level in each tail and zero otherwise, estimated by fixed-effects Poisson. Across both panels, Columns 1 to 3 measure filer income by the above-median tract indicator and Columns 4 to 6 by log tract income, under progressively saturated fixed effects. Variables are defined in Table II.a. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

Panel A: Financial Restitution Likelihood

	solved w/MR$_{i,c,q,r}$					
	(1)	(2)	(3)	(4)	(5)	(6)
High-income tract	0.0101***	0.00951***	0.00868***			
	(0.0019)	(0.0016)	(0.0014)			
ln(tract income)				0.0139***	0.0131***	0.0131***
				(0.0025)	(0.0022)	(0.0022)
Quarter FE	x	x	x	x	x	x
Company FE	x			x		
Company × product FE		x	x		x	x
County FE			x			x
Observations	1,828,730	1,826,276	1,826,203	1,828,730	1,826,276	1,826,203
R^2	0.104	0.137	0.139	0.104	0.137	0.139

Panel B: Financial Restitution Amount

	monetary relief amount$_{i,c,q,r}$					
	(1)	(2)	(3)	(4)	(5)	(6)
High-income tract	0.299***	0.295***	0.271***			
	(0.028)	(0.027)	(0.031)			
ln(tract income)				0.396***	0.384***	0.388***
				(0.029)	(0.028)	(0.033)
Quarter FE	x	x	x	x	x	x
Company FE	x			x		
Company × product FE		x	x		x	x
County FE			x			x
Observations	1,517,592	1,439,004	1,433,561	1,517,592	1,439,004	1,433,561

Table 4. Filer and Company Activity

Description: The table reports OLS estimates of filer- and company-side activity on filer income. Panel A asks whether self-advocating or attaching a supporting document varies with income: the outcome is self-advocacy in the complaint narrative in Columns 1 and 2, self-advocacy in the desired-resolution field in Columns 3 and 4, and attachment in Columns 5 and 6. Panel B asks whether companies assign different personnel to complaints from different tracts: the outcome is employee experience in Columns 1 and 2 and employee login time in Columns 3 and 4. Within each panel, odd-numbered columns measure filer income by the above-median tract indicator and even-numbered columns by log tract income. Variables are defined in Table II.a, and self-advocacy terms are listed in Table II.b. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

Panel A: Filer activity

	I(filer self-advocacy)$_{i,c,q,r}$				I(filer attachment)$_{i,c,q,r}$	
	in description		in desired resolution			
	(1)	(2)	(3)	(4)	(5)	(6)
High-income tract	0.0233***		0.0134***		0.00373***	
	(0.0015)		(0.0016)		(0.0013)	
ln(tract income)		0.0341***		0.0210***		0.00694***
		(0.0021)		(0.0021)		(0.0014)
Quarter FE	x	x	x	x	x	x
Company × product FE	x	x	x	x	x	x
County FE	x	x	x	x	x	x
Observations	1,741,564	1,741,564	1,805,330	1,805,330	1,826,203	1,826,203
R^2	0.0304	0.0307	0.0242	0.0243	0.281	0.281

Panel B: Company activity

	employee experience$_{i,c,q,r}$		employee login time$_{i,c,q,r}$	
	(1)	(2)	(3)	(4)
High-income tract	0.000795		0.0000803	
	(0.0028)		(0.00034)	
ln(tract income)		0.00177		0.000492
		(0.0038)		(0.00044)
Quarter FE	x	x	x	x
Company × product FE	x	x	x	x
County FE	x	x	x	x
Observations	1,427,040	1,427,040	1,330,225	1,330,225
R^2	0.588	0.588	0.437	0.437

Table 5. Financial Restitution Mechanisms

Description: The table reports OLS estimates of how much filer- and company-side activity accounts for the income gradient in financial restitution. Columns 1 and 2 add the filer activity variables to Equation 1; Columns 3 and 4 add the company activity variables; Columns 5 and 6 include both. Even-numbered columns interact each activity variable with filer income. Filer income is measured by the above-median tract indicator; Table III.b repeats the specification with log tract income. Variables are defined in Table II.a. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	Resolved with monetary relief					
	Filer activity		Company activity		Both	
	(1)	(2)	(3)	(4)	(5)	(6)
High-income tract	0.00834***	0.00348***	0.00931***	0.00747*	0.00891***	0.00195
	(0.0014)	(0.00099)	(0.0017)	(0.0039)	(0.0016)	(0.0036)
Self-advocacy	0.0219***	0.0176***			0.0206***	0.0167***
	(0.0031)	(0.0025)			(0.0031)	(0.0027)
Attachment	0.0123***	0.00974***			0.0120***	0.00949***
	(0.0021)	(0.0019)			(0.0019)	(0.0016)
Self-advocacy × High-income tract		0.00841***				0.00777***
		(0.0016)				(0.0016)
Attachment × High-income tract		0.00514***				0.00496***
		(0.0016)				(0.0012)
Employee experience			0.000237	-0.000587	0.000265	-0.000544
			(0.00095)	(0.00093)	(0.00093)	(0.00091)
Employee login time			-0.0120***	-0.00489*	-0.0113***	-0.00420
			(0.0032)	(0.0026)	(0.0031)	(0.0026)
Employee experience × High-income tract				0.00168***		0.00164***
				(0.00058)		(0.00058)
Employee login time × High-income tract				-0.0145***		-0.0143***
				(0.0043)		(0.0043)
Quarter FE	x	x	x	x	x	x
Company × product FE	x	x	x	x	x	x
County FE	x	x	x	x	x	x
Respondent FE			x	x	x	x
Observations	1,820,166	1,820,166	1,328,922	1,328,922	1,311,748	1,311,748
R^2	0.142	0.142	0.168	0.168	0.170	0.170

Table 6. CFPB Monitoring Report Cards

Description: The table presents difference-in-differences estimates around the first time a company receives a confidential monitoring report card from the CFPB. The sample comprises all treated companies and, as controls, companies that were reviewed but did not receive a report and have at least 100 complaints in 2018. The outcome is an indicator equal to one if the complaint was resolved with monetary relief. Columns 1 to 3 report the baseline specification under progressively saturated fixed effects. Columns 4 and 5 interact the post-report indicator with filer self-advocacy, and Columns 6 and 7 with filer attachment; Columns 5 and 7 add the triple interaction with filer income. Variables are defined in Table II.a. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	Resolved with monetary relief						
	Baseline			Self-advocacy		Attachment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-report	-0.00586	-0.00847	-0.00817	-0.00651	-0.0138	-0.00626	-0.0128
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
High-income tract	0.00727***	0.00690***	0.00570***		0.00381***		0.00464***
	(0.0020)	(0.0017)	(0.0014)		(0.0013)		(0.0015)
Post-report × High-income tract	0.0203***	0.0178***	0.0174***		0.0144***		0.0128***
	(0.0040)	(0.0035)	(0.0035)		(0.0035)		(0.0033)
Self-advocacy				0.0178***	0.0154***		
				(0.0040)	(0.0036)		
Post-report × Self-advocacy				0.0203***	0.0168***		
				(0.0051)	(0.0052)		
High-income tract × Self-advocacy					0.00497***		
					(0.0016)		
Post-report × High-income tract × Self-advocacy					0.00456		
					(0.0032)		
Attachment						0.0117***	0.0107***
						(0.0026)	(0.0025)
Post-report × Attachment						0.0162***	0.0111***
						(0.0046)	(0.0040)
High-income tract × Attachment							0.00195
							(0.0016)
Post-report × High-income tract × Attachment							0.00866***
							(0.0032)
Quarter FE	x	x	x	x	x	x	x
Company FE	x						
Company × product FE		x	x	x	x	x	x
County FE			x	x	x	x	x
Observations	1,003,708	1,003,541	1,003,388	991,205	991,205	1,003,388	1,003,388
R^2	0.106	0.132	0.135	0.137	0.137	0.135	0.136

Table 7. Structural Estimates: The Firm's Resolution Policy Function

Description: The table reports maximum-likelihood estimates of the firm's complaint-resolution policy function

$$\sigma_i = \Lambda\big(\alpha_{p(i)} + \gamma \mathrm{FS}_i + (\varphi_q + \delta_{q\omega}\omega_i)\, q_i + (\varphi_\theta + \delta_{\theta\omega}\omega_i)\, \theta_i\big) \tag{13}$$

where $\Lambda(\cdot)$ is the logistic function, σ_i is the probability that complaint i is solved with monetary relief, q_i is a filer self-advocacy indicator, θ_i is an above-median income indicator, ω_i is the firm's CFPB monitoring report card state at the time of the complaint (equal to one if the complaint has been filed after the company has received a monitoring report card from the CFPB), and FS_i is a firm-size indicator (equal to one for companies spanning at least three products *and* clearing at least 250 complaints in every year they appear). All coefficients are in log-odds of σ_i. The α_p are product baselines (the log-odds of resolution for a reference complaint with $q = \theta = \omega = \mathrm{FS} = 0$); γ is the firm-size shift in that baseline; φ_q and φ_θ are the merit gradients in self-advocacy and income; and $\delta_{q\omega}$ and $\delta_{\theta\omega}$ are the monitoring report card-induced steepening of those gradients. Every parameter block is estimated freely by MLE. The estimation sample is the publicly available narrative subsample over the post-2020 CFPB monitoring report cards window, with debt-collection and monetary transfer complaints excluded ($N =$ 266,287 at 2,443 firms). The "implied monitoring DiD" rows report the difference-in-differences the fitted model generates *in resolution-probability space*, computed with the within-company transformation of Table 6 so that it is the same estimator on the same population; these are the objects comparable to the causal benchmarks reported beneath them, and are distinct from the log-odds δ coefficients above. The memo rows re-estimate $\alpha_p, \gamma, \varphi_q, \varphi_\theta$ on the $\omega = 0$ subsample alone ($N =$ 166,473), where no firm is currently under a report card and the δ terms play no role. Standard errors (in parentheses) are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	σ_i = Pr(solved w/MR)
	coefficients in log-odds
Product baselines α_p (complexity)	
stud. loan	-4.6496^{***}
	(0.3982)
car loan	-3.9546^{***}
	(0.2564)
mortgage	-3.8857^{***}
	(0.2125)
other loan	-3.3086^{***}
	(0.2071)
accounts	-2.3582^{***}
	(0.2441)
cards	-2.0471^{***}
	(0.3423)
Firm size (log-odds)	
γ (large firm, FS = 1)	0.5193
	(0.3321)
Merit gradients (log-odds)	
φ_q (self-advocacy)	0.2569^{***}
	(0.0720)
φ_θ (income)	0.0699
	(0.0615)
Monitoring interactions (log-odds)	
$\delta_{q\omega}$ (monitoring × advocacy)	0.0719
	(0.1143)
$\delta_{\theta\omega}$ (monitoring × income)	0.0351
	(0.0866)
Implied monitoring DiD (probability space)	
monitoring × advocacy	+0.0249
causal benchmark (Table 6)	+0.0203
monitoring × income	+0.0096
causal benchmark (Table 6)	+0.0174
log-likelihood	−89,967.9
Observations	266,287
clusters (firms)	2,443
Memo: $\omega = 0$ subsample ($N =$ 166,473, no firm monitored)	
γ	+0.5376
φ_q	+0.3105
φ_θ	+0.0982

Table 8. Implied Product Complexity

Description: The table reports product complexity implied by the estimated firm policy function (Table 7). α_p is the estimated product baseline: the log-odds of resolution for a reference complaint with $q = \theta = \omega = 0$ at a *small* firm (FS $= 0$). $\Lambda(\alpha_p)$ is the corresponding baseline resolution probability and $\Lambda(\alpha_p + \gamma)$ is the same object at a *large* firm (FS $= 1$: at least three products and at least 250 complaints in every year the company appears), so the two columns read the same product at the two enforcement regimes; $1 - \Lambda(\alpha_p)$ is the baseline denial rate—what the firm gets away with—and $\kappa_p = 1 - \Lambda(\alpha_p)/\max_{p'} \Lambda(\alpha_{p'})$ is complexity normalized so that the simplest product (cards) has $\kappa = 0$. Products are ordered from most to least complex. Baseline denial is uniformly higher at small firms, at every product.

