The “Paper Deal—Real Deal” Gap in Consumer Markets: Evidence from A Field Experiment

Meirav Furth-Matzkin[[1]](#footnote-1)\*

*The Gap Theory suggests that in competitive markets, sellers will often deviate from clear and unconditional standard form contracts to please consumers. This Article investigates this theory through the case study of product returns. In a large-scale field experiment, a team of testers (auditors) was trained to follow a uniform bargaining script and was sent to return clothing items without receipts (despite a formal receipt requirement) to ninety-five retail stores in Chicago. The results show that, across a wide range of stores, a significant minority of retailers strategically departs from the formal, standardized policy in favor of consumers. In addition, sellers are almost twice as likely to depart from their contractual arrangements once consumers complain. These results are consistent with the Gap Theory in suggesting that sellers may deliberately adopt rigid policies on paper with the aim of screening out certain consumers, while tailoring the “standard” agreement in favor of others. Yet, at the same time, the results qualify the Gap Theory by showing that most sellers insist on enforcing their rigid policy requirements, and that this insistence largely persists even after consumers complain. These findings, I argue, have important policy implications. Gap Theory proponents have suggested that in competitive markets, policymakers should be reluctant to intervene in consumer transactions because reputation will encourage sellers to depart from clear and unconditional contractual arrangements to please consumers, while the formal terms will protect sellers from potential abuse. These findings suggest that regulation may be warranted because most sellers enforce their policies to the letter, even when dealing with non-opportunistic consumers. Moreover, building on new evidence and on insights from the social sciences, this Article proposes that even sellers’ limited departures from their formal conditions might actually harm consumers and lead to regressive distributive outcomes.*

# Introduction

Most of our everyday transactions are governed by standard form contracts. Virtually every firm selling goods or services includes in its sale contracts certain standard provisions that dictate whether and when a good can be returned, when and how to make payments, whether charges are imposed for services beyond those originally contracted for, as well as various other elements of the sales relationship. Although these clear and unconditional boilerplate provisions considerably facilitate transactions,[[2]](#footnote-2) scholars and commentators have expressed concern that such terms might be excessively rigid and even harmful to consumers.[[3]](#footnote-3)

In response to these concerns, several scholars have suggested that, at least in competitive markets, sellers’ implementation of ostensibly rigid and unconditional terms may be more lenient and flexible in practice. As these scholars argue, while the existence of rigid terms on paper enables sellers to fend off buyers likely to take advantage of a more lenient policy to extract benefits that the seller did not intend to offer, reputational considerations will constrain sellers from enforcing their contracts to the letter vis-à-vis good faith customers.[[4]](#footnote-4) I refer to this approach as the Gap Theory, because it suggests that sellers sensitive to reputational consequences will depart from their formal policies in favor of consumers, thereby generating a gap between the “paper deal”—the written contract governing sellers’ relations with buyers, and the “real deal”—the contract in action.[[5]](#footnote-5)

While the Gap Theory has attracted significant scholarly attention, academics continue to debate both its descriptive accuracy and its normative consequences. In particular, some have questioned the actual ability of competitive forces, and reputation in particular, to deter sellers from adhering to the four corners of the agreement.[[6]](#footnote-6)

This debate has significant policy implications. If reputational constraints are sufficient to discipline sellers and prevent them from enforcing their contracts to the letter when such enforcement is burdensome for consumers, policymakers and courts may not need to intervene in the contents of these agreements in order to protect consumers.[[7]](#footnote-7) However, if such reputational considerations do not effectively constrain sellers’ decisions, then regulators and courts may need to intervene in the contents of consumer transactions to ensure that consumers are adequately protected.[[8]](#footnote-8) Despite the significant practical implications of the debate, there has nonetheless been little systematic empirical research on how sellers actually enforce rigid contractual provisions.[[9]](#footnote-9)

This Article represents an important step towards mitigating this deficiency. Using product returns as a first test case, it reports the results of an original, large-scale field experiment covering nearly one hundred retail stores in Chicago. The experiment examines whether and when retailers allow their store clerks discretion to depart from rigid, unconditional return policy terms in favor of consumers.

The experiment uses an audit technique, in which pairs of testers were trained to follow a uniform bargaining script and were then sent to return clothing items without receipts (despite clear receipt requirements in the policy terms) to 95 clothing retail stores in Chicago. To explore the effect of consumer complaints on sellers’ departure decisions, testers denied a refund were instructed to complain by asking to speak to the store’s management.

The findings reveal that a significant minority of sellers (about 20%) departed from their formal receipt requirement in favor of consumers in response to their initial requests, and almost twice as many sellers departed from this requirement once consumers complained. At the same time, the vast majority of sellers (almost 80% at the initial stage and 63% at the final, post-complaining stage) strictly adhered to their policies’ requirements, refusing to offer any concessions to consumers. This held true regardless of how luxurious or established the store was and regardless of whether the store was local or a part of a chain. These findings qualify the plausible Gap Theory prediction that a majority of sellers will depart from their rigid, formal requirements in favor of consumers, suggesting that, at least in the context studied here, most sellers are not flexible.

The results also show that some terms are stickier than others, with sellers very rarely departing from stickier terms to consumers’ advantage. Sellers were reluctant to apply pro-consumer gaps in contexts of cases of unequivocally rigid terms completely denying refunds, where the Gap Theory would predict that gaps would be the most prevalent.[[10]](#footnote-10) Importantly, these are also the contexts where gaps are arguably needed the most, because these harsher terms significantly limit consumers’ withdrawal rights.

These findings, I argue, carry significant legal ramifications. Many scholars have argued that sellers’ departures from their formal contracts “renders legal intervention less necessary.”[[11]](#footnote-11) According to Gap Theory supporters, “courts would do well to take a hard line in enforcing the terms” of consumer contracts (in the absence of evidence of fraud).[[12]](#footnote-12) Yet, the findings reveal that reputational constraints on businesses are not always sufficient to prevent sellers’ enforcement of strict and unconditional contract terms. Consequently, the gap between the “paper deal” and the “real deal” does not necessarily obviate the need for regulation of contract terms.

This Article also makes a more nuanced argument against the notion that the gap can be relied on to discipline sellers. It suggests that even when sellers exercise tailored forgiveness, the discrepancies between the standardized terms and their actual implementation might generate distortions and regressive distributive effects.

These possible adverse effects would arise because gaps could generate distortions in consumers’ decision-making process. The findings reveal that stores vary significantly in the extent to which they are willing to depart from their standardized policies in practice. While more luxurious, experienced, and chain stores are more likely to exhibit tailored forgiveness *ex post*, substantial unexplained variation remains even after store characteristics are controlled for. This unexplained variation might harm consumers’ ability to distinguish between stores that strictly adhere to their formal return policy and those that offer better terms in practice. Consequently, consumers might make poor return decisions. They might be discouraged from trying to make returns to stores that would be likely to depart from their unconditional paper policies in their favor, or inefficiently attempt to make returns to stores that would not budge.

These distortions are aggravated because of sellers’ use of complaint-based segmentation of consumers. As the findings reveal, store employees are significantly more likely to exercise discretionary, tailored forgiveness once consumers complain.[[13]](#footnote-13) Of course, sellers may apply complaint-based screening benevolently, using consumer assertiveness as a proxy for the merits of the claim or the value of the concession to the consumer. Yet, this Article suggests that such complaint-based segmentation might be harmful to consumers. Accumulating empirical evidence suggests that consumers tend to be contract formalists, with most believing that whatever the contract says is the final word.[[14]](#footnote-14) They are thus unlikely to complain when the seller refers them to the policy’s clear and unconditional language, even if they have a meritorious claim, and possibly even when their benefits from the seller’s more lenient accommodation exceed the costs to the seller of granting it.

Furthermore, this complaint-based mechanism is likely to have regressive distributional effects. Quantitative evidence and original interviews with store clerks indicate that lower income consumers as well as minority group members are less likely to complain than are higher income consumers or those belonging to majority groups.[[15]](#footnote-15) The former will consequently cross-subsidize the insistent complainers who benefit from the gap. Furthermore, store clerks with discretion to deviate from the formal policy on the ground may apply their discretion inconsistently and to the disadvantage of lower income consumers and minorities. In particular, preliminary evidence from a field experiment that I have recently conducted indicates that sellers are significantly more likely to exercise tailored forgiveness when facing white consumers than when facing black consumers, even when they use identical bargaining scripts. The troubling conclusion is that the gap most likely helps those who are already better off, while harming less empowered consumers.

The Article proceeds as follows. Section I provides the background and motivation for the study. It introduces the Gap Theory and the debate surrounding its descriptive validity and welfare implications. It explains the importance of shedding empirical light on the debate in view of the opposing policy prescriptions that emanate from each side. It then describes the paradigmatic example of product returns and the legal framework governing consumers’ withdrawal rights. Section II presents the sample and design of the field experiment. Section III reports the results. Section IV discusses the implications of the findings. It suggests that the results raise some serious concerns regarding the outcomes of the interaction between the formal agreement, sellers’ on-the-ground practices, and consumer psychology.

# I. Background and Motivation

## The Gap Theory and Gap Skepticism

The Gap Theory argues that clear and unconditional contract terms, although often *seemingly* inefficient, may allow sellers to exercise tailored discretion on-the-ground, while segmenting between opportunistic (or low-value consumers) and non-opportunistic (or high-value) consumers.[[16]](#footnote-16)

Under this theory, sellers adopt these terms in light of the inherent asymmetry between sellers and consumers. While sellers are constrained by reputational forces, buyers typically are not. Consumers may therefore abuse a seller’s policy term without incurring reputational costs, while sellers will not be able to insist on enforcing the policy to its fullest extent against the consumer without harming their reputations in the market.[[17]](#footnote-17)

The Gap Theory reflects a belief in the ability of competitive market forces, and reputation in particular, to discipline sellers. According to its proponents, reputational constraints will ensure that sellers selectively enforce rigid, bright-line terms *only* against low-value consumers (defined as consumers whose benefits from the concession do not exceed the cost to the seller of granting it), while departing from the formal agreement to the benefit of the remaining, high-value consumers (as consumers whose benefits from the concession exceed the cost to the seller of granting it).[[18]](#footnote-18)

Sellers’ ability to adopt clear and unconditional standard-form terms on paper is important under this view, as courts are ill-equipped to distinguish between different types of buyers.[[19]](#footnote-19) Given the courts’ limited ability to identify whether the buyer values the concession at more than its cost to the seller, sellers are expected to offer contracts lacking these concessions. However, sellers constrained by reputational forces will concurrently have an informal policy of allowing concessions not required by the contract when encountering high-value consumers. This descriptive theory therefore also has a normative prong: When the value of the benefit to consumers is observable to sellers but non-verifiable, or verifiable only at a high cost to courts, the argument is that sellers should be allowed to use rigid, bright-line terms, since these terms will allow them to behave efficiently (on account of their discretion to forgive).[[20]](#footnote-20)

Sellers’ ability to adopt clear and unconditional terms on paper is important under this view, as courts are ill-equipped to identify the exact circumstances in a given case.[[21]](#footnote-21) Sellers are therefore expected to include rigid, rule-like terms in their contracts. However, sellers will concurrently have an on-the-ground policy of allowing concessions not required by the contract. This descriptive theory therefore also has a normative prong: When consumer misbehavior is observable to sellers but non-verifiable, or verifiable only at a high cost to courts, strict bright-line terms are both efficient and good for (the non-opportunistic) consumers.[[22]](#footnote-22)

Importantly, according to the Gap Theory, sellers’ willingness to depart from their contracts is not limited to interactions with repeat customers. Sellers that are repeat players in the market, with expectations of doing business with other consumers, may be discouraged from enforcing their contracts to the letter even if these sellers do not expect to have further dealings with the particular consumer.[[23]](#footnote-23)

Although the Gap Theory has gained considerable traction, some scholars have raised concerns that competitive forces might not adequately deter sellers from enforcing their strict, rule-like terms. These scholars rely mainly on the manifold evidence that informational flows in consumer markets are far from perfect, because reputational information is neither reliable nor accurate.[[24]](#footnote-24)

This debate has important policy implications. Gap Theory proponent Jason Johnston has argued that “courts should support the standard-form discretionary benefits/forgiveness market equilibrium.”[[25]](#footnote-25) Lucian Bebchuk and Richard Posner have similarly suggested that “courts would do well to take a hard line in enforcing the terms” of standardized consumer contracts in the absence of evidence of fraud.[[26]](#footnote-26) At the same time, those who point to reputational failure highlight “the centrality of the law to the future of the marketplace.”[[27]](#footnote-27) Indeed, if sellers insist on adhering to the contract even when it is not socially desirable to do so, regulation of the contents of standardized agreements may be warranted after all.

In view of the conflicting policy prescriptions stemming from both sides of the debate, exploring them in the field is of the utmost importance. Yet, notwithstanding the practical implications of this debate, empirical investigation into the gap is surprisingly lacking. This Article uses the test case of product returns to shed light on the gap theory.