	α_p (log-odds)	$\Lambda(\alpha_p)$ small firms	$\Lambda(\alpha_p + \gamma)$ large firms	$1 - \Lambda(\alpha_p)$ baseline denial	κ_p complexity
stud. loan	−4.6496	0.0095	0.0158	0.9905	0.917
car loan	−3.9546	0.0188	0.0312	0.9812	0.836
mortgage	−3.8857	0.0201	0.0334	0.9799	0.824
other loan	−3.3086	0.0353	0.0579	0.9647	0.691
accounts	−2.3582	0.0864	0.1372	0.9136	0.244
cards	−2.0471	0.1143	0.1783	0.8857	0.000

Table 9. Counterfactual Policy Experiments: Resolution and Awareness Channels

Description: The table reports every counterfactual from both models as a percentage change against its own model baseline, so that levers acting on different populations can be ranked on a common scale. The eight counterfactuals are defined in Table II.c. Both outcomes are evaluated at merit multiplier $m = 1$, and percentage changes are invariant to m. The two outcome columns measure different margins of the same problem. Restitution is expected relief $\sum_i \sigma_i R_i$ for the resolution channels and cumulative projected restitution for the awareness channels. Both are dollars of relief on the same complaint universe, so all eight levers rank against one another directly. Foregone welfare measures the intensive margin for the resolution channels, a flow over complaints that were filed and denied, and the extensive margin for the awareness channels, a stock over households that never file. Percentages in that column are read within a channel, against the baseline in the memo block. Columns 1 and 2 set each model at its widest scope, the full 2015–2024 complaint sample against the 120-month diffusion path starting in 2024. Columns 3 and 4 fix both to a single year (2024). The resolution counterfactuals are static, so their percentage effects are near-identical across the two horizons. The awareness counterfactuals are dynamic and compound with the horizon, which is the substantive content of the comparison. Appendix VII reports the awareness scenarios across the full λ grid. The memo block gives, for each horizon, the levels every percentage above it is measured against, in the column matching the concept.

	Full horizon		**2024 only**	
	2015–2024 sample vs. 120-month diffusion		2024 cohort vs. first 12 months	
percentage change (% Δ) in	**restitution**	**foregone welfare**	**restitution**	**foregone welfare**
Resolution channel (filer welfare; invariant to λ)				
CF1 universal self-advocacy	+7.67%	−1.43%	+7.54%	−1.44%
CF2 monitoring complementarity ×1.5	+1.17%	−0.26%	+1.07%	−0.23%
CF3 large-firm exposure for all firms	+11.77%	−1.15%	+12.97%	−1.42%
CF4 baseline resolution odds ×1.5	+39.07%	−7.09%	+38.82%	−7.30%
CF5 combined (CF1–CF4)	+67.91%	−10.92%	+68.84%	−11.47%
Awareness channel (non-filer welfare; $\lambda = 0.50$)				
CF6 close the income–awareness gap	+29.70%	−1.16%	+29.74%	−0.33%
CF7 double the diffusion speed	+135.89%	−9.99%	+6.73%	−0.16%
CF8 combined (CF6 + CF7)	+203.21%	−13.77%	+38.49%	−0.54%
*Context: **baseline levels ($ millions)***				
restitution, actual (paid)	176.4		42.3	
restitution, structural model — resolution channel	176.4		43.9	
restitution, SIR model — awareness channel (10-yr cum.)	887.6		44.2	
foregone welfare, structural model — filers		148.4		36.8
yearly foregone welf., SIR model — non-filers		8,283.8		8,509.5

Appendix to:

Fallible Firms:
The Market for Correcting Firm Mistakes

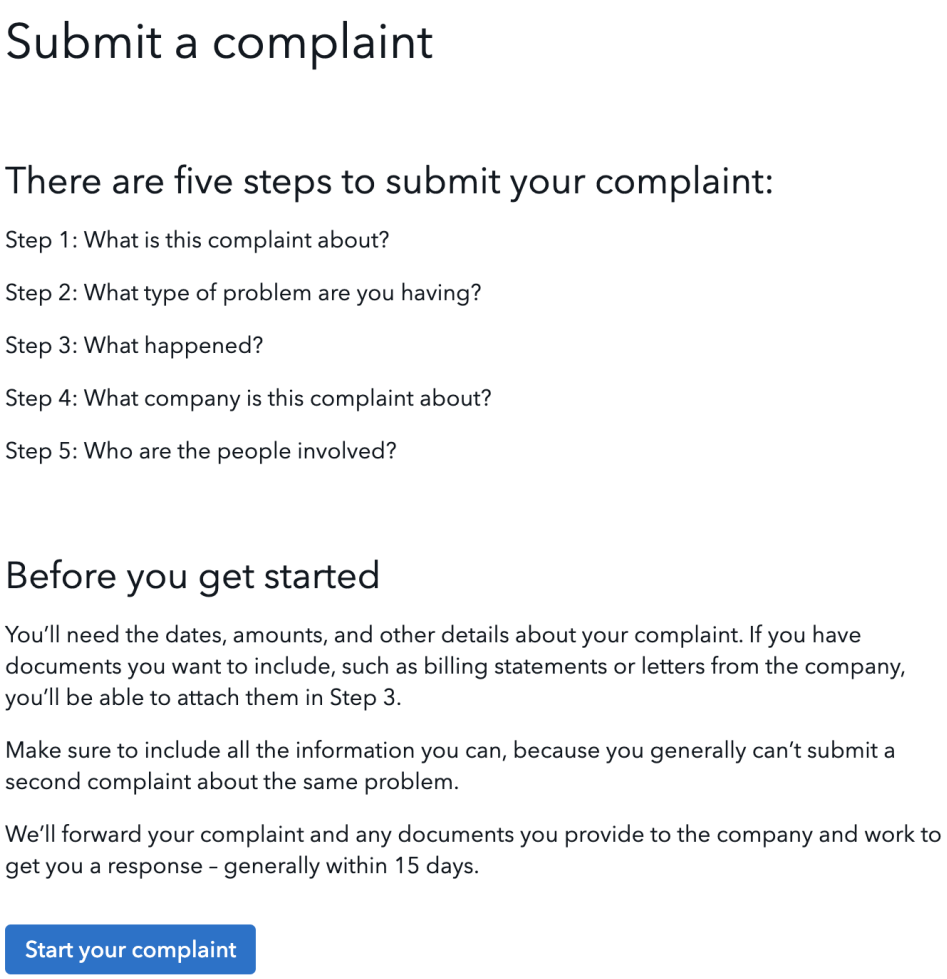
(intended for online publication)

Can be found at this link: Internet Appendix.

I Submitting a Complaint on the CFPB Website

Figure I.1. Filing a complaint on the CFPB website

The Figure shows the different steps for filing a complaint regarding a checking account.



Submit a complaint

What is this complaint about?

Choose the product or service that best matches your complaint.

- ○ Debt collection
- ○ Credit reporting, credit repair services, or other personal consumer reports
- ○ Mortgage
- ○ Credit card or prepaid card
- ○ Checking or savings account
- ○ Vehicle loan or lease
- ○ Student loan
- ○ Payday loan, title loan, or personal loan (installment loan or personal line of credit)
- ○ Money transfer, virtual currency, or money service (check cashing service, currency exchange, cashier's/traveler's check, debt settlement)

‹ | Previous Step 1 of 5 Next | ›

What type of problem are you having?

Most of the **checking account** complaints we get are about one of the following topics. Select the one that best describes your complaint. You will have the chance to explain your complaint in detail in the next step.

- ○ Opening an account
- ○ Managing an account (deposits, withdrawals, using ATM card, making or receiving payments, cashing a check, fees)
- ○ Closing an account
- ○ Problem caused by your funds being low
- ○ Problem with a lender or other company charging your account
- ○ Problem with credit report or credit score

‹ Previous | Step 2 of 5 | Next ›

What happened?

Describe what happened, and we'll send your comments to the companies involved.

- Include dates, amounts, and actions that were taken by you or the company.
- Do not include personal information, such as your name, account number, address, Social Security number, etc. We may ask for some of this information later, to help the company identify you and your account.

☐ I want the CFPB to publish this description on consumerfinance.gov so that others can learn from my experience. The CFPB will take steps to remove my personal information from this description but someone may still be able to identify me. Learn how it works. I consent to publishing this description after the CFPB has taken these steps.

Publishing this description will not affect how the CFPB handles your complaint.

What would be a fair resolution to this issue?

We'll forward this to the companies involved. Be specific so they know what resolution you are looking for. The company may or may not offer to resolve your complaint.

Attach documents (optional)

Include copies or photos of documents related to your issue, such as contracts, letters, and receipts, and we will forward all materials to the company for review.

Select a file to attach

We accept all file formats, except .dll, .dmg, .exe and .msi, up to 10 MB per file. All uploaded files will be scanned for viruses and will be immediately deleted if one is detected.

Select a file to upload

Previous | Step 3 of 5 | Next

What company is this complaint about?

We'll forward your entire complaint to the company and request they respond within 15 days of receiving it.

Bank or credit union

Company name

Company name

We will forward your complaint to this company and ask for a response.

We need this information to help the company find you in their system and respond to your complaint. (optional)

- [ ] Account number
  This number is on your billing statement
- [ ] Billing address

Previous | Step 4 of 5 | Next

Who are the people involved?

Identify who is involved in this complaint. This could include:

- "Just you" if you are the account holder or borrower
- "You and someone else" if you are submitting for yourself and want to include another account holder or co-borrower
- "Someone else" if you are submitting for someone else as an authorized third party, such as a lawyer, advocate, or power of attorney

- ◯ Just you
- ◯ You and someone else
- ◯ Someone else

Previous | Step 5 of 5 | Review

Figure I.2. Google Search Interest in "CFPB Complaint"

Description: The figure plots the monthly Google Trends search volume index (SVI) for the query "CFPB complaint" from January 2010 to December 2024. The index is normalized so that the month of peak search interest over the sample (December 2024) equals 100; months with insufficient search volume are reported by Google as zero. The thin gray line is the monthly index and the solid black line its trailing 12-month moving average. The dashed vertical line marks July 2011, when the CFPB began operations and opened its complaint portal.

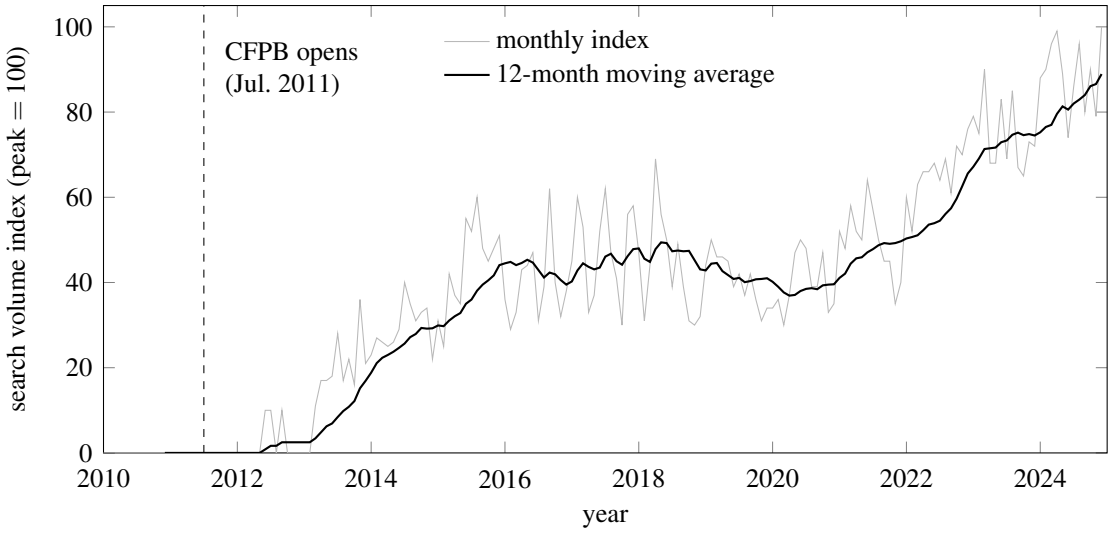


Table I.a. Example of Complaint

Description: The table presents a complaint from the CFPB database, available on the CFPB database webpage. For an explanation of the information attached to each complaint, refer to the CFPB database fields webpage.

Date received	1/17/19
Product	Mortgage
Issue	Trouble during payment process
Company	Ditech Financial LLC
State, Zip code	TX, 781XX
Company response	Closed with explanation

Complaint narrative: I have been trying to get my Private Mortgage Insurance Removed from my mortgage since XX/XX/XXXX when my mortgage dropped below 80 % loan to value. Last year my mortgage was sold from XXXX XXXX (Under mortgage # XXXX) to Ditech Mortgage (account # XXXX). I reached out to Ditech via a email (after being told to do so via phone representative) request to remove my PMI on mortgage on XX/XX/XXXX and received no response at all from them, I even checked my junk box and nothing was there. My mortgage papers that I signed state an "Automatic Termination of PMI" that states once my loan is below 78 % loan to value PMI will automatically terminate (I have attached this document). I reached out again today on XX/XX/XXXX to make this request via phone and was told initially to send the request that I already sent it too. I asked to speak with a supervisor and after being put on hold for about 30 minutes, I finally spoke to one. They told me that my loan to value must be under 70 % loan to value and that was their policy. After reading this document to the supervisor, I was told that "they don't have that document on file". She sent me a link to send her the form I have. I did so and just told me that I'll be hearing from them in 7-10 business days. Given their past history, I highly doubt that I will hear from them. I did mention to the supervisor and ask her why they weren't staying compliant to the homeowners protection act and she said nothing. From my understand this act requires mortgage companies to drop off PMI once loans are below 78 % LTV and the loan is current. I qualify for both of those items and don't understand why this is such a difficult task.

II Definitions

Table II.a. Variable Definitions

Description: The table defines the variables used in the main text. Subscripts denote complaint i, company c, respondent r, quarter of filing q, and filer's Census tract tr. Tract characteristics are matched through the filer's address, and the above-median indicator is defined against the median tract in the quarter of filing.

Variable	Definition
Outcomes	
Resolved with monetary relief	One if complaint i was resolved with monetary relief; zero otherwise. Denoted $solved\ w/MR_{i,c,r,q,tr}$.
Monetary relief amount	Dollar relief awarded on complaint i when it is resolved with monetary relief.
Filer income	
High-income tract	One if complaint i comes from a tract tr with above-median household median income in quarter q; zero otherwise. Denoted $I(above\ med.\ household\ med.\ income)_{i,q,tr}$.
ln(tract income)	Natural logarithm of household median income in Census tract tr in quarter q. Denoted $ln(household\ med.\ income_{i,q,tr})$.
Filer activity	
Self-advocacy (description)	One if complaint i's filer uses a *refund* or *fraud* term in the description field of complaint i; zero otherwise. Terms are listed in Table II.b. Denoted $I(filer\ self\text{-}advocacy\ in\ description)_{i,c,r,q,tr}$.
Self-advocacy (desired resolution)	Defined as above on the desired-resolution field. Denoted $I(filer\ self\text{-}advocacy\ in\ desired\ resolution)_{i,c,r,q,tr}$.
Attachment	One if complaint i's filer attached a supporting document when filing complaint i; zero otherwise. Denoted $I(filer\ attachment)_{i,c,r,q,tr}$.
Company activity	
Employee experience	Natural logarithm of the count of complaints company respondent r had already resolved before handling complaint i. Denoted $employee\ experience_{i,c,r,q,tr}$.
Employee login time	Percentage of days between the submission and the response of complaint i on which respondent r logged into the CFPB complaint system. Denoted $employee\ login\ time_{i,c,r,q,tr}$.
Regulatory scrutiny	
Post-report	One if complaint i was filed against company c in a quarter after c first received a confidential CFPB monitoring report card; zero otherwise. Denoted $I(post\text{-}report)_{i,c,q}$.
Fixed effects	
Quarter	Quarter q in which complaint i was filed.
Company	Company c against which complaint i was filed.
Product	The financial product named in complaint i.
County	County of the filer's Census tract tr.
Respondent	Company employee r assigned to answer complaint i.

Table II.b. Self-advocacy Terms

Description: The table lists the terms behind the two self-advocacy indicators defined in Table II.a. A complaint counts as self-advocating if the filer uses any term below, in the description field or in the desired-resolution field respectively. Matching is case-insensitive and is performed on the raw complaint text.

Stem	Terms matched
refund	refund, refunds, refunding, refunded, repay, repaying, repayment, repayments, reimburse, reimburses, reimbursing, reimbursed, reimbursement, reimbursements, pay back, paying back, paid back, make good, making good, made good, compensate, compensates, compensating, compensated, compensation, compensations, recoup, recoups, recouping, recouped, remunerate, remunerates, remunerating, remunerated, remuneration, remunerations, square accounts with, squaring accounts with, squared accounts with
fraud	fraud, fraudster, defraud, defrauded, deceit, deception, deceive, deceiving, deceived, trickery, trick, tricks, tricking, tricked, cheat, cheats, cheating, cheated, mislead, misleading, misled, misguide, misguiding, misguided, rip-off, fake, con, impostor

Table II.c. Definitions of the Policy Counterfactuals

Description: The table defines the eight counterfactuals reported in Table 9. Resolution channels perturb one primitive of the estimated resolution index η_i of Table 7 and re-evaluate $\sigma_i = \Lambda(\eta_i)$ on the complaints actually filed; they act on filer welfare and are invariant to the filing-given-awareness rate λ. Awareness channels modify the diffusion path of the awareness model and hold the policy function fixed; they act on non-filer welfare and are evaluated at the reference value $\lambda = 0.50$. Structural parameters are defined in Table 7 and complaint-level variables in Table II.a.

	Lever	Definition
Resolution channels (filer welfare; invariant to λ)		
CF1	Universal self-advocacy	Sets $q_i = 1$ for every complaint, so that every filer advocates for themselves.
CF2	Monitoring complementarity $\times 1.5$	Scales both monitoring interactions, $\delta_{q\omega}$ and $\delta_{\theta\omega}$, by 1.5.
CF3	Large-firm exposure for all firms	Holds every firm to the large-firm standard, $\mathrm{FS}_i \to 1$. This moves only the 36.7% of complaints filed against small firms, those failing the three-product and 250-complaint-per-year bar that defines $\mathrm{FS} = 1$.
CF4	Baseline resolution odds $\times 1.5$	Raises every product's baseline resolution *odds* by half, $\alpha_p \to \alpha_p + \ln 1.5$.
CF5	Combined	Applies CF1 through CF4 together.
Awareness channels (non-filer welfare; $\lambda = 0.50$)		
CF6	Close the income–awareness gap	Lifts the below-median awareness seed to the above-median one, $a_0(0) := a_1(0)$, the analogue of a one-time information campaign.
CF7	Double the diffusion speed	Doubles both tiers' diffusion rates, $\beta_\theta \to 2\beta_\theta$, the analogue of sustained outreach.
CF8	Combined	Applies CF6 and CF7 together.

III Ancillary Analysis

Table III.a. Respondent Fixed Effects

Description: The table investigates whether the relationship between filer income and the likelihood of receiving financial restitution is at least partially driven by the precise company employee taking care of the complaint—i.e., the *respondent r*. It does so by adding respondent fixed effects to the baseline specification in Equation 1. Since respondent ID information is not always available, Columns 1 and 3 replicate the baseline results of Table 3 over the subsample for which respondent ID information is available. Then, Columns 2 and 4 add respondent fixed effects. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	solved w/MR$_{i,c,q,r}$			
	(1)	(2)	(3)	(4)
High-income tract	0.00925***	0.00926***		
	(0.0016)	(0.0017)		
ln(tract income)			0.0139***	0.0140***
			(0.0024)	(0.0024)
Quarter FE	x	x	x	x
Company $\times$ product FE	x	x	x	x
County FE	x	x	x	x
Respondent FE		x		x
Observations	1,410,810	1,410,810	1,410,810	1,410,810
R^2	0.141	0.166	0.142	0.166

Table III.b. Financial Restitution Mechanisms, Log Tract Income

Description: The table repeats the specification of Table 5 measuring filer income as *ln*(tract income) rather than the above-median tract indicator. Variables are defined in Table II.a. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	Resolved with monetary relief					
	Filer activity		Company activity		Both	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(tract income)	0.0125***	0.00724***	0.0143***	0.0111**	0.0137***	0.00459
	(0.0021)	(0.0015)	(0.0025)	(0.0056)	(0.0024)	(0.0050)
Self-advocacy	0.0218***	-0.0631***			0.0206***	-0.0682***
	(0.0031)	(0.014)			(0.0031)	(0.016)
Attachment	0.0123***	-0.0666***			0.0120***	-0.0600***
	(0.0021)	(0.022)			(0.0019)	(0.016)
Self-advocacy × ln(tract income)		0.00759***				0.00791***
		(0.0014)				(0.0016)
Attachment × ln(tract income)		0.00702***				0.00641***
		(0.0020)				(0.0015)
Employee experience			0.000232	-0.0278***	0.000258	-0.0272***
			(0.00095)	(0.0077)	(0.00093)	(0.0078)
Employee login time			-0.0120***	0.221***	-0.0113***	0.221***
			(0.0032)	(0.067)	(0.0031)	(0.066)
Employee experience × ln(tract income)				0.00250***		0.00245***
				(0.00069)		(0.00070)
Employee login time × ln(tract income)				-0.0208***		-0.0207***
				(0.0061)		(0.0060)
Quarter FE	x	x	x	x	x	x
Company × product FE	x	x	x	x	x	x
County FE	x	x	x	x	x	x
Respondent FE			x	x	x	x
Observations	1,820,166	1,820,166	1,328,922	1,328,922	1,311,748	1,311,748
R^2	0.142	0.142	0.168	0.168	0.170	0.170

IV Companies Receiving a Monitoring Report

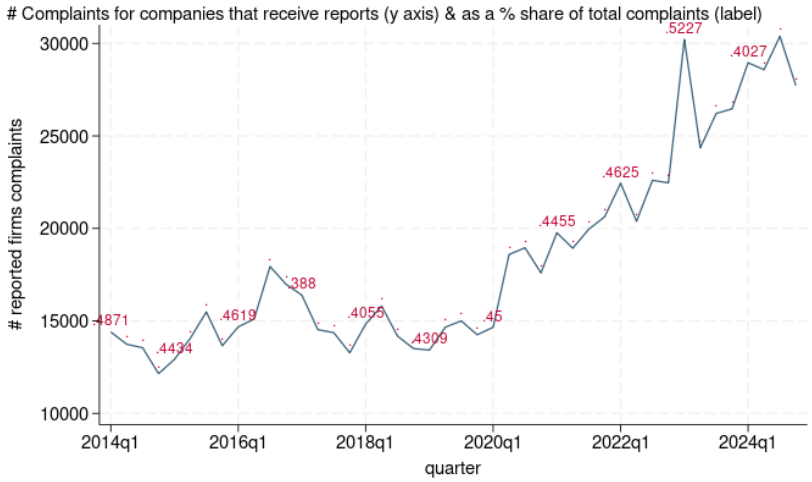
Figure IV.1. Firms Receiving a CFPB Monitoring Report

The figure reports information on CFPB monitoring reports. Panel A illustrates how many firms the CFPB has sent monitoring reports to between 2021 and 2024 (sign of negative feedback from the agency). Panel B plots the evolution over time of the total number of complaints received by firms that the CFPB has sent a monitoring report to between 2014 and 2024. The labels in red represent the share that these complaints represent of overall complaints, in percentage.