The case study of product returns is particularly suitable for scrutiny because the gap literature has used return policies as the poster child example for its hypothesis[[28]](#footnote-28) without any meaningful empirical investigation into sellers’ return practices on the ground.[[29]](#footnote-29) In addition, despite the substantial economic significance of product returns to both consumers and sellers, we know far too little about the contents and actual implementation of retailers’ return policies. The next section provides background to the study by briefly presenting the legal framework governing product returns.

## The Case Study: Retail Stores’ Return Policies

In the U.S. alone, consumers spend more than $10 trillion a year in retail stores.[[30]](#footnote-30) In a significant number of cases, consumers who purchase goods or services end up regretting their purchases after the fact.[[31]](#footnote-31) In 2016, for example, Americans returned $260 billion in merchandise (or eight percent of all purchases made in the United States) to retailers.[[32]](#footnote-32)

The right to withdraw from consumer transactions is therefore an important consumer right.[[33]](#footnote-33) It enables consumers to inspect the product and gain information that they are unable to obtain before the purchase, and allows them to cancel the transaction if they realize that it is no longer desirable or necessary. Yet, despite the enormous economic significance of the right to withdraw from consumer transactions, consumers are not in most cases[[34]](#footnote-34) entitled to return non-defective goods to the seller under current U.S. law.[[35]](#footnote-35) Consumers cannot return items to stores unless the contract with the seller stipulates otherwise.

Against this legal backdrop, scholars and policymakers continuously debate the desirability of regulating consumers’ rights to cancel transactions. While some have proposed adopting either a mandatory or a default right to withdraw,[[36]](#footnote-36) others believe that statutory intervention may not be warranted due to market incentives already in place.[[37]](#footnote-37) That is, sellers will be incentivized to offer efficient terms—if not on paper, then in practice—so long as their reputation is at stake.

Supporters of minimal intervention rely on the liberal return policies adopted by many sellers as standard commercial practice.[[38]](#footnote-38) Yet, in recent years, stores have begun imposing more and more restrictions on consumer returns, to prevent consumers from being able to take advantage of retailers’ generous policies to obtain free product rentals.[[39]](#footnote-39) These shifts in firms’ policies have spurred consumer outrage,[[40]](#footnote-40) reviving the debate over the need to regulate consumers’ withdrawal rights.[[41]](#footnote-41)

Even though underlying this debate are competing views of whether, and to what extent, market forces ensure that sellers offer favorable return policies to consumers—if not on paper, then in practice—empirical studies of retail stores’ return policies and practices are surprisingly scarce.[[42]](#footnote-42) This Article presents the first comprehensive empirical account of on-the-ground retailers’ return practices. Exploring the relationships between contractual language and the contract in action, this Article illustrates how these discrepancies shape the interactions between sellers and consumers and the resulting market outcomes.

# II. Sample & Design

To test whether there is a gap between retailers’ return policy terms and their actual enforcement vis-à-vis consumers, I conducted a field experiment.[[43]](#footnote-43) The study uses an audit technique, in which testers were sent to return clothing items to retail stores located in Chicago.[[44]](#footnote-44) The following sections present the sample and design of the experiment. Section A explains how the stores were selected, and presents their descriptive statistics. Section B describes the experimental design. Section C describes the training and the measures taken to ensure experimental control. Section D explains how the formal return policies were classified, and Section E explains how pro-consumer gaps were defined and measured.

## Store Selection and Descriptive Statistics

The stores were selected from the ReferenceUSA and Hoover’s Company Directories’ databases according to the following selection criteria: They all had a Chicago location, and they all required a receipt—either for any return or exchange or for a refund. The study focuses on the enforcement of “receipt required” return policies because they are both very common and strongly disliked by many consumers, who perceive the receipt requirement as an unnecessary burden.[[45]](#footnote-45) In addition, dollar stores were excluded from the sample in order to make the analysis more tractable. To keep the items’ prices fixed across stores at between $20 to $30, designer stores that did not offer items for $30 or less were also excluded from the sample.[[46]](#footnote-46)

After applying the mentioned selection criteria and discarding tests (due to inconsistent outcomes, deviations from the script, or scheduling difficulties), the final sample includes 190 audits from 95 retailers (two audits per store).

For each of the sampled stores, I collected information on basic company characteristics such as annual revenues (for the year of 2018) and age (defined as 2019 minus the year of establishment). I also collected data on whether the store is local (defined as stores with no more than two locations, both in Illinois) or part of a chain. I obtained most data from Bloomberg and Hoover’s Company Directories. Finally, as a measure of store prestige, I collected data on the median prices of all clothing and accessory items listed on each store’s website.[[47]](#footnote-47)

The stores’ mean annual revenues is $434 million (SD = 2,311,000). The average company age is 52 years (SD = 42).[[48]](#footnote-48) Seventy-one percent of the stores in the sample are national retail chains, and 29% are local stores.[[49]](#footnote-49) The average median price of items (listed on the stores’ websites) is $68 (SD = 70).

## Experimental Design

For the purposes of the audit study, I hired and trained a team of six testers—all White American females of approximately the same age. A pair of testers audited each store, and the composition of pairs varied from audit to audit.[[50]](#footnote-50) The testers all wore casual clothing and followed a memorized script in their interactions with store clerks.

Data collection took place between March and April 2019, well after the holiday season, to avoid potential changes in stores’ return policies that typically occur during that season.[[51]](#footnote-51) Both testers in a pair visited the same store, usually within a few days of one another. After leaving each store, all testers filled out a detailed report, in which they described the outcomes of the attempted returns.

In each store, testers attempted to return a clothing item that had been purchased in advance[[52]](#footnote-52) in its original packaging and condition, with tags attached, but *without the receipt*.[[53]](#footnote-53) They were instructed to wait in line until a store clerk became available and asked them what he or she could do for them. The testers were then instructed to say that they wanted to return the item, and to put the item in its original packaging on the counter. If the store clerk asked them why they wanted to return the item, testers would say that they realized they did not need it.[[54]](#footnote-54) If asked for the receipt, testers would answer that they thought they had the receipt with them, but, after looking for it, they appeared to have lost it. They were then instructed to await the store clerk’s response. If the store clerk agreed to provide a refund, testers were instructed to accept the refund, thank the clerk, and leave the store.

If, however, they were denied the return or were offered anything other than a refund (e.g., exchange or store credit), testers were instructed to ask to speak to a manager, thereby complaining. If the store clerk refused to call a manager or identified as the manager, or if a manager came to see them, testers asked once again for a refund. If still denied a refund, testers would thank the clerk or manager and leave the store.

The testers’ bargaining script was designed to test the interaction between consumer assertiveness and the gap. The study predicted that sellers would be more likely to exercise leniency once consumers insisted and complained, because sellers may use consumers’ complaints or assertiveness as a proxy for the value that consumers attach to the particular concession.

The purpose of the study was to identify *systematic*, rather than sporadic deviations from the paper policy. Yet, it is possible that in some of the stores, store clerks mistakenly deviated from the formal policy, and not as a result of a systematic store policy allowing for store clerk discretion on the ground. To address the concern about random divergences from the paper contract, each store was audited by *two* testers, and the analysis includes only those stores in which both testers obtained identical return outcomes.[[55]](#footnote-55)

Admittedly, it could be the case that even when the two testers obtained identical return outcomes, these outcomes were the result of store clerk error. However, this design gives me confidence that most of the observed deviations were the result of managerial policy allowing clerks to depart from the formal policy in favor of consumers. This methodological approach also mitigates the concern that, despite attempts to control for uniform tester performance and demographics (as described below), differences in testers’ bargaining behavior or other characteristics meaningfully influenced the results.

## Training and Uniformity

Testers were recruited and trained in January and February 2019. In order to minimize the possibility of non-uniform bargaining, particular attention was paid to issues of experimental control. A major goal of the study was to choose uniform testers and to train them to behave in a standardized manner. Testers were chosen to satisfy the following criteria for uniformity:

1. *Race:* All testers were white;
2. *Gender:* All testers were female;
3. *Age*: All testers were twenty-two to thirty-years-old;
4. *Education*: All testers had between one and four years of college education;
5. *Dress*: All testers wore casual attire during the audits: jeans, t-shirt, and minimal make-up;
6. *Contact Information:* If asked for their names when returning the item, testers would use one of two fake names.[[56]](#footnote-56) If asked for an identification card, testers would say that they did not have one with them.[[57]](#footnote-57)

All testers attended a training session with the author and research team before visiting the stores. The training included not only memorizing the tester script, but also participating in mock negotiations designed to help testers gain confidence and learn how to negotiate and answer questions uniformly. The script anticipated that store clerks would ask questions and gave testers a list of contingent responses to questions that might be asked. Testers were accompanied by project coordinators to the stores. The coordinators ensured that testers were following the script and accurately reporting the results. In addition to ensuring bargaining uniformity, the returned items’ prices were also kept constant, at between $20 to $30.

## Classification of Return Policies

In order to study variations between stores’ formal policies and their actual return practices, a database was created of the audited stores’ return policies, as they appeared on the stores’ websites, in-store signs, and receipts.[[58]](#footnote-58) Stores were subsequently classified by two independent research assistants as having harsh, moderate or lenient paper policies.[[59]](#footnote-59)

Stores are classified as having *harsh policies* (on paper)if they do not allow consumers to return items for a refund, and require receipts for any exchange or return. They are classified as having *moderate policies* if they allow consumers to return items for a refund conditional on showing receipts, and require receipts for *any* exchange or return. Finally, they are classified as having *lenient policies* if they allow consumers to return items for a refund conditional on showing receipts, while allowing consumers to return items for exchanges or store credits even without receipts.[[60]](#footnote-60) Table \_\_\_ summarizes this classification method.

*Table 4. Paper Policy Types (n = 95)*

|  |  |  |
| --- | --- | --- |
|  | Are refunds allowed? | Are non-receipted returns (for store credit or exchange) accepted? |
| Harsh Policy Stores  (n =23) | No | No |
| Moderate Policy Stores  (n = 33) | Yes | No |
| Lenient Policy Stores  (n = 39) | Yes | Yes |

The next table reports the basic company characteristics for each type of policy store. I report mean annual revenues (for the year of 2018) and age (defined as 2019 minus the year of establishment). I also report the proportion of chain stores and luxury stores in each category.

*Table \_\_. Summary Statistics*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean Revenues (in millions of dollars)  (SD) | Mean Age  (SD) | Mean of Median Prices  (SD) | Chain |
| Harsh Policy Stores  (n =23) | 0.26  (SD = 0.38) | 21  (14) | 9% | 0.04  (0.21) |
| Moderate Policy Stores  (n = 33) | 751  (SD = 3285) | 46  (32) | 27% | 0.85  (0.36) |
| Lenient Policy Stores  (n = 39) | 199  (SD = 1024) | 71  (48) | 18% | 0.97  (0.16) |

## Analysis of Gaps

The study explores the prevalence of pro-consumer gaps, defined as any of the following deviations from the paper policy in favor of consumers:

1. *Among both Harsh and Moderate Policy Stores (that require receipts for any exchange or return)*—a pro-consumer gap was observed whenever testers were offered any concession—in the form of refund, exchange or store credit—despite their failure to show a receipt.
2. *Among the* *Lenient Policy Stores (that require receipts for a refund, while allowing for non-receipted exchanges or store credits)*—a pro-consumer gap was observed whenever the testers received a refund despite their failure to show a receipt. In these stores, I also document whether store clerks departed from the formal policy to consumers’ detriment, by completely denying the return, while failing to offer store credits or exchanges.

Table \_\_ describes this analysis.

*Table 5: Operationalization of Gaps*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Store Type | Paper Policy | Return Outcomes: Is there a Gap? | | | |
|  |  | **Is the Return Denied?** | **Is Exchange/Store Credit Offered?** | **Is Refund**  **Offered?** |
| Harsh | No Refund + Receipt required for *any* return | Yes = **No Gap** | Yes = **Pro-Consumer Gap** | Yes = **Pro-Consumer Gap** |
| Moderate | Refund Allowed + Receipt required for *any* return |
| Lenient | Refund Allowed + Receipt required for *refunds* | Yes = **Pro-Seller Gap** | Yes = **No Gap** |

# III. Results

The results as reported by the testers in their post-audit surveys provide a rich database for investigating the “paper deal—real deal” gap. The next three sections present the findings of the field experiment. Section A reports on the prevalence of pro-consumer gaps at the initial stage, after testers had requested a refund but before they complained. Section B turns to report the prevalence of pro-consumer gaps at the final stage, after testers had complained. Section C reports the results of multivariate analyses aimed at testing the relationships between several store characteristics and the gap.

### Initial Return Outcomes and the Prevalence of Pro-Consumer Gaps

Figure 6 describes what happened when testers initially asked for a refund—overall and across the different types of stores.

*Figure 6. Return Outcomes at the Initial Stage.*

*Figure \_\_.* Return outcomes at the initial (pre-complaining) stage. The upper bar represents return outcomes overall (n = 95), while the lower three bars represent the return outcomes across stores with different policies.