Panel A: Monitoring reports per year

Year	Reported Firms
2021	14
2022	20
2023	31
2024	20

Panel B: Complaints of reported firms



V CFPB Investigation Difference-in-Differences, Robustness

Table V.a. CFPB Company Monitoring Reviews across Income Groups

Description: The table presents estimations regarding the likelihood of the CFPB flagging a reviewed complaint as either not properly addressed by the company (Columns 1 to 3 across panels), or not having received a substantive response from the company (Columns 4 to 6 across panels). The tables regress these likelihoods on quintiles of the filer's Census tract household median income. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	I(company not address compl.)$_{i,c,q,r}$			I(response not substantive)$_{i,c,q,r}$		
	(1)	(2)	(3)	(4)	(5)	(6)
household med. income quintile$_{i,q,tr}$=1	-	-	-	-	-	-
household med. income quintile$_{i,q,tr}$=2	-0.00216	0.00103	-0.000363	-0.00117	0.00144	0.000171
	(0.0017)	(0.0016)	(0.0011)	(0.0014)	(0.0012)	(0.00097)
household med. income quintile$_{i,q,tr}$=3	-0.00421**	0.00119	-0.00109	-0.00237*	0.00194^{+}	-0.0000329
	(0.0016)	(0.0015)	(0.0013)	(0.0013)	(0.0012)	(0.00095)
household med. income quintile$_{i,q,tr}$=4	-0.00571***	0.00178	-0.00153	-0.00606***	-0.0000156	-0.00299^{+}
	(0.0018)	(0.0015)	(0.0030)	(0.0016)	(0.0012)	(0.0018)
household med. income quintile$_{i,q,tr}$=5	-0.00946***	0.00113	-0.00314	-0.00759***	0.000674	-0.00318
	(0.0025)	(0.0017)	(0.0027)	(0.0021)	(0.0014)	(0.0023)
Quarter FE	x			x		
Company FE		x			x	
product FE			x			x
Observations	194,338	193,914	194,338	194,338	193,914	194,338
R^2	0.0139	0.0583	0.00586	0.0142	0.0653	0.00510

Table V.b. The Financial Restitution Gap - Diff-in-Diffs, Balance Tests

Description: The table compares treated vs. controls (Panel A) and treated early vs. treated late (Panel B) before 2020 (the start of the company monitoring reviews at the CFPB). In panel A, Column 1 reports the pre-2020 quarterly average of the variables outlined across companies that will end up receiving a monitoring report (negative feedback from CFPB). Column 2 does the same for control companies, i.e., companies that will be reviewed but will not receive a monitoring report, and also had at least 100 complaints in 2018. In panel B, Column 1 reports the pre-2020 quarterly average of the variables outlined across companies that will end up receiving a monitoring report (negative feedback from CFPB) late in the company monitoring reviews time period (2023-2024). Column 2 does the same for companies that will end up receiving a monitoring report (negative feedback from CFPB) early in the reviews time period (2020-2022). Across panels, Column 3 reports the difference across the previous two columns. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

Panel A: Treated vs Controls

	Reported	**Control**	**Difference**
	(1)	(2)	(3)
ln(# complaints)	4.60	3.16	1.44***
% deposit-related complaints	39.21%	9.47%	29.74%***
% loan-related complaints	60.78%	90.54%	-29.75%***
% complaints w/filer attachment	13.71%	8.91%	4.80%***
% complaints from bottom income quartile	27.81%	27.83%	-0.02%

Panel B: Treated Late vs Treated Early

	Late Report	**Early Report**	**Difference**
	(1)	(2)	(3)
ln(# complaints)	3.42	5.59	-2.18***
% deposit-related complaints	12.11%	61.75%	-49.64%***
% loan-related complaints	87.89%	38.25%	49.64%***
% complaints w/filer attachment	11.77%	15.32%	-3.55%***
% complaints from bottom income quartile	31.82%	24.47%	7.35%***

Table V.c. CFPB Investigations - Difference in Differences, Testing Complaints Composition

Description: The table presents difference-in-differences estimates around the first time a company receives an investigation report (negative feedback from the CFPB). The sample comprehends all treated companies and, as controls, those companies that were also reviewed but did not receive a report and have at least 100 complaints in 2018 (pre-reviews). The outcome variable is $I(low\text{- }income\ complaint)_{i,c,q,r}$, a dummy variable equal to one if complaint i against company c has been filed in quarter q by a filer from a Census tract that is in the bottom quartile of Census tracts by household median income, zero otherwise. Standard errors are clustered at the company level. ***, **, and * denote 1%, 5%, and 10% statistical significance, respectively.

	Low-income complaint		
	(1)	(2)	(3)
Post-report	0.00523	0.00244	0.00591*
	(0.0040)	(0.0040)	(0.0031)
Quarter FE	x	x	x
Company FE	x		
Company × product FE		x	x
County FE			x
Observations	1,003,708	1,003,541	1,003,388
R^2	0.0228	0.0274	0.208

VI The Structural Model of Complaint Resolution: Full Details

This appendix sets out the structural model of Section V in full: the environment and the two agents' problems, the closed-form mapping from the firm's first-order condition to the estimated policy function, the identification of each parameter block, the estimator and its sample, the welfare accounting, the counterfactual design, and the robustness exercises.

The Setting

Time. The firm's decision is modelled statically: each complaint is a single grant/deny choice. Households are static one-period decision-makers, filing once per error. The regulator's monitoring technology is exogenous from the perspective of both firms and households.

Products. A finite set $\mathscr{P}$ of financial products—accounts, credit cards, mortgages, auto loans, student loans, and other (consumer) loans. Each product $p \in \mathscr{P}$ carries a complexity κ_p: the difficulty of obtaining relief on it at baseline. Complexity is not a free input. It is read off the estimated product intercept α_p, so it moves with the model rather than being imposed on it.

Agents and observables. Household i chooses complaint quality $q_i \in \{0,1\}$, observed as a self-advocacy indicator coded from the complaint narrative—whether the filer mentions refund- or fraud-related terms in the complaint description. Firm f chooses whether to grant relief, observed as the closed-with-monetary-relief indicator in the complaints database, with the conditional restitution amount observed on granted complaints. The regulator/CFPB's actions enter through two firm-level objects: the monitoring report card state and the firm's size class.

Household types. θ_i is filer income, a binary above-/below-median census-tract indicator. $q_i = 1$ marks a strong complaint—one using self-advocacy language, citing fraud, requesting a refund—and $q_i = 0$ a basic one.

The latent grievance. g_i is the objective dollar magnitude of the underlying financial error at complaint i: the size of an overcharge, the discrepancy on a balance, the value of a contractual

violation. It is a property of the underlying transaction, set before any complaint is filed and independent of complaint quality. A complaint is meritorious if $g_i \geq \bar{g}_{p_i}$, with $\bar{g}_p$ a product-specific threshold separating substantive grievances from trivial ones.

The grievance is never estimated. The data record realized restitution r_i on the resolved subsample only, and even there r_i is a noisy signal of g_i rather than g_i itself. The model therefore works with g_i only through the merit probability $\phi(q, \theta, p) = \Pr(g_i \geq \bar{g}_p \mid q, \theta, p)$—the probability that a denial was wrongful, which is what enters the welfare formula. In this model the merit probability is taken proportional to the firm's own resolution probability, on the reasoning that the firm grants when the complaint looks meritorious; the fitted $\widehat{\sigma}_i$ therefore doubles as the per-complaint merit weight. The level g_i appears in no moment and no objective.

Firm monitoring state. $\omega_{f,t} \in \{0,1\}$ indicates that firm f has received a CFPB monitoring report card over the previous year. It is the payoff-relevant firm state: under a report card the firm's returns to resolving engaged complaints steepen. The systematic monitoring campaign begins in 2021, so the post-2021 window is the identifying period for monitoring-related effects.

Firm size. $\text{FS}_f \in \{0,1\}$ equals one for companies that span at least three products *and* clear at least 250 complaints in every year they appear. It is a firm-size indicator, and it proxies for two things jointly: scale itself, and the supervisory intensity that comes with scale, since these are the firms the CFPB examines most closely and the ones to which report cards are almost exclusively issued. Nothing in the cross-section separates the two readings; we write "large" and "small" for the two classes throughout. It is a time-invariant firm attribute and enters as a regressor: the two classes differ in the baseline cost of denying, which is a parameter of the firm's problem.

Complaints perimeter. Credit report and debt-collection complaints rarely seek an in-kind service obligation—their relief rate is under one percent—and are largely mis-filed credit-reporting issues, so reading their resolution rate as a firm deny-cost parameter is suspect. They are dropped at load, before the product set is built, so they never enter the policy function, the welfare accounting, or the diffusion merge. We also exclude monetary transfer complaints, as they should entail more

transparent restitution mechanics, and the related input data for the following awareness model were not publicly available.

The Household's Problem

A household holds product p_i, has income type θ_i, and suffers a financial error of objective size g_i. Conditional on being aware of the CFPB, it makes two nested choices: whether to file, and—if filing—at what quality. The expected utility of filing at quality q is

$$U_i(q) \;=\; \underbrace{\sigma(q,\theta_i,p_i;\omega_i)}_{\text{prob. of relief}} \cdot \underbrace{\mu(g_i,p_i)}_{\text{prize if granted}} \;-\; \underbrace{c(q,\theta_i,\kappa_{p_i})}_{\text{documentation cost}} \;+\varepsilon_i^q, \tag{VI.1}$$

with $\sigma(\cdot)$ the firm's resolution policy, taken as given by the household—equilibrium consistency, in that the household best-responds to the policy function derived below—and rising in q; $\mu(g_i,p_i) = \mathbb{E}[r_i \mid g_i,p_i,\text{grant}]$ the expected restitution conditional on a grant (a property of the underlying error and independent of q); $c(q,\theta,\kappa_p) = c_q\,q + c_{q\theta}\,q\theta + c_{q\kappa}\,q\kappa_p$ the cost of assembling a complaint of quality q, with $c_q > 0$, $c_{q\theta} < 0$, and $c_{q\kappa} > 0$; and ε_i^q an i.i.d. Type-I extreme-value taste shock, the outside option of not filing carrying payoff zero plus its own shock.

The quality margin. The household upgrades from $q = 0$ to $q = 1$ when the marginal expected benefit weakly exceeds the marginal cost,

$$\big[\sigma(1,\theta_i,p_i;\omega) - \sigma(0,\theta_i,p_i;\omega)\big]\,\mu(g_i,p_i) \;\geq\; c_q + c_{q\theta}\,\theta_i + c_{q\kappa}\kappa_{p_i}, \tag{VI.2}$$

net of the taste shocks. With extreme-value shocks this is a binary logit in the net benefit of upgrading, $\Pr(q_i = 1 \mid \text{file}) = \Lambda\big(\Delta\sigma_i\,\mu(g_i,p_i) - \Delta c_i\big)$.

The participation margin. The household files if and only if $\max_q U_i(q) \geq 0$. Conditional on awareness, this participation probability is exactly the filing-given-awareness rate the diffusion model of Appendix VII carries, marginalized over the grievance.

Testable sign predictions. The comparative statics are unambiguous: the probability of self-advocacy rises with income (which both lowers Δc through $c_{q\theta} < 0$ and raises $\Delta\sigma$), falls with product complexity (which raises Δc through $c_{q\kappa} > 0$), and rises with the size of the prize. All three are visible directly in the data.

How the household problem feeds the estimation. Participation cannot be recovered from complaint records, which contain filed complaints only—households that experience an error but never file never appear—so this margin is supplied by the awareness-diffusion model, whose product of awareness and filing-given-awareness is precisely $\Pr(\text{aware}) \times \Pr(\text{file} \mid \text{aware})$. Quality is observed directly on the narrative subsample; the cost parameters $c_q, c_{q\theta}, c_{q\kappa}$ are not separately estimated. The household problem instead does two jobs: it supplies the economic reading of q and the conditional exogeneity that lets q enter the firm's policy function as a covariate.

The Firm's Problem

For each complaint of type (q, θ, p) at a firm in monitoring report card state ω and size class $c \in \{\text{large}, \text{small}\}$, the firm chooses grant or deny to minimize expected cost. Granting costs the restitution paid, R_p—product-specific and observed. Denying is costly through regulatory exposure: denying a complaint that is in fact meritorious raises the firm's poorly-handled rate and hence the probability the regulator escalates. That exposure is not uniform across firms, because the regulator's supervisory attention is concentrated on firms of scale, so the size class shifts both how sensitively a poorly-handled rate translates into a report and how large an escalation is. The expected dollar cost of denying is

$$D(q, \theta, p, \omega, c) = \underbrace{\phi(q, \theta, p)}_{\text{merit prob.}} \cdot \underbrace{\rho_0'(c)\,\Phi(c)}_{\substack{\text{report sensitivity} \\ \times \text{ penalty}}} \cdot \underbrace{s(q, \theta, \omega)}_{\text{monitoring scrutiny}}, \tag{VI.3}$$

with $\rho_0'(c)$ the baseline marginal report probability per unit of poorly-handled rate and $\Phi(c)$ the dollar penalty of an escalation. For a small firm, which the regulator rarely examines, $\rho_0' \approx 0$: a denial generates almost no regulatory exposure, whatever the complaint looks like.

The threshold rule. The firm grants if and only if the deny cost weakly exceeds the grant cost, equivalently if and only if the cost ratio D/R_p is at least one.