As figure \_\_ illustrates, the formal terms of the contract had a strong and significant effect on testers’ initial return outcomes. Stores with lenient return policies (formally allowing non-receipted returns for store credits or exchanges) were significantly more likely to accept non-receipted returns (for exchange or store credit) than stores with moderate return policies, and the latter were marginally significantly more likely to accept such returns than were stores with harsh return policies.[[61]](#footnote-61)

That said, across all stores, testers received more favorable treatment than the official policy required in a significant minority of cases. Overall, the audited stores exhibited a pro-consumer gap in 22% of the cases. Figure 3 illustrates the prevalence of pro-consumer gaps in each type of store.

*Figure 3: Gaps across Stores*

*Figure 3. Gaps at the initial (pre-complaining) stage*. The upper bar represents return outcomes overall (n = 95), while the lower three bars represent the return outcomes across stores with different policies.

As figure 3 shows, stores with moderate return policies were much more likely than stores with harsh return policies to exhibit a pro-consumer gap.[[62]](#footnote-62) One plausible explanation for this finding is that harsh policy stores almost exclusively consist of local retailers operating only one or two shops in Illinois.[[63]](#footnote-63) These local stores probably incur higher depreciations costs from accepting returns, in view of their lower ability to resell items or to return them to the supplier.

Unexpectedly, a small subset of stores (3%) departed from their return policies to the consumers’ *detriment*. Store clerks in these stores not only refused to refund the testers. They also refused to accept the non-receipted item for store credit or exchange, even though they were contractually required to do so. Although the experiment was not designed to explore pro-seller gaps, it is puzzling that sellers allow themselves to depart from their express policies to consumers’ disadvantage. I leave these findings for future investigation.[[64]](#footnote-64)

### Final Return Outcomes and the Effect of Complaining on the Gap

This study hypothesized that sellers may be significantly more likely to depart from their formal requirements after consumers complain. By waiting for the consumer to complain, the company can effectively let the high-value, high-information consumers identify themselves.[[65]](#footnote-65) Sellers may consequently use consumer assertiveness to determine whether to depart from the paper policy in favor of consumers.

This Section explores the impact of consumer bargaining strategy on the gap, by testing whether sellers are more likely to depart from their policies after consumers complain. For this purpose, testers were instructed to continue to bargain with the store clerk if denied a refund at the initial stage. Testers escalated assertiveness levels throughout the bargaining process, asking to speak to a manager and then asking the manager for a refund. This experimental design enables testing for the interaction between consumer assertiveness and the gap by comparing testers’ initial return outcomes to their outcomes after complaining. Figure 4 reports testers’ initial and final return outcomes across the different policy stores.

*Figure 4: Initial and Final Return Outcomes by Policy Type*

*Figure 4. Return outcomes at the initial (pre-complaining) and final (post-complaining) stages.* The upper pair of bars represent return outcomes overall (n = 95), while the three lower pairs of bars represent the return outcomes across stores with different policies.

Importantly, as figure \_\_ shows, even after testers had complained, the paper policy had a significant and strong effect on return outcomes, such that lenient stores were significantly more likely to accept non-receipted returns (for exchange, store credit or refund) than moderate policy stores,[[66]](#footnote-66) and moderate stores were significantly more likely to accept such returns than harsh policy stores.[[67]](#footnote-67)

At the same time, the results reveal that consumer assertiveness plays a major role in determining sellers’ leniency in practice. As expected, complaining significantly affected the on-the-ground behavior of stores. Yet, the magnitude of the effect is striking. While only 22% of the stores had treated testers more leniently than the policy required before testers complained, as many as 36% of the stores treated testers more favorably after testers had complained.[[68]](#footnote-68) Importantly, complaining significantly improved consumers’ likelihood of both receiving refunds and of having their returns accepted (in stores with a receipt requirement), notwithstanding their failure to show a receipt.[[69]](#footnote-69)

Figure 3 shows the effect of complaining on the gap across policy types, by reporting the return outcomes, at both the initial and final stages, across policy types.

*Figure 3. Pro-Consumer Gaps: Before and After Complaining*

As Figure 3 shows, complaining significantly improved testers’ return outcomes among both the harsh and moderate policy stores.[[70]](#footnote-70) Complaining also operated in the expected direction of improving consumers’ outcomes among the lenient policy stores, yet the effect was not significant.[[71]](#footnote-71) This is plausibly the case because of a “ceiling effect:” The lenient policy stores were already relatively generous towards consumers.

Importantly, complaining had no effect on consumers’ chances of obtaining a refund notwithstanding a clear “no refund” policy: All of the harsh (“no refund”) policy stores refused to provide refunds to testers both before and after they had complained. Yet, complaining significantly increased consumers’ chances of receiving an exchange or store credit (in the harsh and moderate policy stores, requiring a receipt for any exchange or return), from 23% to 39%. Finally, complaining marginally significantly increased consumers’ chances of receiving non-receipted refunds in lenient and moderate policy stores that required receipts for refunds, from 11% to 17%.

To summarize, the results reveal that, as expected, complaining significantly increased the probability that the seller would depart from the formal policy to the consumer’s advantage. The magnitude of the effect is remarkable: Complaining almost doubled the consumer’s chances of obtaining concessions not otherwise required by the contract.

What can explain the large role of assertiveness in shaping stores’ leniency on the ground? As explained before, sellers may use the gap to attract and keep the business of insistent consumers who complain without giving every consumer the same benefits they give complainers. Indeed, some consumers may not even realize that they can obtain concessions by complaining to the seller.

Put differently, the strategy of allowing employees to respond to consumer complaints with various forms of concessions not required by the contract may be seen as a cost-minimizing way for a store to increase the probability that it will both keep the business of insistent consumers and prevent such consumers from tarnishing its reputation (through social media, for example). By segmenting the insistent and the non-insistent buyers this way, the seller can save costs by offering the benefits only to those insistent buyers who expect them and would not buy the product without them.

A series of qualitative interviews I conducted in preparation for the field experiment further shed light on sellers’ use of the complaint-based mechanism.[[72]](#footnote-72) For example, in an interview conducted with a former store clerk at Abercrombie & Fitch (Chicago), she explained as follows:

Policy goes out the door when you have an unhappy customer. Policy goes out the door because you’re trying to keep that customer happy, so [you do] whatever you can do. [[73]](#footnote-73)

Ultimately, sellers will engage in cost-benefit analyses and decide accordingly about whom to treat more favorably than the contract requires. These decisions will necessarily hinge on the inevitable compromise between the reputational harm and inconvenience sellers will incur from enforcing unwanted terms on complaining buyers and the costs saved from refusing to make concessions. Consumers who complain and continue to bargain might increase sellers’ costs of adhering to the contract for several reasons. First, in the short term, insistent consumers might create more confusion and inconvenience for the clerks and managers in the store. Second, insistent consumers are more likely to generate reputational harm to the seller than those who relent and acquiesce. And third, complainers are plausibly more likely to refrain from entering into future transactions with the seller.

Recall, however, that the effect of complaining depended on the type of return policies the store had, so that in the context of the harsh “no refund” policies, complaining did not improve consumers’ chances of receiving a refund at all, while in the context of the other, more lenient policy terms, it had a significantly greater impact. This finding reveals, once again, that harsh “no refund” terms are sticky, such that retailers are reluctant to depart from them even after consumers complain. It suggests that the costs of departing from these rigid terms are substantial, such that they exceed even the costs of refusing to cater to consumers’ complaints.

The next Section shifts attention from the role of consumer assertiveness in shaping the gap to other determinants of outcomes on the ground. In particular, it focuses on store characteristics and their relationship with the gap.

### Store Characteristics and the Gap

This Section explores whether various store characteristics are associated with a higher likelihood of pro-consumer gaps. In particular, it tests the predictions that more luxurious stores will be more likely to depart from their formal policies even when keeping the paper policy fixed, and that chain stores will be more likely to apply pro-consumer gaps than will local stores.

The first proposition, that more luxurious stores will be more likely to exercise leniency on the ground, even when the paper policy is controlled for, is based on the premise that these stores’ reputations derive not only from their formal policies, but also, perhaps even mainly, from their willingness to deviate from these policies when facing dissatisfied consumers.

This prediction was supported by several of the store clerks I interviewed. For example, a former store clerk at Saks Fifth Avenue mentioned, in response to my question as to whether she has ever deviated from the store’s formal return policy, that: “It’s up-scale so they want to be very customer-friendly. So they were much more bendable with the rules, and they are an expensive store so they don’t want to lose a customer fighting over $30.”[[74]](#footnote-74) Similarly, a former store clerk at the Bally luxury shoe store stated that:

In a high-end luxury you didn’t want to lose a good client. If you have a good client come back in, more so with the women than men, saying “you know, I’ve worn these shoes for a couple of weeks, and they really itch, and I tried to make adjustments, but they’re just not really working for me,” you’re not gonna [*sic*] lose someone spending $25,000 a year at your store over a $500 pair.[[75]](#footnote-75)

More casual stores, in contrast, build their reputations by offering lower prices. Therefore, they will likely offer no discretionary benefits but charge lower prices.[[76]](#footnote-76)

The second prediction, that chain stores will be more likely to engage in pro-consumer gaps than local stores, is based on the premise that chain stores, like more luxurious stores, are likely to suffer higher reputational losses from refusing to cater to their customers’ demands. Local stores, at the same time, are likely to deviate from their policies when facing repeat customers in order to maintain their loyalty. But, when dealing with one-time customers, chain stores are more likely to behave leniently than are local stores, either because, unlike many local sellers, it would be difficult for them to distinguish between one-time and recurring customers, or because even one-time customers might harm their reputation by telling others what happened through social media. These informational channels are likely to affect local sellers to a lesser degree. Finally, local sellers typically incur higher depreciation costs in terms of their lower ability to resell or return non-defective items to the manufacturer, and are therefore likely to adhere to their policy terms, while chain stores have more flexibility to deviate from the same terms.

A series of multivariate regression analyses was used to explore these hypotheses and to evaluate the determinants of the gap.

In the first set of regressions, the dependent variable receives 1 if a pro-consumer gap was applied by the audited store, and 0 otherwise. The explanatory variables are the following store characteristics:

1. *Policy Type*: Harsh, moderate, or lenient;
2. *Luxuriousness:* Stores were classified as discount, mainstream, or luxury stores based on the median prices of clothing items posted on their websites. Discount stores were defined as all stores with median prices in the lower 25th percentile (i.e., lower than $26). Mainstream stores were defined as all stores with median prices between the 25th and 75th percentiles (i.e., between $26 and $70), and luxury stores were defined as all stores with median prices in the upper 25th percentile (i.e., between $70 and $350);
3. *Store Type:* Whether the store is local or part of a chain;
4. *Experience (Age):* Calculated as logarithm of years of operation since establishment (until 2019);
5. *Size*: As calculated by logarithm of annual revenues from the year of 2018 (included as control).

Table 7 reports the results of six multivariate regression models. In Models One and Two, the dependent variable is assigned 1 if a pro-consumer gap was applied by the audited store (i.e., if non-receipted refunds, exchanges, or store credits were provided) 0 otherwise. Model One looks at return outcomes at the first, pre-complaining stage, while Model Two looks at the final return outcomes, post-complaining. In models Three and Four, the dependent variable is whether a refund was provided notwithstanding the formal paper policy of either demanding receipts or denying refunds, while Model Three looks at outcomes at the first stage and Model Four looks at outcomes at the final stage. Finally, in models Five and Six, the dependent variable is whether non-receipted returns were accepted despite testers’ failure to show a receipt, while looking only at the moderate and harsh policy stores that require a receipt for any return. Model Five looks at first stage outcomes, while Model Six looks at the final stage. Across all regression models, the dependent variable is regressed on all store characteristics, while controlling for policy type.

*Table 7: The Effects of Store Characteristics on the Gap*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1)  Pro-Consumer  Gaps at Initial Stage as DV | (2)  Pro-Consumer  at Final Stage as DV | (3)  Refunds at Initial Stage as DV | (4)  Refunds at Final Stage as DV | (5)  Non-Receipted Returns at Initial Stage as DV | (6)  Non-Receipted Returns at Final Stage as DV |
| Moderate Policy | 0.355 | 0.257 | 0.220 | 0.181 | 0.182 | 0.0241 |
|  | (0.433) | (0.584) | (0.511) | (0.641) | (0.737) | (0.966) |
|  |  |  |  |  |  |  |
| Lenient Policy | 0.123 | -0.236 | 0.301 | 0.224 |  |  |
|  | (0.793) | (0.627) | (0.386) | (0.579) |  |  |
|  |  |  |  |  |  |  |
| Mainstream | 0.102 | 0.157 | -0.0492 | 0.0419 | 0.371\* | 0.447\* |
|  | (0.413) | (0.216) | (0.595) | (0.689) | (0.088) | (0.053) |
|  |  |  |  |  |  |  |
| Luxury | 0.593\*\*\* | 0.470\*\*\* | 0.193 | 0.237\* | 0.678\*\* | 0.476\* |
|  | (0.000) | (0.005) | (0.102) | (0.078) | (0.011) | (0.078) |
|  |  |  |  |  |  |  |
| Chain | 0.448 | 0.688\*\* | 0.0785 | 0.173 | 0.455 | 0.711\* |
|  | (0.123) | (0.024) | (0.711) | (0.483) | (0.236) | (0.084) |
|  |  |  |  |  |  |  |
| Log Age | 0.157 | 0.237\*\* | 0.260\*\*\* | 0.211\*\* | -0.0422 | 0.0669 |
|  | (0.122) | (0.025) | (0.001) | (0.017) | (0.801) | (0.703) |
|  |  |  |  |  |  |  |
| Log Revenues | -0.0204 | -0.0365 | -0.046\*\*\* | -0.041\*\* | 0.0225 | -0.00088 |
|  | (0.384) | (0.134) | (0.010) | (0.044) | (0.555) | (0.982) |
|  |  |  |  |  |  |  |
| Constant | -1.148\*\* | -1.348\*\* | -1.200\*\*\* | -1.079\*\* | -0.411 | -0.658 |
|  | (0.033) | (0.017) | (0.003) | (0.021) | (0.600) | (0.427) |
| Observations | 89 | 89 | 89 | 89 | 51 | 51 |
| R-squared | 0.373 | 0.426 | 0.390 | 0.265 | 0.348 | 0.360 |
|  |  |  |  |  |  |  |

*Notes*. The first column (Model 1) is a multivariate linear regression of pro-consumer gaps at the initial stage (taking the value of “1” if non-receipted refunds, exchanges, or store credits were provided, and “0” otherwise) on store policy type (with “Harsh” as reference category) and store characteristics. The second column (Model 2) is a multivariate linear regression of pro-consumer gaps at the final stage on policy type and store characteristics. The third column (Model 3) is a multivariate linear regression of refund outcomes at the initial stage on policy type and store characteristics. The fourth column (Model 4) is a multivariate linear regression of non-receipted returns at the initial stage on policy type and store characteristics, and the fifth column (Model 5) is a multivariate linear regression of non-receipted returns at the final stage on policy type and store characteristics. The reference category for policy type is “harsh,” the reference category for prestige is “discount,” and the reference category for store type is “local.” P-values in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

The results reveal that, consistent with the study’s predictions, luxury stores were significantly more likely to depart from their formal policies than discount stores, at both the initial (pre-complaining) and final (post-complaining) stages. In addition, and as expected, chain stores were significantly more likely to apply pro-consumer gaps than local stores.[[77]](#footnote-77) The results also show a significant relationship between a store’s age (as proxied by years since establishment) and its likelihood to depart from the paper policy in favor of consumers: The older, more experienced the store was, the more likely it was to behave more forgivingly than the formal policy dictated.[[78]](#footnote-78) Surprisingly, controlling for all other determinants, smaller stores were more likely to provide refunds despite testers’ failure to meet the policy requirements than larger stores. It is possible that, all else equal, smaller stores need to behave more forgivingly in order to be able to compete with the larger stores, yet this relationship should be further explored in future studies.

# IV. Discussion & Implications

## Pro-Consumer Gaps Persist, but only in a Minority of Stores

The results show that pro-consumer gaps persist across various types of retail stores, including chain and local, large and small, luxury and casual. While gaps are more prevalent in some contexts and stores than in others, a significant minority of retailers behaves more leniently towards consumers than their formal policies require. These results lend support to the theory that sellers may intentionally use ostensibly rigid, unconditional terms in their standardized agreements, but selectively enforce these terms in order to distinguish between different types of consumers. As a result, clear, bright-line terms may operate as standards in practice, while store clerks exercise discretion on the ground.

Why do some stores rather than others deviate from their formal policies in favor of consumers? The results show that several store characteristics are strongly associated with a higher likelihood of applying pro-consumer gaps. In particular, more luxurious, established, and chain stores were more likely to depart from their policies when facing one-time consumers than were less luxurious, less established, and local stores.

Indeed, these sophisticated stores are typically significantly associated with more generous return policies on paper. However, even when keeping the paper policy constant across stores, more luxurious and established stores, and stores operating as part of a chain, were more likely to depart from their policies in order to satisfy consumers’ demands. These findings suggest that product and store quality is signaled by, and reflected in, more generous return policies, both on paper and in practice. These findings also imply that “real deal” leniency and “paper deal” leniency are both correlated with product quality.

The strategy of allowing employees the discretion to grant case-specific benefits beyond those required by standard-form contracts can be seen as a sophisticated way for the firm to increase its revenues by gaining the loyalty of existing customers and also establishing a good reputation attracting new customers.

Yet, importantly, most stores do not depart from their formal receipt requirements in favor of consumers. In fact, in 78% of the stores, store clerks refused to grant consumers *any concessions* absent a receipt; and non-receipted refunds were denied in 92% of the stores. To the extent that gap theorists predict that all, or even a majority, of sellers will depart from the paper deal (at least when facing non-opportunistic consumers), these findings are inconsistent with such predictions.

While more luxurious and established stores were more likely to exercise forgiveness, even among those stores, the majority of sellers were unwilling to depart from the four corners of the agreement. And while consumers’ complaints significantly increased their chances of securing more lenient behavior, the majority of sellers refused to depart from the formal requirements even after consumers had complained, and in most cases, even after consumers had bargained with the stores’ managers.

The fact that sellers often refuse to grant concessions does not, in and of itself, imply that sellers’ adherence to their formal contract terms is inefficient. In competitive markets with no informational asymmetries, sellers have strong incentives to satisfy consumers. It therefore seems unlikely that sellers will insist on enforcing their rigid, rule-like terms to the letter at the risk of driving away customers unless these terms reflect an efficient risk allocation between sellers and consumers.[[79]](#footnote-79)

Admittedly, sellers’ insistence on following the contract to the letter may be efficient, even if the particular contract term *seems* excessively rigid or unfavorable to the consumer. Yet, the results suggest that policymakers should be cautious of inferring that seemingly one-sided terms in standard form contracts do not warrant regulatory scrutiny based solely on the premise that sellers will depart from these terms in practice.[[80]](#footnote-80)

## Gaps Might Generate Distortions

Even when gaps persist, they might generate distortions—both *ex ante* and *ex post*.

### *Ex Ante* Distortions of Consumers’ Decisions

At the *ex ante,* pre-purchase stage,consumers might be unable to distinguish between sellers that strictly enforce their paper policies and those that, by not doing so, provide higher quality services. This informational lacuna could distort consumers’ purchasing decisions. For example, consumers might refrain from buying items from particular sellers because consumers may wrongly assume that these sellers enforce their policies to the letter, while, in fact, the latter often behave more leniently than their policies dictate.

Firms that enforce their terms to the letter will have no incentive to let consumers know that they provide lower quality services by virtue of their refusal to grant concessions. Sellers that systematically deviate from their policies when facing non-opportunistic consumers are apparently incentivized to advertise this practical leniency, thereby distinguishing themselves from the stores that strictly adhere to the written agreement vis-à-vis all consumers. But if a store’s policy to systematically under-enforce its policy is made public, store clerks might not be able to fend off those opportunistic consumers against whom the formal policy was adopted to protect in the first place. The fact that non-lenient firms benefit from the behavior of lenient firms by cloaking themselves as such might, in turn, lower the incentives of firms to be lenient, thereby resulting in a “lemons equilibrium.”[[81]](#footnote-95)

So why do gaps persist? One reason is that even if most consumers could not distinguish between high quality, lenient stores and low quality stores *ex ante*, stores may gain consumer loyalty and create a customer base by applying a gap *ex post* when consumers complain, thereby informing their more sophisticated consumers about their lenient policies in practice. Still, less sophisticated and one-time customers might not be able to distinguish among stores, and might make inefficient consumption decisions because of this informational asymmetry.

### *Ex Post* Distortions of Consumers’ Decisions

The observed gaps may also lead to several inefficiencies at the post-contract stage. First, some consumers might be discouraged from even trying to obtain concessions from the seller that either vary from or directly contravene its formal policy. In the specific context of returns, consumers may be discouraged by the clear language of the policy and may consequently refrain from trying to return items to the seller if, according to the explicit terms of the policy, the items are not eligible for returns or if the consumers otherwise fail to meet the conditions set forth in the policy. In fact, sellers may adopt harsh contract terms precisely in view of this *in terrorem* effect on consumers.[[82]](#footnote-96)

Second, even if consumers do request concessions from sellers, they may relinquish their claims once sellers refer them to the contracts that they had “agreed” to enter into. Store clerks referred testers to the formal policy in the vast majority of stores, including half of the stores in which a gap was ultimately applied. Testers were instructed to continue to bargain and complain despite initial rejections of their claims. Yet, recent evidence suggests that other purchasers, perhaps most, would react to the store clerks’ initial negative responses dismissing their requests by acquiescing and accepting the formal contractual terms without dispute.[[83]](#footnote-97)

This reaction can be attributed to the fact that consumers are contract formalists.[[84]](#footnote-98) They tend to believe that the contract is the final word, and this preconception may be particularly strong in the context of standardized consumer agreements.[[85]](#footnote-99) Consumers are often demoralized by harsh and unconditional contractual language, and consequently refrain from bringing claims to the seller.[[86]](#footnote-100) In fact, previous research has shown that consumers rarely question the validity of contracts that disclaim their mandatory rights and remedies.[[87]](#footnote-101) Consumers are similarly unlikely to challenge contracts induced by fraud because they feel bound by contracts they signed.[[88]](#footnote-102)

In the specific context of the gap, consumers are not likely to realize that sellers often depart from their policies upon consumers’ requests. Indeed, consumers may reasonably assume that if the seller’s policy was one of leniency, the seller would put this policy in writing in order to increase sales. If substantial proportions of high-value, non-opportunistic consumers are unlikely to complain, even when their gains from the concession exceed the costs to the seller, the observed complaint-based segmentation of consumers might lead to inefficient outcomes.

Yet, even if consumers do not relent and continue to bargain and complain, ultimately, if the sellers remain firm in their decision to strictly enforce their agreements, consumers do not have any real ability to discipline them. Having no legal entitlement, consumers cannot even threaten to take legal action.[[89]](#footnote-103)

Some of these concerns may be addressed and mitigated through informational flows. For example, consumers may post online reviews praising stores that exercise leniency while criticizing or shaming those that do not. Sellers, in turn, will be incentivized to depart from one-sided terms, fearing the risks of reputational harm, either from attempting to impose unwanted terms on buyers or from becoming known as inflexible when disputes arise. Yet, there are reasons to fear that these informational flows will be inadequate in disciplining sellers.

First, consumers may not feel wronged by a seller’s adherence to the written policy, even if a seller is known for exercising leniency towards other consumers. In such cases, consumers may believe that the seller has a legitimate reason for failing to exercise leniency or to respond to their complaints, reasoning that a store that enforces the terms of its contracts, even selectively, is merely doing what it has a contractual right to do.

Concurrently, when sellers depart from the contract in favor of consumers, consumers might not report this more lenient behavior because they may believe, perhaps erroneously, that the seller’s leniency was a one-time occurrence rather than strategic, albeit concealed, behavior. Consumers are unlikely to realize that sellers use these harsh terms in order to fend off opportunistic consumers. For why would sellers advertise strict policies or terms that might discourage or scare away consumers, rather than more lenient terms that could encourage consumers to buy more, if sellers actually intended to behave leniently towards consumers at the post-contract stage? As a result, consumers may not share their experiences of sellers’ leniency with others.

But even if they did provide information about sellers’ deviations from their formal policies on social media, other consumers may fail to realize that these deviations reflect a systematic inclination, rather than one-time deviations, for the same reasons mentioned earlier.

### *Ex Ante* Distortions of Sellers’ Decisions

In addition to distorting consumers’ purchasing decisions, the ability to apply a gap might also lead sellers to adopt inefficient contractual risk allocations. The traditional assumption in the gap literature is that complaint-based segmentation of consumers may be welfare enhancing. For example, Jason Scott Johnston has argued that firms’ practice of providing “benefits to consumers who complain” beyond those that their standard forms oblige them to provide help maintain “cooperative, value-enhancing relationships” between firms and their customers. Johnston further maintains that “were firms legally required to extend such benefits . . . —as would result either from judicial invalidation of the tough standard-form performance terms or legislatively mandated generous standard-form performance terms—then both firms and their customers would be worse off.”[[90]](#footnote-104)

Yet, this may not be true because of the *ex post* chilling effect of rigid, bright-line terms on consumers. Sellers may find it profitable to begin with a strict term, even if it is harmful to consumers, under the assumption that a sufficient number of buyers will be discouraged from complaining about the contractual arrangement (or at least resign once the seller refers them to its written policy). If the combined costs of accommodating buyers who complain and any reputational loss from disappointed buyers who fail to complain are lower than the costs of accommodating the claims of all disappointed consumers (including the non-complainers), sellers will adopt the strict term denying the accommodation and depart only selectively in favor of those consumers who complain.[[91]](#footnote-105) Importantly, sellers may adopt the strict term even if a more lenient term reflects a more efficient risk allocation from a social welfare perspective (to the extent that these terms are not sufficiently salient to consumers *ex ante*, and therefore cannot be translated into higher prices).