Two log-linear primitives. Merit rises with quality and income, and monitoring concentrates the regulator's scrutiny on exactly those engaged dimensions:

$$\phi(q,\theta,p) = \phi_0(p)\, e^{\varphi_q q + \varphi_\theta \theta}, \qquad s(q,\theta,\omega) = e^{(\delta_{q\omega} q + \delta_{\theta\omega}\theta)\,\omega}, \tag{VI.4}$$

with $\varphi_q, \varphi_\theta \geq 0$ the merit log-gradients and $s \equiv 1$ whenever $\omega = 0$ or $q = \theta = 0$. Monitoring scales up the penalty for denying an *engaged* complaint but leaves a baseline complaint's deny cost untouched. This normalization is not a convenience: it is imported from the paper's own reduced-form evidence, in which both the two-way fixed-effects regressions and the Callaway–Sant'Anna estimators show no post-report effect for low-income, low-advocacy complaints.

The index, and closed forms for α_p and γ. Taking logs of the cost ratio turns "grant iff $D/R_p \geq 1$" into "grant iff $\eta \geq 0$," where

$$\eta(q,\theta,p,\omega,c) \;=\; \log \frac{D(q,\theta,p,\omega,c)}{R_p(c)} \;=\; \alpha_p + \gamma \mathrm{FS} + (\varphi_q + \delta_{q\omega}\omega) q + (\varphi_\theta + \delta_{\theta\omega}\omega)\theta, \tag{VI.5}$$

which is the estimated index. The baseline log cost ratio separates additively into a product component and a firm-size component,

$$\begin{aligned} \alpha_p &= \log \frac{\phi_0(p)\,\rho_0'\,\Phi}{R_p} \;=\; \log \frac{\text{baseline expected deny cost at } p}{\text{grant cost at } p}, \\ \gamma &= \underbrace{\log \frac{\rho_0'(\text{large})\,\Phi(\text{large})}{\rho_0'(\text{small})\,\Phi(\text{small})}}_{\text{enforcement exposure}} \;-\; \underbrace{\log \frac{R_p(\text{large})}{R_p(\text{small})}}_{\text{grant cost}}, \end{aligned} \tag{VI.6}$$

with α_p evaluated at a small firm. The two are the same object read along two dimensions: the log ratio of what denying costs to what granting costs, indexed by product and by firm size.

Reading α_p. The intercept is the log ratio of the firm's baseline enforcement exposure to the restitution it would pay, at a reference complaint of product p at a small firm. $\Lambda(\alpha_p)$ is the baseline resolution probability; $\alpha_p > 0$ would mean denying a reference complaint costs the firm more than granting. In the complaints data, baseline resolution is well below one half everywhere, so we expect $\alpha_p < 0$ throughout: at a typical complaint the firm's expected enforcement cost from denying is only a fraction of the restitution it would pay, so it denies. The four dollar primitives $\phi_0(p), \rho_0', \Phi, R_p$ enter only through this single ratio, so the data identify α_p but none of the four separately (in particular R_p is absorbed here and reappears, in levels, only in welfare). Equivalently, the complement $1-\Lambda(\alpha_p)$ is the baseline denial rate—what the firm gets away with. A very negative α_p means denying carries little enforcement cost relative to paying, so the firm can deny even meritorious complaints with near-impunity; this is the high-complexity, low-resolution end of the product space. So α_p, the complexity index

$$\kappa_p \;=\; 1 - \frac{\Lambda(\alpha_p)}{\max_{p' \in \mathscr{P}} \Lambda(\alpha_{p'})}, \tag{VI.7}$$

and the denial rate $1-\Lambda(\alpha_p)$ are three monotone readings of one baseline number, with "gets away with more" equivalent to lower α_p. Table 8 reports all of them.

Reading the remaining terms. The merit log-gradients $\varphi_q, \varphi_\theta$ say that a documented or articulated claim, or a higher-income filer, is more likely to be a substantiated grievance the firm prefers to settle. The interactions $\delta_{q\omega}, \delta_{\theta\omega}$ are the monitoring report card-induced steepening of those gradients coming from $s(q,\theta,\omega)$: under stricter scrutiny, the expected penalty for denying an engaged complaint rises, so the returns to engagement go from $\varphi_q, \varphi_\theta$ to $\varphi_q + \delta_{q\omega}, \varphi_\theta + \delta_{\theta\omega}$. Lastly, γ captures the effect of higher complaint-handling capacity and supervisory standards, and it sits alongside, and not against, the monitoring report card effects. This happens for two reasons: (i) firms react to the enforcement regime they are in, and to scrutiny they can see, not to scrutiny

they cannot foresee; (ii) among small firms, only 24 of 2,407 ever receive a monitoring report card, against 25 of 36 large ones, so there is almost no variation from which to learn how a small firm responds to a report card (92.7% of all complaints at firms that have received a monitoring report card are filed against large firms).

From the threshold to the logit. "Grant iff $\eta \geq 0$" is deterministic and would make every cell all-grant or all-deny; observed within-cell resolution rates lie strictly inside $(0,1)$. We therefore add an idiosyncratic, econometrician-invisible shock ζ_i to the firm's perceived *log cost ratio*—the case officer's private read, internal documentation, processing idiosyncrasies—so the firm grants if and only if $\eta_i + \zeta_i \geq 0$ with ζ_i standard logistic. Integrating over ζ_i,

$$\sigma_i \;=\; \Pr(\zeta_i \geq -\eta_i) \;=\; \Lambda(\eta_i), \tag{VI.8}$$

which is the bridge from first-order condition to estimator.

Identification

Each term of the index maps to a distinct piece of identifying variation:

Index term	Structural reading	Identified from
α_p	product baseline / complexity	product-level resolution rates
γ	firm-size baseline shift	loads on 36 large against 2,407 small firms
φ_q	self-advocacy (merit) gradient	variation in self-advocacy
φ_θ	income (merit) gradient	variation in tract income
$\delta_{q\omega}$	monitoring $\times$ advocacy steepening	cross-sectional variation under monitoring state
$\delta_{\theta\omega}$	monitoring $\times$ income steepening	cross-sectional variation under monitoring state

The Estimator and the Estimation Sample

Data. For the modeling, we use the 2015-2024 complaint-level dataset. Credit reporting, debt collection, and monetary transfer complaints are dropped, taking the raw data to 1,034,613 complaints across six products and 4,618 firms. This is the population on which welfare is evaluated.

Sample. Model identification requires two overlapping restrictions. Complaint quality is coded from the narrative, so the advocacy channel is identified only on the publicly available narrative subsample (486,787 complaints, about 47% of the panel); and the monitoring report card state has meaningful, non-selectively-assigned variation only after the systematic review campaign begins in 2021, so the monitoring channel requires the post-2021 window (509,052 complaints). Their intersection, 266,287 complaints at 2,443 firms, is the structural estimation sample.

Estimator. We use a binary logit by maximum likelihood, with analytic gradient and an L-BFGS-B optimizer. The parameter vector is laid out as the product baselines, then φ_q, φ_θ, $\delta_{q\omega}$, $\delta_{\theta\omega}$, and the firm-size shifter last, standard errors are company-clustered. Maximum likelihood is the natural estimator here because the logistic form is part of the economic specification—the distribution of the firm's idiosyncratic grant/deny payoff—rather than an auxiliary statistical assumption; and because the model is a static binary logit on a linear index, its log-likelihood is globally concave, so the estimator has a unique maximizer and the convergence, multiple-starting-value, and local-optimum concerns that attend full-solution dynamic models do not arise.

Estimates. A single unconstrained fit over all six blocks delivers the estimates in Table 7, with log-likelihood $-89{,}967.9$. The product baselines are sharp, with $|t|$ between 6 and 18: the maximum-likelihood score conditions match product means exactly, so α_p is pinned by the product-level resolution rates. The self-advocacy gradient is sharp ($t = 3.57$). The income gradient is indistinguishable from zero ($t = 1.14$). The firm-size shifter is a between-firm contrast, so company clustering bites hardest there and it lands at $t = 1.56$. The monitoring interactions are weakest by far ($t = 0.63$ and 0.41): they are identified off cross-sectional variation in a selectively assigned treatment, and only 49 firms are observed switching state at all.

Fit Assessment

A logit's score conditions mechanically reproduce the cells that identify its free parameters—the product means, the $q = 1$ and $\theta = 1$ marginals, the FS $= 1$ marginal, and the two monitoring report card state interaction cells. Therefore, the fit diagnostics report only moments the additive

specification is not free to match. Given sample sizes in the hundreds of thousands, the absolute gaps rather than the standardized residuals are the informative quantity. Table VI.a collects them.

The pooled monitoring report card effect. The additive index has no standalone ω term, so it cannot generate a pooled monitoring report card jump directly. The empirical jump is $+8.97$ percentage points against a model jump of $+8.68$; the bulk of it is product composition working through α_p.

The low-engagement split. At $q = \theta = 0$ both interaction terms are identically zero, so the index has *no* term able to lift these cells. Empirically, monitoring lifts $q = 0$ cells by 7.37 percentage points and $\theta = 0$ cells by 8.76; the model delivers 6.42 and 8.14, short by 0.95 and 0.62 points respectively.

The income-by-advocacy interaction. This is the cleanest test, because the additive index omits any $q \times \theta$ term and is not monitoring-selection-confounded. The empirical difference-in-differences on this sample is $+0.0017$ and the model delivers $+0.0026$—so the curvature of Λ alone reproduces the complementarity present here, and overshoots it.

Product-by-monitoring report card patterns. The index has no product-by-ω term, so the monitoring report card lift is common in log-odds. In probability space this under-delivers for the low-baseline products—student loans, consumer loans, mortgages—where the data show large *relative* monitoring responses the additive index cannot generate; the mean absolute miss across products is 1.65 percentage points. This block is descriptive and cross-sectional, hence selection-confounded.

Welfare Accounting

Filer welfare. Per denied complaint the expected dollar loss is $W_i = m \cdot \widehat{\sigma}_i \cdot R_{p,t,\omega}$: the monetary relief the complainant would have received in expectation had the firm resolved $R_{p,t,\omega}$, times $\widehat{\sigma}_i$ pinned at the maximum-likelihood estimate and doubling as the merit weight, scaled by the merit multiplier m. Summing over the complaints that were actually denied gives W^{denied}. The formula decomposes into two clean pieces. The first is the dollar stake from the complaints data, which

enters linearly because a wrongfully denied complainant loses exactly the restitution owed. The second is the expected merit of the complaint, evaluated at the complaint's *actual* covariates and pinned at the estimate: it does not move when a counterfactual changes income, advocacy, or the baseline. This is deliberately distinct from the denial *weight* $(1-\sigma)$ in the counterfactuals, which does respond to policy.

The merit multiplier grid. The fitted $\widehat{\sigma}_i$ is the model's *positive* prediction of how likely resolution was. How *wrongful* a denial is—whether each denial was fully wrongful or only partly—is a normative quantity the data cannot pin down. Welfare is therefore reported across $m \in \{1, 1.5, 2\}$: $m = 1$ values a denial at exactly the expected relief foregone from the firms' point of view according to complaints data, and $m > 1$ treats denials as proportionately more wrongful. Foregone welfare scales linearly in m, while distributional shares by income and product, and counterfactual percentage changes, are invariant to it.

Dollar stakes. The per-product stake R_p is the mean of realized relief over granted complaints, by product. The per-complaint stake used in welfare is allowed to vary with the monitoring report card state, because monitored complaints carry materially larger realized stakes;[29] it is the product-year-monitoring cell mean with a three-tier fallback—the cell mean where it rests on at least 30 observations, else the product-year mean, else the product mean. Because the post-2021 estimation sample carries materially larger stakes than the full panel, R_p is recomputed on whichever sample the model is fit on, with the full-sample value retained for reference.