### Distributional Concerns

Stores were nearly twice as likely to depart from the paper contract once consumers complained. This complaint-based segmentation of consumers effectively leads to the cross-subsidization of the complaining consumers by the more acquiescent, non-complaining consumers.

If consumer assertiveness is correlated with socio-economic status, this complaint-based discrimination might have troubling distributive implications.[[92]](#footnote-106) There is abundant evidence that people from lower socio-economic backgrounds exhibit a lower sense of entitlement, a lower willingness to bargain over payoffs, and a higher inclination to accept unfavorable offers than do people from higher socio-economic strata.[[93]](#footnote-107) Similarly, race and gender were found to influence what people expect and feel they deserve, with blacks and females feeling significantly less entitled than do whites and males.[[94]](#footnote-108)

In the particular context of contracts, evidence suggests that consumers belonging to minority groups (e.g., black and female buyers) and to lower socio-economic status are more likely to feel bound by standardized agreements and to view the four corners of these agreements as the final word.[[95]](#footnote-109)

In the context of the gap, it is plausible therefore, that lower-income, less-educated consumers, as well as members of disadvantaged groups, will feel more discouraged and demoralized by the formal policy or by sellers’ initial refusals to grant concessions. This, in turn, might lead to a troubling conclusion that due to sellers’ selective enforcement of their formal agreements, poorer, less educated consumers, and racial and gender minority groups ultimately cross-subsidize those consumers who are better-off.

The interviews conducted for this research provide further support for this prediction. For example, a former store clerk at a Chicago rug store explained, in response to the Author’s question as to whether he has ever deviated from the store’s formal return policy, that:

“Our policy was to charge a $100 delivery fee, but there might be something in the conversation […] where I’d say: ‘Ok, I’ll waive it for you’ if they ask. […] Those who managed to get their fees waived were typically white baby-bloomers. […] There are plenty wealthy people of color who buy rugs, but to my memory, the people who would get their fees waived were mainly white. The black customers wouldn’t typically ask for their fees to be waived.”[[96]](#footnote-110)

In addition, allowing store clerks discretion in exercising leniency toward certain consumers might produce regressive distributional outcomes to the extent that store clerks would use this discretion to the systematic disadvantage of certain group members.[[97]](#footnote-111) Store clerks might use their discretion in a discriminatory fashion either for animus-based or statistical-based reasons (for example, store clerks could potentially use race, gender, or perceived socioeconomic status as proxies of value to the firm of keeping such customers satisfied). At least in the context of race, such discrimination violates the Civil Rights Act of 1991, which prohibits racial discrimination in making, performing, modifying, and terminating contracts and in restricting access to “the enjoyment of all benefits, privileges, terms, and conditions of the contractual relationship.”[[98]](#footnote-112)

The potentially harmful effects of sellers’ departures from their contractual arrangements, both to consumers as a group and to members of disadvantaged groups, illustrate, once again, that if the “paper deal” is excessively harsh, regulators may need to intervene in the terms of consumer transactions in order to enhance consumer welfare, even when pro-consumer gaps are relatively prevalent.

# Conclusion

reveal

Consistent with the predictions of the Gap Theory, at least in the context of product returns, the “real deal” departs in meaningful ways from the “paper deal” in a non-negligible minority of stores. Yet, concurrently, the findings demonstrate that the majority of sellers enforce their formal policies to the letter, even after consumers complain.

Furthermore, departure patterns vary across stores and policy terms. Some terms are stickier than other terms, some stores are less likely to depart from their paper terms than other stores, and some customers are more likely to benefit from these departures than other customers. In particular, rigid “no refund” terms are less likely to be departed from than other, “softer,” policy requirements. Luxury stores, more established stores and chain stores are more likely to exercise tailored forgiveness than are more casual, less established and local stores. In addition, consumers who bargain and complain are more likely to obtain concessions than are non-insistent consumers.

Given the observed variation in the divergence between the terms of the paper deal and the terms of the “real deal,” it is important for policymakers and courts to identify the terms of the “real deal” before concluding that regulatory intervention is unnecessary.

This Article not only shows that, as a descriptive matter, gaps are not as prevalent as could be expected. It also questions, on a normative level, the prevailing assumption that “pro-consumer” gaps are welfare-enhancing and conducive to non-opportunistic consumers as a group. It proposes, rather, that these gaps can be harmful to consumers.

The Article points to two types of distortions: *ex ante* and *ex post*. *Ex ante*, consumers might not enter into beneficial transactions because they fail to realize that it may be possible to benefit from a more lenient treatment than the policy requires. *Ex post*, consumers might fail to realize that the seller can depart from the paper deal in their favor, and might consequently refrain from requesting the benefit, or relent once the seller refers them to the standardized agreement. At the same time, sellers will not be incentivized to offer more lenient terms in their standard form contracts, even if such terms are efficient from a social welfare perspective, as long as the costs sellers incur from having to provide the benefit to *all* consumers, including opportunistic buyers, exceed the aggregate costs of discouraging certain consumers from buying at the store.

If the costs of these distortions to consumers exceed the gains to the sellers from adopting the more rigid paper terms alongside a “gap strategy,” regulators might ultimately need to intervene in the contents of the paper deal, to ensure that the paper deal accurately reflects the real deal.

In addition, regulators could try to promote information sharing about on-the-ground leniency. The gap strategy harms the non-opportunistic consumers only if they are unaware of its existence. Yet, if consumers learn about firms’ actual leniency through social networks or online, the mentioned distortions may be corrected, or at least, their effect substantially mitigated.

Finally, this Article suggests that gaps might lead to regressive distributional outcomes, both because insistence and assertiveness are plausibly correlated with higher socio-economic status, gender, and race, and because store clerks may apply their on-the-ground discretion discriminatorily. If these propositions are corroborated in future research, it may be warranted to prohibit sellers from using a complaint-based mechanism to screen consumers, and to monitor sellers’ behavior through the use of audit techniques similar to the technique used in this study. More generally, this Article suggests shifting scholarly and regulatory attention from looking almost exclusively at the terms of the “paper deal” to looking more critically at the terms of the “real deal.”

This study represents an attempt to shed empirical light on the contract in action, using the case study of product returns. As any first step, the study inevitably leaves many questions, including questions concerning the generalizability of the findings, for future research. In particular, this study has focused on a single policy condition, (the receipt requirement) in a specific type of contract (return policies) in a particular product market (the retail clothing market) in one city (Chicago). In future studies, it will be worthwhile to explore whether and when gaps persist in other types of consumer markets and contracts, and to further investigate the determinants of the gap and the effect of consumer complaints on sellers’ tailoring decisions.

Admittedly, future studies of different terms, markets, and geographical locations may produce different results. Yet, the findings provide preliminary evidence that, at least in the context studied here, pro-consumer gaps persist, sellers use complaint-based mechanisms to segment consumers, and most sellers enforce their harsh policy requirements to the letter.

It is, of course, possible that in more competitive markets, with stronger reputational incentives, better informational flows, and more sophisticated consumers, pro-consumer gaps will proliferate. Yet, the case of product returns should be the context in which the effect of Gap Theory is most conspicuous and gaps are most prevalent. Indeed, anecdotal evidence pertaining to sellers’ departures from their return policies have been most commonly used by Gap Theory proponents to argue that, because of the existence of the gap, we should not be worried about buyer-unfriendly terms. This study casts serious doubt on the assumption that market forces can unquestionably be relied upon when the paper deal is excessively or unnecessarily strict.

[Appendices --- to be added]

Appendix 1

Return Outcomes at the initial Stage:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Store Type | Return Denied[[99]](#footnote-113) | Exchange/Store Credit | Refund Offered | Pro-Consumer Gaps |
| Harsh  (No Refund)  )n = 23) | 91%\*\*\* | **9%\*\*\*** | 0% | **9%\*\*** |
| Moderate  (Receipt required for *any* return)  (n = 33) | 64%\*\*\* | **33%\*\*\*** | **3%\*\*** | **36%\*\*** |
| Lenient (Receipt required for refund only)  (n = 39) | 8%\*\*\* | 74%\*\*\* | **18%\*\*** | **18%\*** |

*Note:* When the differences in return outcomes between two policy type categories are significant at the 1% level, they are marked as \*\*\*, when they are significant at the 5% level they are marked as \*\*, and when they are significant at the 10% level, they are marked as \*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Store Type | Stage | Return Denied[[100]](#footnote-114) | Exchange/Store Credit | Refund Offered | Pro-Consumer Gaps |
| Harsh  (No Refund)  (n = 23) | Initial | 91% | **9%** | 0% | 9%\* |
| Final | 70% | **30%** | 0% | 30%\* |
| Moderate  (Receipt required for *any* return)  (n = 33) | Initial | 64% | **33%** | **3%** | 36%\* |
| Final | 42% | **45%** | **12%** | 57%\* |
| Lenient (Receipt required for refund only)  (n = 39) | Initial | 8% | 74% | **18%** | 18% |
| Final | 8% | 72% | 21% | 21% |