Scaling to the full sample. The publicly available narrative subsample is about 47% of complaints; the rest have product, income, year, monitoring report card status, size class, and outcome, but no coded advocacy. To reach the full panel we compute bucket advocacy rates $\Pr(q = 1 \mid p, t, \omega)$ on the publicly available narrative subsample with a three-tier fallback requiring at least 20 obser-

[29] Two independent checks are run to validate this choice. The first models the relief amount with a multiplicative conditional mean, $\mathbb{E}[r_i \mid X_i] = \exp(\beta_0 + \beta_p + \beta_q q_i + \beta_\theta \theta_i + \beta_\omega \omega_i)$, estimated by Poisson pseudo-maximum-likelihood with product fixed effects on the narrative subsample. It returns $\widehat{\beta}_q = +0.273$, $\widehat{\beta}_\theta = +0.097$, and $\widehat{\beta}_\omega = +0.539$, so the monitoring multiplier on expected restitution is $e^{0.539} \approx 1.71$. The second check is the raw granted-cell ratio $R_{p,\omega=1}/R_{p,\omega=0}$, product by product: 1.82 for accounts, 1.15 for auto loans, 1.44 for cards, 1.27 for mortgages, 1.84 for consumer loans, and 1.94 for student loans. The two checks share direction, and a positive β_ω together with ratios above one, justifying the modeling choice.

vations per bucket, let each full-sample complaint inherit its bucket rate, and evaluate the policy function at that *continuous* imputed value. Foregone welfare is then summed over complaints that were *actually* denied rather than over a model-implied $(1-\sigma)$, because the outcome is observed for every complaint so the denied set is known exactly; this keeps the full-sample table on the same definition as the estimation-sample table and imports no policy-function fit error into the distributional shares.

Results. On the estimation sample the model reproduces the dollars actually paid to within 0.4% ($66.7 million expected against $67.0 million disbursed) and yields $55.6 million of foregone filer welfare at $m = 1$, a ratio of 0.83 dollars foregone per dollar paid. On the full panel the corresponding figures are $176.4 million expected against $176.4 million paid (a ratio of 1.000) and $148.5 million foregone, a ratio of 0.84; 56.0% of the loss accrues to above-income median filers and 44.0% to below-income median ones. Because size class enters the fitted probability through γ, the projection carries each firm's own baseline outward rather than extrapolating the large firms' resolution rate onto the rest, and the model-versus-actual relief ratio is reported separately by class as a direct check that it does: 0.994 for large firms and 1.025 for small ones. Both classes reproduce within three percent. Absent γ only one baseline would be on offer, and the model would under-predict the large-firm block and over-predict the small-firm one by a wide margin—which is why γ is retained on fit grounds, corroborated by its t-statistic.

The monitoring report card effect and the δ band. Since the $\delta_{q\omega}$ and $\delta_{\theta\omega}$ are the least sharply identified block, they are the main channel through which sampling noise reaches the welfare headline. For robustness, we therefore set up a band to understand how much welfare calculations are skewed by this noise. The band is is $\widehat{\delta} \pm 1.96\,\mathrm{SE}$ in log-odds on the company-clustered standard errors, with both interactions moved to the same edge—the widest, perfectly correlated band. For each band point the whole pipeline re-runs: pin both interactions at that edge, refit the baselines, the firm-size shifter, and the merit gradients by maximum likelihood so that the trade-offs among them are respected rather than frozen, then recompute the headline. Table VI.b reports what the band does to the two quantities that are pure model output on the estimation

population—foregone filer welfare and model-predicted restitution. The headline moves by -5.4% from end to end, which is small relative to the merit grid and an order of magnitude smaller than the spread the multiplier m opens up, for reference. Restitution is reported beside foregone welfare because the two move in lockstep—both are built from the same fitted probability—but only restitution has an observed counterpart. The predicted-to-actual ratio is closest to one at the point estimate and drifts away from it in both directions by at most 2.8 percentage points, a reassuring finding in as the two parameters' dispersion does not impact model fit and predictions as much.

Counterfactual Design

Five scenarios perturb one structural primitive of the index each and let the resolution probability respond.

CF	What changes	Policy interpretation
CF1	$q_i = 1$ for all complaints	universal self-advocacy
CF2	$\delta_{q\omega}, \delta_{\theta\omega} \times 1.5$	monitoring report card steepens the engagement gradients
CF3	$\text{FS}_i \to \max(\text{FS}_i, \chi)$, $\chi = 1$	small firms held to the large-firm standard
CF4	$\alpha_p \to \alpha_p + \ln 1.5$	baseline product resolution odds raised by half
CF5	CF1–CF4 together	combined

CF1 explores full self-advocacy across filer, as an educational campaign on how to file complaints and properly self-advocate. CF2 mimics a 50% additional steepening of the income and self-advocacy gradients under monitoring report card, similar to more perceived pressure on the firms from receving the report.[30] CF3 acts on coverage, mimicking a policy of holding small firms to the standard the large ones already meet—implemented as $\text{FS}_i^{\text{CF}} = \max(\text{FS}_i, \chi)$ with coverage $\chi \in [0,1]$, so that only the 37.0% of complaints filed against small firms move when χ=1 and large firms/γ are untouched by construction. CF4 is an additive log-odds shift that, because resolution

[30] To be noted that this is the a weaker counterfactual, as the structural model estimated $\delta_{q\omega}$ and $\delta_{\theta\omega}$ were not significant to start with, albeit reproducing a similar effect as the reduced-form identified one (see Table 7, subpanel *Implied monitoring DiD*).

is rare, almost exactly mimics a +50% increase in the baseline resolution probability within each product—meaning firms can get away with 50% less of what they are currently getting away with (on complaints filed from low-income tracts, without self-advocacy, and directed to small firms). CF5 applies CF1 to CF4 at the same time. Three additional awareness counterfactuals complete the set: CF6, closing the income–awareness gap; CF7, doubling the diffusion speed; and CF8, both CF6 and CF7 together. Since they act on the diffusion path rather than on the policy function, they are set out in Appendix VII.

How foregone filer welfare and projected restitution are computed under each. For each scenario we recompute $\sigma_i^{\mathrm{CF}} = \Lambda(\eta_i^{\mathrm{CF}})$ from the perturbed primitive and evaluate

$$W^{\text{foregone, CF}} = \sum_i \left(1 - \sigma_i^{\mathrm{CF}}\right) \cdot \left[m \cdot \widehat{\sigma}_i \cdot R_{p,t,\omega}(i)\right]. \tag{VI.9}$$

$$Restitution^{\text{granted, CF}} = \sum_i \sigma_i^{\mathrm{CF}} \cdot \left[R_{p,t,\omega}(i)\right]. \tag{VI.10}$$

The two roles of σ in equation VI.9 are deliberately distinct. The denial weight $(1 - \sigma_i^{\mathrm{CF}})$ *responds* to the policy—a stronger baseline or more advocacy resolves more complaints, so fewer remain denied, which is the point of the counterfactual—while the per-denial value uses the *pinned* estimate at the complaint's actual covariates, held fixed. By the same token, only the approval rate σ_i^{CF} that responds to the policy counterfactual enters restitution projections in equation VI.10. In Table 9, each change is differenced against the *observed* baseline, so every counterfactual reads as "counterfactual versus the real-world status quo"; the model-baseline row is reported alongside so that its gap against the observed figure *is* the model's baseline fit error.

Results. Table 9 reports all five counterfactuals at two horizons: the full 2015-2024 complaint database, and 2024 alone. On the full sample the ranking is CF4 (+39.1% of expected relief) well ahead of CF3 (+11.8%) and CF1 (+7.7%), with CF2 (+1.2%) negligible; combining all four gives +67.9%. Because the resolution counterfactuals are static, the 2024 cut returns near-identical percentages (+38.8%, +13.0%, +7.5%, +1.1%, +68.8%). The foregone-welfare column moves

far less than the restitution column—CF5 raises relief paid by more than two thirds while reducing foregone filer welfare by only 10.9%—which is the arithmetic of a low-resolution system: raising the resolution probability moves complaints out of the denied set, but the vast majority remain denied under every scenario, and those that remain are still valued at their own merit weight.

Table VI.a. Fit Diagnostics: Non-Targeted Moments

Description: The table compares empirical and model-implied resolution rates on moments the additive index is *not* free to match. Model σ is the fitted $\Lambda(\eta_i)$ of Table 7 averaged over the cell; the last column is the empirical rate less the model rate, in resolution-probability units. All cells are on the estimation sample ($N = 266{,}287$). Panel A reports the pooled monitoring report card jump, which the index has no standalone ω term to generate. Panel B asks whether the monitoring report card effect reaches complaints with no engagement: at $q = \theta = 0$ both δ terms are identically zero, so the index has *no* term able to lift these cells. Panel C is the cleanest test—the index omits any $q \times \theta$ term, and unlike every ω moment it is not monitoring report card-selection-confounded. Panel D reports the monitoring report card jump product by product; the index has no product-by-ω term, so the lift is common in log-odds.

Moment	N	Empirical σ	Model σ	Emp. − model
Panel A: pooled monitoring level				
$\sigma(\omega = 0)$	166,473	0.0894	0.0905	−0.0011
$\sigma(\omega = 1)$	99,814	0.1791	0.1772	+0.0018
jump Δ_ω		+0.0897	+0.0868	+0.0029
Panel B: low-engagement monitoring split				
$\sigma(\theta = 0, \omega = 0)$	80,555	0.0821	0.0844	−0.0022
$\sigma(\theta = 0, \omega = 1)$	45,321	0.1698	0.1658	+0.0040
monitoring lift on $\theta = 0$		+0.0876	+0.0814	+0.0062
$\sigma(q = 0, \omega = 0)$	50,293	0.0753	0.0789	−0.0036
$\sigma(q = 0, \omega = 1)$	30,717	0.1490	0.1431	+0.0059
monitoring lift on $q = 0$		+0.0737	+0.0642	+0.0095
Panel C: income × self-advocacy interaction				
$\sigma(q = 0, \theta = 0)$	40,142	0.0953	0.0956	−0.0003
$\sigma(q = 1, \theta = 0)$	85,734	0.1223	0.1222	+0.0002
$\sigma(q = 0, \theta = 1)$	40,868	0.1110	0.1107	+0.0003
$\sigma(q = 1, \theta = 1)$	99,543	0.1397	0.1398	−0.0001
$q \times \theta$ DiD		+0.0017	+0.0026	−0.0010
Panel D: product × monitoring (jump Δ_ω by product)				
accounts		+0.0390	+0.0329	+0.0061
car loan		+0.0095	+0.0118	−0.0023
cards		+0.0407	+0.0444	−0.0037
mortgage		+0.0336	+0.0152	+0.0185
other loan		+0.0575	+0.0235	+0.0339
stud. loan		+0.0420	+0.0077	+0.0343
mean |gap|				0.0165

Table VI.b. Robustness: Welfare and Restitution Across the δ Band

Description: The table traces the two model-output quantities of the welfare accounting—foregone filer welfare and model-predicted restitution—across the 95% confidence band of the steepings of the income and self-advocacy gradients under monitoring report cards, the least sharply identified block of the index. The band is $\widehat{\delta} \pm 1.96\,\mathrm{SE}$ in log-odds on the fitted, company-clustered standard errors ($\mathrm{SE}_{q\omega} = 0.1143$, $\mathrm{SE}_{\theta\omega} = 0.0866$, giving half-widths of 0.2240 and 0.1698), with *both* interactions moved to the same edge—the widest and perfectly correlated case. At each band point the pipeline re-runs in full: both interactions are pinned at that edge and the product baselines, the firm-size shifter, and the merit gradients are refit by maximum likelihood, so the trade-offs among the blocks are respected rather than frozen. All figures are on the full sample ($N = 1{,}034{,}613$) at merit multiplier $m = 1$; foregone welfare scales linearly in m and the ratios are invariant to it. Restitution actually paid is data and does not move with the band; it is reported so that the predicted-to-actual ratio can serve as a fit check the foregone-welfare rows have no observed counterpart for.

	Lower edge $-1.96\,\mathrm{SE}$	Point estimate (MLE)	Upper edge $+1.96\,\mathrm{SE}$	Range rel. to centre
Monitoring interactions (log-odds), pinned at each band point				
$\delta_{q\omega}$	-0.1521	$+0.0719$	$+0.2960$	
$\delta_{\theta\omega}$	-0.1347	$+0.0351$	$+0.2048$	
Foregone filer welfare ($m = 1$)				
W^{foregone} (\$M)	152.5	148.5	144.5	-5.43%
ratio (foregone / paid)	0.864	0.842	0.819	
Restitution				
model-predicted, $\sum_i \widehat{\sigma}_i R_i$ (\$M)	181.2	176.4	171.5	-5.51%
actually paid (\$M)	176.4	176.4	176.4	invariant
predicted / actual	1.027	1.000	0.972	

VII The Awareness-Diffusion Model: Full Details

The structural model of Appendix VI prices the welfare lost by consumers who file a complaint and are denied. It is silent on the consumers who never file, because complaint records contain filed complaints only: a household that experiences a financial error and never learns that the CFPB channel exists leaves no trace in the data. That population is where the household's participation first-order condition (VI.1) lives, and this appendix develops the model that supplies it.

The awareness-diffusion model projects awareness of the CFPB complaint channel forward ten years, decomposed by income tier and by product, and returns for every product-income tier cell the annual stock of unaware problem-households, the flow of new filers the diffusion creates, and the cumulative restitution that flow generates. Those objects are then merged, cell by cell, with the structural model's pinned resolution probabilities and dollar stakes to produce non-filer welfare.

What the Model Has to Separate

The probability that a single problem episode results in a filed complaint is not one object but three, and the model's substantive content is that it keeps them apart:

$$\Pr(\text{file} \mid \text{problem with } p,\ \text{income tier } \theta,\ \text{month } t) \;=\; a_\theta(t) \cdot \zeta_p \cdot \lambda \;=\; a_\theta(t) \cdot f_p, \tag{VII.1}$$

with the product filing rate $f_p \equiv \lambda \zeta_p$. Here $a_\theta(t)$ is the share of income tier-θ households that know the CFPB channel exists, a permanent household-level stock diffusing logistically over a fixed household population; ζ_p is the product's complaint-relevance, how readily a problem with p is treated as a CFPB matter relative to the average product, normalized to mean one; and λ is the filing-given-awareness rate, the common conversion from "aware, and the problem is salient" to "complaint submitted." The income tier is indexed by $\theta \in \{0,1\}$, matching the structural model's above-/below-median tract income indicator. Three real-world properties motivate this factorization: (i) CFPB awareness is a household trait and does not vary by product (a household learns about the CFPB once and never unlearns); (ii) awareness is a permanent stock while problem-

households are a rotating annual flow (each year a fresh set of households experiences a problem with each product); (iii) filing entails being aware of the CFPB, having a problem, but also being willing to file conditional on awareness and problem (the separate λ). In order to separate these three elements we set-up a modified SIR model—Susceptible, Infectious, Recovered model, likening CFPB awareness to an infectious disease in a population—explained in detail in what follows.

The Model touches Three Populations

Population	Symbol	Nature	Source
Households in income tier θ	N_{θ}^{tot}	fixed stock	national household counts
Holders of product p in income tier θ	$N_{p,\theta}$	fixed stock	SCF 2022 product penetration
Problem episodes on p in income tier θ, per year	$N_{p,\theta}^{\text{prob}}$	annual flow	problem incidence $\times$ holders

Awareness is a property of the first population and is permanent: learned once, through any product, and never unlearned, applying to every product the household holds. Problem episodes are the third, and re-draw every year. Awareness is stored on the household stock; filings are read off the rotating problem flow.

The Model has Two Layers

Layer A — awareness diffusion. For income tier θ, the aware share $a_{\theta}(t)$—drawn from the fixed household population—diffuses logistically by word of mouth at the tier rate β_{θ}:

$$a_{\theta}(t+1) \;=\; a_{\theta}(t) \;+\; \beta_{\theta}\, a_{\theta}(t)\big(1-a_{\theta}(t)\big). \tag{VII.2}$$

Three features follow. There is *one curve per income tier*: all products in income tier θ read the same $a_{\theta}(t)$, which is the household-level spillover made structural, with the income split retained because both diffusion speed and starting awareness are likely to genuinely differ by income. It *diffuses over the whole household population*: the $(1-a_{\theta})$ term is the pool of still-unaware tier-θ

households and the a_θ term the aware households who transmit, and the relevant population is every household—whether or not it currently has a problem—because word of mouth is not confined to people with live grievances. And it is *permanent and monotone*, since $\beta_\theta \geq 0$: awareness, once acquired, is never lost.

Layer B — problem episodes. Each year $N^{\text{prob}}_{p,\theta}$ problem episodes arise in cell (p,θ), a stationary annual flow. Because that flow is a fresh sample of the income tier's households each year, and problem incidence is taken to be independent of awareness, the aware share among this year's problem-households equals the population share. This is the assumption that licenses reading the episode layer straight off Layer A:

$$S_{p,\theta}(t) = N^{\text{prob}}_{p,\theta}\big(1 - a_\theta(t)\big), \qquad I_{p,\theta}(t) = N^{\text{prob}}_{p,\theta}\, a_\theta(t), \tag{VII.3}$$

with S the *susceptible* unaware problem-households—the non-filers the welfare calculation needs—and I the *infectious* aware ones, when read in the key of a *SIR model*. An aware problem-household files at the product filing rate $f_p = \lambda\,\zeta_p$, so the monthly filing flow is $I_{p,\theta}(t)\, f_p/12$, cumulative filers accumulate that flow, and the restitution flow is the monthly filings scaled by the cell resolution probability $\pi_{p,\theta}$ and the mean award $\mu_{p,\theta}$:

$$\Delta R_{p,\theta}(t) = I_{p,\theta}(t)\frac{f_p}{12}, \qquad \Delta\text{Restitution}_{p,\theta}(t) = \mu_{p,\theta}\,\pi_{p,\theta}\, I_{p,\theta}(t)\frac{f_p}{12}. \tag{VII.4}$$

Laws of motion. Writing $N = S + I = N^{\text{prob}}_{p,\theta}$ for the stationary annual stock of problem-households, the headcounts obey, with a one-month Euler-forward step and β_θ a monthly rate,

$$\frac{dS}{dt} = -\beta_\theta \frac{SI}{N}, \qquad \frac{dI}{dt} = +\beta_\theta \frac{SI}{N}, \qquad \frac{dR}{dt} = \frac{f_p}{12} I, \tag{VII.5}$$

the standard epidemiological system of a SIR model with one change: the contagion rate is the *income tier* rate, common to all products in the tier, because the underlying awareness state is shared. Dividing the first two equations by N recovers the scalar awareness law of Layer A ex-

actly. Cumulative filings are a read-only accumulator—they do not enter the first two equations or the denominator—so $S + I = N$ is conserved and the awareness system is closed and convergent. Because $a_\theta(t)$ is a logistic S-curve it decelerates as it approaches one, so the restitution flow tends to a constant and cumulative restitution is S-shaped and then linear.

Data Sources

The model draws on four inputs outside the complaint records, assembled at the product-by-income tier level.

Household counts by income tier. National household counts split at the income median identified within the CFPB filers universe, from public national statistics: 85.9 million below-median and 45.4 million above-median households, roughly 131 million in total—signaling CFPB filers skewing towards higher household median income census tracts and further validating the need to split by income tier. This is the population over which awareness diffuses.

Product penetration. The share of households in each income tier holding each product, from the 2022 Survey of Consumer Finances. Penetration differs sharply by income tier and product—from near-universal deposit accounts in both tiers to consumer loans held by four to seven percent of households—and multiplying it by the tier household count gives the product-holder stock $N_{p,\theta}$.

Problem incidence. The share of product-holders experiencing a problem with that product in a twelve-month window, drawn from consumer-problem surveys and related external sources. Multiplying the share by the holder stock gives the annual problem flow $N^{\text{prob}}_{p,\theta}$. Table VII.a lists each share with the survey it is taken from and the question or statistic it is built on. These are the most assumption-laden inputs in the model, drawn from heterogeneous surveys, and the level of every welfare *stock* scales one for one with them—which is why we treat the unaware-population welfare figure as an upper bound rather than a point estimate, and emphasize the restitution *flows*, which are disciplined by the complaint cross-section instead.

Cell resolution probabilities and mean awards. The per-cell resolution probability $\pi_{p,\theta}$ and mean restitution $\mu_{p,\theta}$ used inside the restitution flow come from the 2019-2023 complaint data, averaged monthly.[31]

The tier diffusion rate. The diffusion rate β_θ is the growth rate of income tier-θ *awareness*. It is not observed directly, but it can be inferred from complaint growth. Total tier-θ complaints are $C_\theta^{\text{tot}}(t) = \sum_p N_{p,\theta}^{\text{prob}} a_\theta(t) f_p$, so if the problem flow and the filing rates are stationary then total tier complaints are proportional to awareness—they move one for one with it—and the growth rate of total tier complaints equals the growth rate of awareness. The rate is also computed from tier-aggregate complaint counts over 2019–2023, a five-year window rather than a single pair of years, and deliberately stopping short of 2024 so that the first projection year can be compared against data the calibration did not see. It gives $\beta_0 = 0.01195$ per month for below-median households (14.3% annualized) and $\beta_1 = 0.01156$ for above-median households (13.9%). The two tiers diffuse at nearly the same rate, so the logistic *widens* the absolute awareness gap for a long time before saturation eventually closes it.

Identification and Implementation

The terminal complaint cross-section gives one equation per cell,

$$C_{p,\theta} \;=\; N_{p,\theta}^{\text{prob}} \cdot a_\theta(0) \cdot \lambda \cdot \zeta_p. \tag{VII.6}$$

The unknowns $a_\theta(0)$, ζ_p, and λ enter *only as a product*, so their overall level is not separately identified from complaint counts. Our awareness-diffusion model separately pins $a_\theta(0)$ and ζ_p, but not λ. For this reason, we report awareness-diffusion model findings across $\lambda \in \{0.10, 0.25,$

[31] Note that when the diffusion output is merged into the welfare calculation, the resolution probability used is instead the structural model's pinned cell-level fitted value—the mean of the estimated per-complaint probability within each product-tier cell, with a product-level fallback for missing cells—so that the dollar value of an unaware household's foregone complaint is anchored to the estimated policy function rather than to a raw cell mean. Both the pinned probabilities and the per-product stakes R_p are held fixed across the baseline and every counterfactual, so cross-scenario comparisons are read off identical coordinates: only the diffusion-side counts vary.

0.33, 0.50, 0.75, 0.90} or at the reference value $\lambda = 0.5$ throughout. We identify $a_\theta(0)$ and ζ_p in three steps.

Step A — relative tier awareness. Define a cell's complaint intensity as complaints per problem-household, $r_{p,\theta} \equiv C_{p,\theta}/N^{\text{prob}}_{p,\theta} = a_\theta(0)\lambda\zeta_p$, and take its ratio across tiers *within* a product. Both λ and ζ_p cancel, leaving

$$\rho \equiv \frac{r_{p,0}}{r_{p,1}} = \frac{a_0(0)}{a_1(0)}, \tag{VII.7}$$

so relative income-tier awareness is identified from data alone—and it is *over-identified*, since each product yields an independent estimate. Under the model's premise that awareness is a universal household trait rather than a product-specific one we retain the pooled estimate (their geometric mean, since these are ratios) from Table VII.b. Below-median households start at roughly 77% of the above-median awareness level.

Step B — product weights. With relative awareness fixed at $\tilde{a}_\theta = (\rho, 1)$, the cross-product pattern of complaint intensity *within* a tier identifies the weights up to scale. Pooling both tiers and normalizing to mean one,

$$\zeta_p \propto \frac{\sum_\theta C_{p,\theta}}{\sum_\theta \tilde{a}_\theta N^{\text{prob}}_{p,\theta}}, \qquad \overline{\zeta_p} = 1. \tag{VII.8}$$

Deflating by $\tilde{a}_\theta$ is what makes this a measure of filing conditional on awareness rather than a measure of the two combined. The resulting weights are reported in Table VII.b and span an order of magnitude: 2.54 for mortgages, 1.26 for consumer loans, 1.08 for deposit accounts, 0.56 for cards, 0.33 for student loans, and 0.23 for auto loans. A mortgage problem is taken to the Bureau some eleven times more readily than an auto-loan problem, relative to how often each arises—the pattern one would expect from dollar stakes and from which products read as regulatory matters.

Step C — the seed. With the λ-free objects in hand, the seed is a direct back-out, one division per income tier rather than a global scaling. Summing the cell identity over the products of tier θ

and solving,

$$a_\theta(0) \;=\; \frac{C_\theta}{\lambda \sum_p \zeta_p N^{\text{prob}}_{p,\theta}}, \tag{VII.9}$$

clamped to $[10^{-6}, 0.999]$. Each tier is fit independently from its own complaint total, so the model reproduces the below-median and above-median complaint counts *exactly*. The weighting of the problem base by ζ_p is not optional: products do not convert problems into complaints at a common rate—complaint intensity spans roughly two orders of magnitude across products—so a raw unweighted sum would fit the tier total but badly mis-split it across products. A raw sum is exactly the special case $\zeta_p \equiv 1$, which the observed spread decisively rejects. Note that ρ itself does not enter the seed; it is used to deflate the weights in Step B, and otherwise remains the specification diagnostic of Step A.