1. \* Olin Law & Economics Research and Teaching Fellow, University of Chicago Law School. [acknowledgments to be added]. This project benefited from the generous support of the University of Chicago Coase-Sandor Institute for Law & Economics and the Harvard Institute for Quantitative Social Science. [↑](#footnote-ref-1)
2. *See*Henry Butler, Christopher Drahozal, & Joanna Shepherd, Economic Analysis for Lawyers 183 (3d ed. 2015) (explaining that *“*[f]orms reduce transactions costs and benefit consumers because, in competition, reductions in the cost of doing business show up as lower prices”); Melissa T. Lonegrass, *Finding Room for Fairness in Formalism-the Sliding Scale Approach to Unconscionability*, 44 Loy. U. Chi. L.J. 1, 27 (2012) (explaining how standard form contracts reduce transaction costs, promote efficient transactions, and benefit consumers); Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 Harv. L. Rev. 1173, 1221 (1983) (explaining how form contracts save money and promote transactions).  [↑](#footnote-ref-2)
3. *See, e.g.*,Margaret J. Radin, Boilerplate: The Fine Print, Vanishing Rights, and the Rule of Law (2013) (noting that non-negotiable boilerplate terms regularly deprive non-drafting parties of their most basic rights); Edith Warkentine, *Beyond Unconscionability: The Case for Using “Knowing Assent” as the Basis for Analyzing Unbargained-for Terms in Standard Form Contracts*, 31 Seattle U. L. Rev. 469, 515 (2007) (observing that drafting parties often hide one-sided terms in their boilerplates); Oren Bar-Gill, Seduction by Contract: Law, Economics, and Psychology in Consumer Markets (2012) (showing how sellers exploit consumers’ bounded rationality and systematic cognitive biases through contract design); Russell Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. Chi. L. Rev. 1203 (2003) (arguing that drafting parties have an incentive to introduce self-serving terms in view of the non-drafting parties’ bounded rationality); Richard Craswell, *Taking Information Seriously: Misrepresentation and Nondisclosure in Contract Law and Elsewhere*, 92 Va. L. Rev. 565, 591 (2006) (“[I]f consumers . . . have no information (or only poor information) about the effect of the contract terms used by any individual seller, each seller will . . . have an incentive to degrade the ‘quality’ of its terms.”); Nancy S. Kim, Wrap Contracts: Foundations and Ramifications 29 (2013) (suggesting that sellers use one-sided clauses, such as dispute resolution provisions, to hinder buyers’ access to the judicial system); J. Maria Glover, *Beyond Unconscionability: Class Action Waivers and Mandatory Arbitration Agreements*,59 Vand. L. Rev. 1735 (2006); Omri Ben-Shahar, *The Myth of the ‘Opportunity to Read’ in Contract Law*, 5 Eur. Rev. Contract L. 1 (2009) (questioning consumers’ ability to understand and comprehend contract terms and discussing the widespread use of boilerplate or one-sided terms); David A. Hoffman, *Relational Contracts of Adhesion*, 85 U. Chi. L. Rev. 1395, 1396 (2018) (explaining that “[b]ecause consumers don’t read their contracts, firms can make ‘hidden’ terms worse without lowering prices”).  [↑](#footnote-ref-3)
4. *See, e.g.*, Avery Katz, *The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation*, 89 Mich. L. Rev. 215, 281 (1990) (“Having the terms [unfavorable to the consumer] in the writing gives a seller the discretion to invest in goodwill in circumstances where it is most valuable to do so, while leaving him the option of enforcing the contract to the letter at other times.”); Robert A. Hillman & Jeffrey Rachlinski, *Standard Form Contracting in the Electronic Age*, 77 N.Y.U. L. Rev. 429, 441 (2002); Clayton P. Gillette, *Rolling Contracts as an Agency Problem*,Wis. L. Rev. 679, 704–12 (2004) (suggesting that sellers may use a “contract clause that assigns an entitlement to the seller, but that the seller may under-enforce when it is dealing with a good claimant”); Clayton P. Gillette, *Pre-Approved Contracts for Internet Commerce*, 42 Houston L. Rev. 975, 977 (2005) (observing that sellers may use “ostensibly oppressive terms” to allow themselves “discretion to treat buyers who appear to be acting in good faith differently from those who appear to be acting opportunistically”); Lucian A. Bebchuk & Richard A. Posner, *One-Sided Contracts in Competitive Consumer Markets*, 104 Mich. L. Rev. 827, 828 (2006) (suggesting that “reputational considerations” may “induce the seller to treat the buyer fairly even when such treatment is not contractually required”); Jason Scott Johnston, *The Return of the Bargain: An Economic Theory of how Standard Form Contracts Enable Cooperative Negotiation Between Businesses and Consumers*, 104 Mich. L. Rev. 857 (2006) (suggesting that firms use “clear and unconditional standard-form contract terms not because they will insist upon these terms, but because they have given their managerial employees the discretion to grant exceptions from the standard-form terms on a case-by-case basis”); Douglas Baird, Reconstructing Contracts 129 (2013) (“For all I knew, Norm had a form that disclaimed the implied warranty of merchantability, but such a disclaimer was irrelevant as long as reputational forces ensured that he would make amends if his goods did not pass in his trade.”); Eyal Zamir, *Contract Law and Theory: Three Views of the Cathedral*, 81 U. Chi. L. Rev. 2077 (2014); Lisa Bernstein & Hagay Volvovsky, *Not What you Wanted to Know: The Real Deal and the Paper Deal in Consumer Contracts: Comment on the Work of Florencia Marotta-Wurgler*, 12 Jrsl. Rev. Legal Stud. 128, 129 (2015). For a similar claim in the context of franchise agreements, see Klein, note 3, at 358–60. [↑](#footnote-ref-4)
5. The terms “paper deal” and “real deal” are derived from Stewart Macaulay, who used these terms in his work on divergences between formal agreements and their actual implementation in business-to-business transactions. *See, e.g.*,Stewart Macaulay, *The Real and the Paper Deal: Empirical Pictures of Relationships, Complexity and the Urge for Transparent Simple Rules*, 66 Mod. L. Rev. 44, 79 (2003); Stewart Macaulay & William Whitford, *The Development of Contracts: Law in Action*, 87 Temple L. Rev. 793 (2014).Since then, several scholars have applied this terminology in the consumer contracts context. *See, e.g.*, Bernstein and Volvovsky, *supra* note 3, at 129 (suggesting that “studies of consumer contracts in particular contexts should move from looking almost exclusively at the terms of the paper deal to looking at the terms of the real deal”). [↑](#footnote-ref-5)
6. *See, e.g.*,David A. Hoffman, *From Promise to Form: How Contracting Online Changes Consumers*, 91 N.Y.U. L. Rev. 1595, 1641 (2016) (explaining that “[t]he problem is that firms might be able to insist (in the law’s shadow) that consumers comply with unenforceable [one-sided] terms, simply because those consumers misconstrue the operative rules”); Yonathan A. Arbel, *Reputation Failure: The Limits of Market Discipline in Consumer Markets*, Wake Forest L. Rev. (forthcoming 2019) (arguing that reputation fails to adequately discourage sellers from enforcing harsh terms to the letter); Manisha Padi, *The Exercise of Contract Rights* (draft; on file with the Author) (finding that, in the context of foreclosures, “market forces, such as competition across lenders or consumer bargaining, are not strong enough to discipline” lenders from exercising contract rights inefficiently); Shmuel I. Becher & Tal Z. Zarsky, *Minding the Gap*, 51 Conn. L. Rev. 1, 12 (forthcoming 2019) (arguing that online information flow is less powerful when the firm’s conduct is not aligned with its contractual language). [↑](#footnote-ref-6)
7. *See, e.g.*, Gillette, *supra* note 3, at 712; Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-7)
8. *See, e.g.*, Arbel, *supra* note 5; Becher & Zarsky, *supra* note 5, at 14; *see also* Eyal Zamir & Doron Teichman, Behavioral Law & Economics 304 (2018) (noting the “ongoing, heated debate over the appropriate legal treatment of one-sided clauses in standard-form contracts”); Eyal Zamir & Ian Ayres, *Mandatory Rules* (Hebrew Univ. of Jerusalem Legal Research Paper No. 19-12, 2019), https://ssrn.com/abstract=3420179 (noting that while “some are inclined to conclude that regulation (or much of it) should be abandoned altogether, leaving the scene to market forces of reputation and competition,” others call “for more serious and systematic consideration of the use of mandatory regulation of the content of transactions”). [↑](#footnote-ref-8)
9. Scholars who discussed the gap and its implications for public policy have acknowledged the empirical deficiency, calling for future empirical work on these issues. *See, e.g.*,Gillette, *supra* note 3, at 698; Becher & Zarsky, *supra* note 5, at 13 (suggesting, for example, that “further empirical and analytical work is required” in order to distinguish sellers’ systematic deviations from their formal contracts from more sporadic departures). [↑](#footnote-ref-9)
10. *See, e.g.*, Gillette, *supra* note 3, at 705 (suggesting that “the seller may offer a full refund to a buyer . . . notwithstanding that the terms of the contract permit a lesser remedy”); Becher & Zarsky, *supra* note 5, at 1 (suggesting that “a vendor may stipulate a ‘no refund and no returns’ policy, yet exhibit—at least in some circumstances—accommodating, lenient behavior”). [↑](#footnote-ref-10)
11. Schmuel I. Becher & Esther Unger-Aviram, *The Law of Standard Form Contracts: Misguided Intuitions and Suggestions for Reconstruction*, 3 DePaul Bus. & Com. L.J. 199, 208 (2010). [↑](#footnote-ref-11)
12. Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-12)
13. This finding is consistent with some of the predictions made in the gap literature. *See,* e.g., Gillette 2004, *supra* note 3, at 707 (hypothesizing, without providing empirical data, that sellers might use the gap to distinguish between insistent and non-insistent buyers); Johnston, *supra* note 3, at 881 (suggesting that firms use customer complaints to determine the value of tailored concessions to consumers). [↑](#footnote-ref-13)
14. *See, e.g.*, Tess Wilkinson-Ryan, *Intuitive Formalism in Contract*, 163 U. Pa. L. Rev. 2109 (2015); Tess Wilkinson-Ryan & David A. Hoffman, *The Common Sense of Contract Formation*, 67 Stan. L. Rev. 1269, 1281–98 (2015) (finding that laypeople put excessive weight on written terms compared to oral agreements, believe that contracts are formed primarily through formalities such as signature and payment (even though contract law does not require such formalities for a contract to be formed), and generally feel obligated to abide by terms that follow formalized assent processes); Yuval Feldman & Doron Teichman, *Are All Contractual Obligations Created Equal?*, 100 Geo. L.J. 5, 5 (2012) (arguing that laypeople feel they are bound to the signed contract); Meirav Furth-Matzkin, *The Harmful Effects of Unenforceable Contract Terms: Experimental Evidence*, 4 Ala. L. Rev. 1032 (2019) (finding that tenants perceive their lease terms as enforceable and binding, even when they contain unenforceable contract terms); Meirav Furth-Matzkin & Roseanna Sommers, *Consumer Psychology and the Problem of Fine Print Fraud*, 72 Stan. L. Rev. (forthcoming 2020) (finding that consumers are demoralized by fraudulent fine print even when it contradicts what they were promised at the pre-contractual stage). [↑](#footnote-ref-14)
15. *See infra* notes 108-112. [↑](#footnote-ref-15)
16. *See* sources cited *supra* note 3. [↑](#footnote-ref-16)
17. *See, e.g.*, Gillette, *supra* note 3, at 704–07; Bebchuk & Posner, *supra* note 3, at 827. [↑](#footnote-ref-17)
18. Bebchuk & Posner, *supra* note 3, at 827–28 (“A seller concerned about its reputation can be expected to treat consumers better than is required by the letter of the contract. But the seller’s right to stand on the contract as written will protect it against opportunistic buyers.”). [↑](#footnote-ref-18)
19. Gillette, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 831. On problems of observability by courts in the context of franchise agreements, see Klein, *supra* note 3. [↑](#footnote-ref-19)
20. Gillette, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 834; Johston, *supra* note 3, at 882. [↑](#footnote-ref-20)
21. Gillette, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 831. [↑](#footnote-ref-21)
22. Gillette, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-22)
23. Bebchuk & Posner, *supra* note 3, at 828 (“As our analysis highlights, . . . the distinction is also relevant to contracts that businesses enter into with consumers who are not repeat players. As long as the business is a repeat player with the consumer side of the market, its expectation of doing business with other consumers in the future may dissuade it from enforcing a one-sided contract to the hilt against a particular customer even though the business does not expect to have further dealings with that consumers.”). [↑](#footnote-ref-23)
24. Zamir & Teichman, *supra* note 7, at 311; Becher & Zarsky, *supra* note 5, at 14 (arguing that “there is no guarantee that online information flow indeed features the seller’s lenient treatment” because “consumers may experience the lenient treatment yet refrain from posting it online due to insufficient motivation”); Arbel, *supra* note 5, at 5 (suggesting that we have reached a point of “reputational failure”). [↑](#footnote-ref-24)
25. Johnston, *supra* note 3, at 887. [↑](#footnote-ref-25)
26. Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-26)
27. Arbel, *supra* note 5, at 1. [↑](#footnote-ref-27)
28. *See, e.g.,* Bebchuk & Posner, *supra* note 3 (using the case of return policies in their model); Gillette, *supra* note 3, at 705 (suggesting that “the seller may offer a full refund to a buyer . . . notwithstanding that the terms of the contract permit a lesser remedy”); Becher & Zarsky, *supra* note 5, at 1 (suggesting that “a vendor may stipulate a ‘no refund and no returns’ policy, yet exhibit—at least in some circumstances—accommodating, lenient behavior”). [↑](#footnote-ref-28)
29. To date, commentators have relied only on limited anecdotal evidence suggesting that retailers behave more leniently towards consumers. *See, e.g.*, Johnston, *supra* note 3, at 873 (suggesting that “[r]etail-return policies. . . dramatically illustrate the reality and significance” of what he terms “two-part standard-form contracts”—the contract on paper and the contract on the ground, while relying solely on anecdotal evidence). [↑](#footnote-ref-29)