The identified awareness paths. The simulator runs at monthly resolution over a 120-month horizon. For a given filing rate it backs out the two tier seeds via (VII.9), diffuses each tier curve forward through (VII.2), and reads the episode layer off it through (VII.3) and (VII.4). Each run writes one row per filing rate, product, and tier, carrying the awareness fraction at $t = 0$, 12, and 120 months; the unaware problem-household stock at the same three dates; cumulative filers at 12 and 120 months; the implied annual new-filer flow at the terminal date; cumulative restitution over one year and over ten years; and the product weight, alongside reference household and problem stocks. Table VII.c reports the simulated awareness and restitution across the λ grid. At the reference filing rate $\lambda = 0.5$ the seeds are $a_0(0) = 0.0075$ and $a_1(0) = 0.0122$—under one percent of below-median households against about one and a fifth percent of above-median households—rising over 120 months to 0.0306 and 0.0469. At $\lambda = 0.25$ the same paths run at twice the level (0.0150 and 0.0245) and at $\lambda = 0.90$ at roughly half (0.0042 and 0.0068), with the λ-invariant content being the roughly 1.6-to-one gap between tiers and the shape of the path.

Merging into Non-Filer Welfare

An unaware problem-household never files, so it forgoes the relief it would have obtained in expectation had it filed. Per cell and snapshot,

$$W^{\mathrm{nf}}(t) \;=\; S_{p,\theta}(t)\cdot\left[m\cdot\widehat{\sigma}_{p,\theta}\cdot R_p\right], \tag{VII.10}$$

where $S(t)$ is the unaware problem-household stock, R_p the per-product mean restitution, and $\widehat{\sigma}_{p,\theta}$ the *pinned* structural resolution probability for the cell. The division of labour is deliberate: the diffusion model changes *how many* households are unaware, while the dollar value of each unaware household's foregone complaint stays anchored at the structural model's estimate, which is the same pinned-merit convention used throughout the welfare accounting. The gross diffusion gain is the reduction in unaware welfare loss between two snapshots, reported in dollars—which scale with the merit multiplier—and as a fraction of the initial baseline, which does not.

Alongside the welfare stock we report the restitution actually delivered to newly aware households who file. This is the quantity we emphasize, because it is disciplined by the complaint cross-section rather than by the problem-incidence shares, and because it is directly comparable to the structural model's expected-relief figures: both are dollars of relief on the same complaint universe. It is also invariant to the merit multiplier, being an observed-dollar flow rather than a welfare unit.

Baseline Results

The unaware stock. At $m = 1$ and $\lambda = 0.5$ the base-year unaware problem-household population carries roughly \$8.52 billion a year of foregone expected relief, moving between \$8.44 billion and \$8.55 billion across the interquartile filing range. This establishes the order of magnitude of the participation margin against the roughly \$148 million of foregone filer welfare, but it scales one for one with the problem-incidence shares and with both grid parameters.

What organic diffusion buys. Ten years of diffusion at the current complaint data-implied rates reduces that stock to $8.28 billion at $\lambda = 0.5$—a gross gain of $238 million a year, or 2.79% of the base-year level, as reported in Table 9. Nonetheless, organic diffusion delivers about $44 million of restitution in the first projection year of 2024 (close to the actual 2024 restitution of $42.3 million, which validates model fit) and $888 million cumulatively over ten years at $\lambda = 0.5$ (see Table VII.d for a detailed breakdowns by product) ($875 million and $892 million at the λ interquartile endpoints), reaching an annual rate of about $88 million by the terminal year. The ten-year cumulative figure is roughly five times the $176 million the complaint system actually paid out over 2015–2024—from organic diffusion alone, with no intervention. Its incidence is close to even, with 48.3% accruing to below-median tracts at $\lambda = 0.5$, which reflects the offsetting facts that below-median households are the larger population but start from the lower awareness seed.

Counterfactual Scenarios

Three scenarios—CF6, CF7 and CF8, continuing the numbering of the five resolution counterfactuals of Appendix VI—are run across the full filing-rate grid, each producing the identical per-cell schema as the baseline. Results are reported in Table 9 and Figure VII.1 at $\lambda = 0.5$.

CF6 — close the income–awareness gap. Below-median households start with the same awareness as above-median households, $a_0(0) := a_1(0)$ at every filing rate, with the diffusion rate unchanged. This is a one-time information campaign that closes the income gap in starting awareness without altering subsequent word-of-mouth dynamics. It raises ten-year cumulative restitution from $888 million to $1.151 billion at $\lambda = 0.5$, a gain of 29.7%, and cuts the terminal unaware-welfare stock by 1.2%. Its incidence is progressive by construction, since only the below-median seed moves: the below-median share of cumulative restitution rises from 48.3% at baseline to 60.1%.

CF7 — double the diffusion speed. Both tiers' diffusion rates are doubled, with the seeds unchanged—an ongoing outreach effort that doubles the word-of-mouth amplification rate while

leaving starting points where the data put them. Over ten years this is the larger lever, raising cumulative restitution to \$2.094 billion (+135.9% at $\lambda = 0.5$) and cutting the terminal unaware stock by 10.0%. But it is close to distributionally neutral—the below-median share is 49.6%—because it scales both tiers equally, and over a single year it is much the weaker of the two, delivering only +6.7% of restitution against CF6's +29.7%.

CF8 — combined. Both interventions together raise ten-year cumulative restitution to \$2.691 billion, +203.2% at $\lambda = 0.5$, cut the terminal unaware stock by 13.8%, and direct 60.8% of the cumulative total to below-median tracts. Over twelve months the combined figure is +38.5%.

Table VII.a. Problem-Incidence Shares and Their Sources

Description: The table documents the per-product problem-incidence share s_p—the fraction of households holding product p that experience a problem with it in a twelve-month window—together with the external survey each share is drawn from and the question or statistic that defines it. "Holders" is the stock of households holding the product, obtained by applying 2022 Survey of Consumer Finances penetration rates to the national household count in each income tier and summing the two tiers; problem "episodes" is s_p times that stock, i.e. the annual flow $N^{\text{prob}}_{p,\theta}$ summed over tiers (the incidence share is applied identically to both income tiers). Other loans carry no survey of their own and are assigned the unweighted mean of the four loan types that do—cards, mortgage, auto loan and student loan.

Product	Share s_p	Holders (M)	Episodes (M/yr)	Source	Question / basis
accounts	0.0800	124.3	9.94	J.D. Power Survey 2024	Have switched banks over the previous year
car loan	0.2570	44.8	11.50	J.D. Power Survey 2025	Average overall satisfaction among financially healthy automotive loan customers is 743 on a 1,000-point scale; the share is its complement, $1 - 0.743$
cards	0.1910	119.8	22.87	SHED 2025	Credit and debit card fraud
mortgage	0.0346	52.9	1.83	National Survey of Mortgage Originations, waves 25 onwards (2020–)	X28A — "Overall, how satisfied are you with the mortgage lender/broker you used"; share answering "not at all"
other loan	0.1806	6.7	1.20	Average of the card, mortgage, car and student loan shares	-
stud. loan	0.2400	28.5	6.84	TICAS / Data for Progress 2025	Servicer has at some point provided them with inaccurate information

Table VII.b. Identified Awareness Structure from the Complaint Cross-Section

Description: The table reports the two objects identified from the terminal complaint cross-section. $\rho_p = r_{p,0}/r_{p,1}$ is the within-product ratio of complaint intensity across income tiers; both λ and ζ_p cancel from it, so each product delivers an *independent* estimate of the same quantity, the relative tier awareness $a_0(0)/a_1(0)$. Under the model's premise that awareness is a universal household trait rather than a product-specific one we retain the pooled estimate (their geometric mean, since these are ratios). ζ_p is the product complaint-relevance weight of equation (VII.8), normalized to mean one across products: $\zeta_p > 1$ means a problem with p is taken to the Bureau more readily than the average product, relative to how often such problems arise. Complaint counts are the terminal-year cell totals used in the identification, and problem episodes are carried over from Table VII.a. Products are ordered by ζ_p.

	Complaints		Problem episodes	ρ_p	ζ_p
Product	$\theta = 0$	$\theta = 1$	(M per year)	relative awareness	complaint relevance
mortgage	11,019	11,841	1.83	1.164	2.544
other loan	4,105	3,142	1.20	1.045	1.264
accounts	22,938	26,942	9.94	0.487	1.076
cards	26,981	32,468	22.87	0.475	0.557
stud. loan	5,004	5,629	6.84	0.696	0.327
car loan	7,004	5,796	11.50	1.068	0.232
Total / pooled	77,051	85,818	54.19	0.771	1.000

Table VII.c. Simulated Awareness and Restitution Across the λ Grid

Description: The table reports the baseline simulation at each value of the filing-given-awareness rate λ. The seeds $a_\theta(0)$ are backed out from each tier's own complaint total by equation (VII.9), so they fall as λ rises—the two enter the cell identity only as a product, and a higher conversion rate requires fewer aware households to generate the same observed complaints. Terminal awareness $a_\theta(120)$ is the seed diffused forward 120 months at the tier rate through equation (VII.2). Restitution is the cumulative flow to newly aware households who file, over one year and over ten. The last column is the below-median share of the ten-year total.

	Seed $a_\theta(0)$		Terminal $a_\theta(120)$		Restitution ($M)	
λ	$\theta = 0$	$\theta = 1$	$\theta = 0$	$\theta = 1$	1 year	10 years
0.10	0.0376	0.0612	0.1399	0.2060	44	840.0
0.25	0.0150	0.0245	0.0597	0.0907	44	875.0
0.33	0.0114	0.0185	0.0458	0.0698	44	881.0
0.50	0.0075	0.0122	0.0306	0.0469	44	887.6
0.75	0.0050	0.0082	0.0205	0.0317	44	891.9
0.90	0.0042	0.0068	0.0172	0.0265	44	893.3

Table VII.d. Ten-Year Simulation Output by Product and Income Tier ($\lambda = 0.50$)

Description: The table reports the per-cell state of the baseline simulation after 120 months at the reference filing rate. $S(120)$ is the stock of problem-households still unaware of the complaint channel, the population that carries the non-filer welfare loss; $R(120)$ is the cumulative count of problem-households that have filed over the decade; and the last column is the cumulative restitution their filings generate, $\mu_{p,\theta}\pi_{p,\theta}$ applied to the monthly filing flow.

	Below median ($\theta = 0$)			Above median ($\theta = 1$)		
Product	$S(120)$ (M)	$R(120)$ (000s)	Restit. ($M)	$S(120)$ (M)	$R(120)$ (000s)	Restit. ($M)
accounts	6.13	556.9	246.5	3.45	501.5	251.2
cards	14.11	664.1	151.7	7.93	597.3	146.0
car loan	5.92	116.0	5.6	5.14	161.3	11.7
mortgage	0.79	169.6	10.1	0.97	333.6	25.6
other loan	0.65	69.2	4.3	0.51	87.1	7.7
stud. loan	3.72	102.6	10.1	2.86	126.4	17.1
Total	31.31	1,678.4	428.3	20.87	1,807.1	459.3

Figure VII.1. Simulated Awareness Diffusion, $\lambda = 0.50$

Description: Awareness paths from equation (VII.2) at the reference filing rate $\lambda = 0.50$, in percent of tier households. Panel (a) shows the two baseline income tiers. Because they start apart and diffuse at almost the same rate ($\beta_0 = 0.0119$ against $\beta_1 = 0.0116$ per month), the logistic *widens* the absolute gap over the decade—from 0.47 to 1.64 percentage points—even though the ratio narrows slightly; saturation would eventually close it, but far beyond this horizon. Panel (b) shows the below-median income tier under each awareness counterfactual. CF6 shifts the initial level and pays immediately, but the path is otherwise parallel to baseline; CF7 leaves the starting point alone but bends the curve, so it is worth less than CF6 in year one and far more by year ten.