30. Lucas Reilly, *By the Numbers: How Americans Spend Their Money*, Mental Floss(July 17, 2012), available at http://mentalfloss.com/article/31222/numbers-how-americans-spend-their-money. [↑](#footnote-ref-30)
31. *See, e.g.*,Zamir & Teichman, *supra* note 7, at 290–91; Shmuel Becher & Tal Zarsky, *Open Doors, Trap Doors and the Law*, 74 L. & Contemp. Probs. 63, 73 (2011). [↑](#footnote-ref-31)
32. Courtney Reagan, *A $260 Billion ‘Ticking Time Bomb’: The Costly Business of Retail Returns*, CNBC (Dec. 16, 2016), available at https://www.cnbc.com/2016/12/16/a-260-billion-ticking-time-bomb-the-costly-business-of-retail-returns.html. [↑](#footnote-ref-32)
33. According a recent consumer poll, 91% of consumers consider return policies as very important in their purchasing decisions. *See* Rimma Kats, *Many Consumers Avoid Retailers with Strict Return Policies*, eMarketer (Jan. 1, 2018), available at https://retail.emarketer.com/article/many-consumers-avoid-retailers-with-strict-return-policies/5a4c05a7ebd40008a852a26c; AllBusiness.com, *The Importance of a Good Return Policy*, N.Y. Times (July 10 2007), https://archive.nytimes.com/www.nytimes.com/allbusiness/AB4353479\_primary.html. On the importance of the right to withdraw to consumers, see, e.g.,Zamir & Teichman, *supra* note7, at 290–91. [↑](#footnote-ref-33)
34. Federal law provides for a cooling-off period in a handful of door-to-door transactions. *See* 16 C.F.R. § 429 (setting forth the federal “Cooling-Off Rule,” which stipulates that buyers are entitled to a three-day “cooling-off period” during which the buyer can cancel certain door-to-door sale transactions). The Truth in Lending Act of 1968 similarly allows consumers three business days to cancel credit transactions (see 15 U.S.C. §§ 1601). *See also* 27 C.F.R. § 11.32 (granting consumers the right to return defective products); 27 C.F.R. § 11.33 (granting consumers the right to return products in case of “any discrepancy between products ordered and products delivered”). [↑](#footnote-ref-34)
35. At the state level, some jurisdictions provide little additional protections beyond those specified federally, while others (e.g., Illinois, Connecticut, Washington, Pennsylvania, and Michigan) have expanded the federal three-day right to cancel transactions, extending it to certain purchases not covered under federal law. Still others require stores to disclose their return policies, or else mandate a right to withdrawal for stores where policies prohibiting returns are not clearly displayed. [↑](#footnote-ref-35)
36. *See, e.g.*, Omri Ben-Shahar & Eric A. Posner, *The Right to Withdraw in Contract Law*, 40 J. Legal Stud. 115, 139–40 (2011) (advocating for a default right to withdraw); Zamir & Teichman, *supra* note 7, at 292 (discussing the desirability of regulating the right to withdraw from a behaviorally informed perspective, and suggesting that “[a]t the very least, contract terms that unreasonably raise the costs of exercising the return option appear to warrant regulation.”); Shmuel I.Becher & Tal Z. Zarsky, *Open Doors, Trap Doors, and the Law*, 74 L. & Contemp. Probs. 63, 63–64, 89 (2011) (suggesting that regulators who embrace “the open door dynamic”—i.e., those who promote mandatory or default rights of withdrawal—may misunderstand “crucial elements” of consumer psychology that explain consumers’ reluctance or inability to invoke those rights in practice). In support of regulating consumer contracts more generally, see, *e.g*., Jeff Sovern, *Toward a New Model of Consumer Protection: The Problem of Inflated Transaction Costs*, 47 Wm. & Mary L. Rev. 1635 (2006) (arguing that sellers are often financially incentivized to inflate—rather than reduce—consumer transaction costs). [↑](#footnote-ref-36)
37. *See, e.g.*, Jan M. Smits, *Rethinking the Usefulness of Mandatory Rights of Withdrawal in Consumer Contract Law: The Right to Change Your Mind?*, 29 Pa. St. Int’l L. Rev. 671, 678–83 (2011) (questioning the utility of imposing mandatory withdrawal rights, due to their ability to undermine sellers’ incentives to grant withdrawal rights anyway for the purposes of “creating trust and attracting consumers”). [↑](#footnote-ref-37)
38. Johnston, *supra* note 3, at 873–74; Zamir & Teichman, *supra* note **Error! Bookmark not defined.**, at 291; Ben-Shahar & Posner, *supra* note 33, at 120–21; Becher & Zarsky, *supra* note 5, at 73. [↑](#footnote-ref-38)
39. *See, e.g.*, Tiffany Hsu, *L.L. Bean, Citing Abuse, Tightens its Generous Policy on Returns*, N.Y. Times (Feb. 9, 2018), available at https://www.nytimes.com/2018/02/09/business/ll-bean-returns-policy.html (explaining L.L. Bean’s decision to amend its lifetime return policy to a one-year return policy, with proof of purchase, in response to consumer abuse); *see also* Khadeeja Safdar & Laura Stevens, *Banned From Amazon: The Shoppers Who Make Too Many Returns*, Wall Street J. (May 22, 2018), available at https://www.wsj.com/articles/banned-from-amazon-the-shoppers-who-make-too-many-returns-1526981401 (explaining Amazon’s decision to close accounts of consumers taking advantage of its generous return policy). [↑](#footnote-ref-39)
40. *See. e.g.*, Ariella Gintzler, *L.L. Bean’s Lifetime Return Policy is No More*, Outside (Feb. 9, 2018), available athttps://www.outsideonline.com/2280581/ll-bean-tightens-generous-return-policy (describing a class action suit filed against L.L. Bean in response to the changes it made to its return policy); Shirley v. L.L. Bean, Inc., No. 3:18-cv-02641 (filed May 4, 2018) [class action complaint], available at https://www.documentcloud.org/documents/4475198-LLBean2.html. At the same time, some stores are adopting more generous return policies. *See* Donna L. Montaldo, *Target Bucks the Trend with a New Super Lenient Return Policy*, The Balance (Aug. 6, 2019), available at https://www.thebalance.com/target-s-new-return-policy-939859. [↑](#footnote-ref-40)
41. *See, e.g.*, Jeanine Skowronski, *Shopper Outrage: Refund Loopholes*, The Street (Feb. 28, 2011), available at https://www.thestreet.com/slideshow/12795160/1/shopper-outrage-refund-loopholes.html. [↑](#footnote-ref-41)
42. Retailing and marketing studies offer some insight into retailers’ return policies, but they typically focus on the relationship between return policy leniency and consumers’ purchasing and withdrawal decisions. For a meta-analytic review of this literature, see Narayan Janakiraman, Holly A. Syrdal, & Ryan Freling, *The Effect of Return Policy Leniency on Consumer Purchase and Return Decisions: A Meta-analytic Review*, 92 J. Retailing 226 (2016); *see also* Scott Davis, Michael Hagerty, & Eitan Gerstner, *Return Policies and the Optimal Level of Hassle*, 50 J. Econ. & Bus. 445 (1998). [↑](#footnote-ref-42)
43. Field experiments are increasingly used in legal scholarship, and are considered “one of the most powerful empirical tools for identifying causal relationships.” *See* Jacob Kopas & Dane Thorley, *Experiments in the Court: The Legal and Ethical Challenges of Running Randomized Field Experiments in the Courtroom* (June 20, 2018), available at https://ssrn.com/abstract=2994298. For literature on the methodological value and limitations of field experiments, see, *e.g.*, Kosuke Imai, Luke Keele, Dustin Tingley, & Teppei Yamamoto, *Unpacking the Black Box: Learning about Causal Mechanisms from Experimental and Observational Studies*, 105 Pol. Sci. Rev. 765 (2011); Kosuke Imai, Dustin Tingley, & Teppei Yamamoto, *Experimental Designs for Identifying Causal Mechanisms*, 176 J. Royal Statistical Soc’y 5 (2012); Alan S. Gerber & Donald P. Green, Field Experiments: Design, Analysis, And Interpretation (2012). For articles specifically addressing the use of field studies in legal research, see, e.g., Donald P. Green & Dane R. Thorley, *Field Experimentation and the Study of Law and Policy*, 10 Annual Rev. L. & Soc. Sci. 53 (2014) (providing an overview of the history of field experiments in the legal context and a number of helpful examples of well-conducted studies); Michael Abramowicz, Ian Ayres, & Yair Listokin, *Randomizing Law*, 159 U. Pa. L. Rev. 929 (2011); Laurens Walker, *Protecting Federal Civil Rules: A Proposal for Restricted Field Experiments*, 51 L. & Contemp. Probs. 67, 67 (1988); D. James Greiner & Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 Annual Rev. L. & Soc. Sci. 295 (2016). [↑](#footnote-ref-43)
44. The audit technique used in this study is similar to audit techniques used in discrimination studies. For discrimination studies using an audit technique, *see, e.g.*, Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Care Negotiations*, 104 Harv. L. Rev. 817 (1991); Ian Ayres and Peter Siegelman, *Race and Gender Discrimination in Bargaining for a New Car*, 85 Am. Econ. Rev. 304 (1995). Audit studies necessarily involve a certain degree of deception, and therefore inevitably raise important questions of research ethics. On the ethical concerns raised by field studies, see, e,g., Kopas & Thorley, *supra* note 40; Abramowicz et al., *supra* note 44. The study’s design sought to minimize the effects of the tests on sellers by conducting tests at off-peak hours (mid-afternoons during the week). In addition, in accordance with IRB instructions (and with the law in Illinois), the testers’ conversations with store clerks were not recorded or videotaped. Rather, testers recorded their audit results using survey forms prepared by the Author. [↑](#footnote-ref-44)
45. In a larger sample of 192 retail stores (randomly selected from the ReferenceUSA and Hoover’s Company Directories’ databases for a separate study), receipt requirements were mentioned in 84% of the return policies. For evidence that receipt requirements are perceived as a hassle by consumers, *see, e.g.*, Janakiraman, Syrdal, & Freling, *supra* note 43. [↑](#footnote-ref-45)
46. While the study’s sample includes luxury stores (as long as at least one of the items they offered in store met the price criterion), it can admittedly shed only limited light on the return practices of the most luxurious designer stores. I leave this issue for future research. [↑](#footnote-ref-46)
47. Python was used to scrape the stores’ websites. Some stores blocked access to their websites, and were therefore manually coded. Coders and programmers were instructed to derive the median prices of the items based on clothing items only, in order to keep the analysis tractable across stores with different offerings. Median prices were chosen instead of mean prices, as mean prices, unlike median prices, are affected by outliers (i.e., extremely expensive or very cheap products). [↑](#footnote-ref-47)
48. Firm data, including revenue, public versus private status, and years since establishment, were obtained primarily from Bloomberg Law and Hoover’s Company Directories’ databases. Age refers to the number of years of operation since establishment. [↑](#footnote-ref-48)
49. This generally reflects the market share division between local and chain stores in Chicago. The most current study has found that individual retailers control about 30% of the market share. *See* Nicole Leinbach-Reyhle, *Celebrating Independent Retailers: Their Surprisingly Strong Future*, Forbes (July 3, 2014), available at https://www.forbes.com/sites/nicoleleinbachreyhle/2014/07/03/celebrating-independent-retailers-their-strong-future/. [↑](#footnote-ref-49)
50. Rather than matching tester A with tester B for all tests, A was sometimes matched with B, sometimes with C, and so on. [↑](#footnote-ref-50)
51. For consistency, testers were instructed to audit the stores on weekdays in the afternoons. [↑](#footnote-ref-51)
52. Research assistants (purchasers) were sent to purchase the items in advance. The items were then returned by different members of the research team—the testers. This design was chosen for several reasons. First, it allowed for the purchase of two identical items from each store, so that items would not vary within stores. Second, it mitigated the concern that in some stores, store clerks would remember the person making the return whereas in others, testers would encounter different store clerks. Still, this design raises the possibility that in some stores, particularly the smaller ones, store clerks were more suspicious of the testers making the returns because the clerks did not remember the testers making the purchase at the store in the first place. It could therefore be the case that, at least in the smaller stores, the findings underestimate the likelihood of a gap. The results should therefore be interpreted as testing whether, in view of reputational constraints, sellers depart from the formal policy even when they do not remember the particular, one-time buyer. [↑](#footnote-ref-52)
53. The study was designed to minimize differences in treatment *among* stores. For that purpose, purchasers were instructed to buy a clothing accessory (i.e., a hat, gloves, socks, scarves, purses, or bags). If no accessories were available, they were instructed to buy a shirt, pants, or another clothing item. They were specifically instructed to refrain from buying underwear, swimwear, jewelry, electronic devices, “final”/clearance/sale items, or any item that was specifically not eligible for returns according to the stores’ formal return policies. Prices were kept constant at between $20 to $30. [↑](#footnote-ref-53)
54. It is possible that store clerks’ willingness to depart from the formal policy would vary depending on the reason offered by testers. For example, if testers had said that they had bought the wrong size, or received the item as a gift and did not like it, store clerks may have responded differently. The generic excuse used in the study was meant to allow testers to request a refund, rather than merely an exchange or store credit. Unlike returning a gift or exchanging an item for a different size, explicitly explaining that the return is because they do not need the product makes asking for a cash refund rather than an exchange or store credit more credible and reasonable. [↑](#footnote-ref-54)
55. Testers obtained inconsistent return outcomes in 15% (n = 20) of the stores sampled (n = 132). [↑](#footnote-ref-55)
56. In order to keep testers’ fictitious names fixed across stores, the names used in this study were Emily Baker and Allison O’Brian. These names were identified as “white-sounding” in recent discrimination field experiments. *See, e.g.*, Marianne Bertrand and Sendhil Mullainathan, *Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, 94 Am. Econ. Rev. 991 (2004); Benjamin Edelman, Michael Luca, & Dan Svirsky, *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, 9 Am. Econ. J. 1 (2017). [↑](#footnote-ref-56)
57. In 4% of the stores (n = 4), store clerks asked testers to show an ID to process the return. In these cases, different testers came to the stores with IDs and provided their real names. The return outcomes of the latter pair of testers (with IDs) were then used for the analysis. The analysis does not substantially change if these stores are excluded from the sample. [↑](#footnote-ref-57)
58. For the purpose of documenting in-store signs and receipts, research assistants (RAs) were sent to the sampled stores. The RAs were instructed to look for a return policy sign and take a clear photo of any sign they managed to locate. Subsequently, they purchased a clothing item or accessory in each store and scanned the receipt. [↑](#footnote-ref-58)
59. Whenever one or more of the informational sources—receipts, in-store signs, or websites—included a receipt requirement or a “no refund” policy term, the return policy was coded accordingly. The two coders were in agreement 91% of the time. Whenever the two coders were not in unanimous agreement about the proper classification to assign to a policy, a third RA coded the policy and the coding given by the two-person majority was used. Cohen’s kappa measure of inter-grader disparities is 0.89. [↑](#footnote-ref-59)
60. Importantly, the sample did not include any stores that did not allow for refunds, while at the same time allowing for non-receipted returns for exchanges or store credits. Similarly, by design, the sample did not include any stores that did not explicitly require receipts (either for a refund or for any return or exchange). [↑](#footnote-ref-60)
61. Under a simple regression of initial return outcomes (defined as a categorical variable assigned 0 if returns are denied, 1 if exchange or store credit is offered, and 2 if refund is provided) on store policy type (harsh, moderate, or lenient), *b* = 0.26, SE = 0.14, *p* < 0.1 for moderate policies, and *b* = 0.97, SE = 0.13, *p* < 0.001 for lenient policies (with harsh policies serving as the reference category). The difference between lenient policy stores and harsh policy stores remains significant once store demographics (i.e., age and size) are controlled for, but the difference between moderate and harsh policy stores becomes insignificant. [↑](#footnote-ref-61)
62. [↑](#footnote-ref-62)
63. 61 Stores with moderate policies were also marginally significantly more likely (at the 10% level) to depart from their policies than were stores with lenient return policies. Yet, this finding should not be surprising in view of the fact that lenient policy stores formally allow consumers to return non-receipted items for exchanges or store credits, whereas the other stores do not. Indeed, in terms of overall return outcomes, consumers obtained significantly higher rates of returns among the lenient policy stores (92%) than among the mainstream stores (36%), *b =* 0.83, SE = 0.09, *p* < 0.001. In this study, only one of the harsh policy stores (4%) belongs to a chain, while the remaining 96% belong to independent retailers. The correlation between store type (chain or local) and policy type is 0.75. A regression of formal policy type on stores’ characteristics reveals that local stores are significantly more likely to adopt harsh policies than are chain stores, even when controlling for the store’s age and size and for whether the store is public or private (*b =* 0.22, SE = 0.05, *p* < 0.001). [↑](#footnote-ref-63)
64. In the script, consumers were instructed to complain and to ask to speak to the manager if denied the return, but they were instructed not to mention the contractual language. In future research, it would be worthwhile to explore how sellers react when testers/consumers point out to store officials the language that entitles them to the denied benefit. [↑](#footnote-ref-64)
65. This prediction is consistent with previous writing on the gap. *See, e.g.*, Gillette 2004, *supra* note 3, at 707 (hypothesizing, without providing empirical data, that sellers might distinguish between insistent and non-insistent buyers); Johnston, *supra* note 3, at 881 (suggesting that firms use customer complaints to determine the value of the concessions to consumers). For research about how firms more generally treat consumers who complain, *see*, *e.g.*, Rory Van Loo, *The Corporation as Courthouse*, 33 Yale J. on Reg. (2016) (arguing that corporations’ responses to consumers’ complaints are influenced by consumers’ characteristics, including past behavior, social influences, and buying power). [↑](#footnote-ref-65)
66. *b* = 0.35, SE = 0.09, *p* < 0.001. [↑](#footnote-ref-66)
67. *b* = 0.27, SE = 0.11, *p* < 0.05. More particularly, lenient stores were significantly more likely to offer exchanges or store credits than both moderate (*b* = 0.26, SE = 0.11, *p* < 0.05) and harsh policy stores (*b* = 0.46, SE = 0.12, *p* < 0.001). They were also marginally significantly more likely than harsh policy stores to offer refunds (*b* = 0.16, SE = 0.09, *p* < 0.1). [↑](#footnote-ref-67)
68. Under a regression of return outcomes (where 1 = pro-consumer gap, 0 otherwise) on the complaining treatment, *b* = 0.14, SE = 0.07, *p* < 0.05. [↑](#footnote-ref-68)
69. A paired t-test analysis reveals that for refund the effect was marginally significant: t = 1.64, df = 94, *p* = 0.052; for non-receipted returns: t = 1.78, df = 55, *p* < 0.05. [↑](#footnote-ref-69)
70. The effect of complaining on the gap among the moderate policy stores is significant at the 10% level, *b* = 0.21, SE = 0.12, *p* = 0.09. For moderate stores, complaining significantly improved the chances of receiving a non-receipted return or exchange (*p* < 0.05), while its effect on refund rates was marginally significant (*p* < 0.1). The effect of complaining on the gap among the harsh policy stores is significant at the 10% level, *b* = 0.21, SE = 0.11, *p* = 0.06. For harsh policy stores, complaining significantly improved the chances of receiving a non-receipted return or exchange (*p* < 0.05) but there was no effect on refund outcomes. [↑](#footnote-ref-70)
71. *b =* 0.02, SE = 0.09, *p* = 0.8. [↑](#footnote-ref-71)
72. Before embarking on the field experiment, I interviewed fifteen store clerks from Chicago about their interactions with customers and management. Store clerks were recruited for the interviews through two online platforms: Craigslist and University of Chicago’s Marketplace. The interviews were semi-structured, and focused mainly on exploring whether and when store clerks are given discretion to depart from the stores’ formal policies in practice. Please see interview questionnaire in Appendix \_\_\_. [↑](#footnote-ref-72)
73. Interview #12 with Sarah (Abercrombie) (interview script on file with the Author). [↑](#footnote-ref-73)
74. Interview #7 with Leila (Saks Fifth Avenue) (recorded interview on file with the Author). [↑](#footnote-ref-74)
75. Interview #4 with Michael (Bally shoe store) (recorded interview on file with the Author). [↑](#footnote-ref-75)
76. This, in turn, might lead consumers who do not value the concessions that certain stores offer to switch to buying at stores that do not provide concessions, but offer lower prices. Still, some stores will continue to offer concessions as long as a sufficiently large group of consumers values these concessions and is willing to pay higher prices for them. These stores will be unlikely to change their formal policies, however, as they will still need to fend off opportunistic consumers by relying on the harsh paper terms to deny their claims. [↑](#footnote-ref-76)
77. While chain stores were not significantly more likely to provide refunds, they were more likely (at the 10% significance level) to accept non-receipted returns at the final (post-complaining) stage. [↑](#footnote-ref-77)
78. This effect is driven by older stores’ greater likelihood of providing refunds notwithstanding testers’ failure to show a receipt. Older stores were not significantly more likely to allow for non-receipted returns, as Models Five and Six show. [↑](#footnote-ref-78)
79. In the case of return policies, for example, strict “no refund” policies are often adopted by local sellers. It is reasonable to assume that these sellers typically incur high depreciation costs from facilitating returns because they have less ability to resell used items or return them to the supplier. [↑](#footnote-ref-79)
80. , , ,, casting doubt on the ability of the market forces of reputation and competition to adequately incentivize sellers to offer favorable terms. [↑](#footnote-ref-80)
81. *See* George A. Akerlof, *The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*, 84 Q. J. Econ. 488 (1970). [↑](#footnote-ref-95)
82. *See, e.g.*, Gillette, *supra* note 3, at 706 [↑](#footnote-ref-96)
83. *See, e.g.*, Meirav Furth-Matzkin, *The Harmful Effects of Unenforceable Contract Terms: Experimental Evidence*, 4 Ala. L. Rev. 1032 (2019) (finding that most tenants in an experimental survey acquiesced to whatever was stated in their lease terms, even when these terms contained unenforceable liability disclaimers); Meirav Furth-Matzkin & Roseanna Sommers, *Consumer Psychology and the Problem of Fine Print Fraud*, 72 Stan. L. Rev. \_\_ (forthcoming 2020); Wilkinson-Ryan 2017, *supra* note 82. [↑](#footnote-ref-97)
84. *See, e.g.*, Tess Wilkinson-Ryan, *Intuitive Formalism in Contract*, 163 U. Pa. L. Rev. 2109 (2015); Tess Wilkinson-Ryan & David A. Hoffman, *The Common Sense of Contract Formation*, 67 Stan. L. Rev. 1269, 1281–98 (2015) (finding that laypeople put excessive weight on written terms compared to oral agreements, believe that contracts are formed primarily through formalities such as signature and payment (even though contract law does not require such formalities for a contract to be formed), and generally feel obligated to abide by terms that follow formalized assent processes); Yuval Feldman & Doron Teichman, *Are All Contractual Obligations Created Equal?*, 100 Geo. L.J. 5, 5 (2012) (arguing that laypeople feel they are bound to the signed contract due to “moral commitments, social norms, and motivated reasoning”). [↑](#footnote-ref-98)
85. *See, e.g.*, Furth-Matzkin 2019, *supra* note 95; Furth-Matzkin & Sommers, *supra* note 95; Tess Wilkinson-Ryan, *A Psychological Account of Consent of Fine Print*, 99 Iowa L. Rev. 1745 (2014) (finding that people maintained that it was fair to hold signees to fine print terms they had not read, even if the terms were buried in a contract that they believed to be unreasonably lengthy). [↑](#footnote-ref-99)
86. *See, e.g.*, Furth-Matzkin 2019, *supra* note 95; Furth-Matzkin & Sommers, *supra* note 95; Wilkinson-Ryan 2017, *supra* note 82. [↑](#footnote-ref-100)
87. Furth-Matzkin 2019, *supra* note 95. In a similar vein, Dennis P. Stolle & Andrew J. Slain find that consumers are reluctant to file meritorious suits if their contracts include legally dubious disclaimers of tort liability. *See* Dennis P. Stolle & Andrew J. Slain, *Standard Form Contracts and Contract Schemas: A Preliminary Investigation of the Effects of Exculpatory Clauses on Consumers’ Propensity to Sue*, 15 Behav. Sci. & L. 83 (1997). [↑](#footnote-ref-101)
88. Furth-Matzkin & Sommers, *supra* note 95. [↑](#footnote-ref-102)
89. Several scholars have suggested that even when sellers grant concessions to consumers, legal entitlements are preferable because, “from a welfare perspective receiving something as a matter of entitlement is more conductive to one’s welfare . . . than receiving the same thing as a favor” (Zamir, *supra* note 3, at 2100; Zamir & Teichman, *supra* note 7 at 311–12). For a similar argument, see Daphna Levinson-Zamir, *In Defense of Redistribution through Private Law*, 91 Minn. L. Rev. 326, 358–65 (2006). [↑](#footnote-ref-103)
90. Johnston, *supra* note 3, at 858. [↑](#footnote-ref-104)
91. For a similar assertion, *see Gillette, supra* note 3, at 706. [↑](#footnote-ref-105)
92. Regressive distributional concerns might yield different policy prescriptions than an analysis that addresses only overall efficiency (or welfare). *See, e.g.*, Levinson-Zamir, *supra* note 102; Lee Fennell & Richard McAdams, *The Distributive Deficit in Law and Economics*, 100 Minn. L. Rev. 1051 (2016). [↑](#footnote-ref-106)
93. *See, e.g.*, John T. Jost, Mahzarin R. Banaji, & Brian A. Nosek, *A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo*, 25 Pol. Psych. 881 (2004); Brett W. Pelham & John J. Hetts, “Underworked and Overpaid: Elevated Entitlement in Men’s Self Pay,” 37 J. Exper. Soc. Psych. 92 (2001); Paul K. Piff, *Wealth and the Inflated Self: Class, Entitlement and Narcissism*, 40 Personality & Soc. Psych. Bulletin 34 (2014); Candace N. Joyner, *Entitled to Expect: System Justification Theory, Socioeconomic Status, and the Ultimatum Game* (2017) (showing, based on an ultimatum game experiment, that socioeconomic status predicts expectations and sense of entitlement); Brenda Major, *From Social Inequality to Personal Entitlement*, 26 Advances in Exper. Soc. Psych. 293 (1994); Jie Hu, Yuan Cao, Philip R. Blue, & Xiaolin Zhou, *Low Social Status Decreases the Neural Salience of Unfairness*, 8 Frontiers in Behav. Neurosci. 402 (2014); Laurie T. O’Brien & Brenda Major, *Group Status and Feelings of Personal Entitlement: The Roles of Social Comparison and System-justifying Beliefs*, in John T. Jost, Aaron C. Kay, & Hulda Thorisdottir, Series in Political Psychology: Social and Psychological Bases of Ideology and System Justification(2009); Annette Lareau, *Invisible Inequality: Social Class and Childrearing in Black Families and White Families*, 67 Am. Soc. Rev. 747 (2002)(suggesting that middle and upper income white families raise their children with a sense of entitlement and assertiveness enabling them to gain what they want later in life, while childrearing strategies among the lower class populations and racial minorities tend to result in a lack of assertiveness or lack of a sense of entitlement, thereby limiting their access to educational and job opportunities later in life). [↑](#footnote-ref-107)
94. *See, e.g.*, Laurie T. O’Brien, Brenda N. Major, & Patricia N. Gilbert, *Gender Differences in Entitlement: The Role of System-Justifying Beliefs*, 34 Basic & Applied Soc. Psych. 136 (2012); Patricia J. Williams, *Alchemical Notes: Reconstructing Ideals from Deconstructed Rights*, 22 Harv. C.R.-C.L. L. Rev. 401 (1987). [↑](#footnote-ref-108)
95. *See, e.g.*, Furth-Matzkin & Sommers, *supra* note 95 (finding that nonwhite participants were inclined to see the consumer as more bound by the fine print than were white participants); Matthew A. Seligman, *The Error Theory of Contract*, 78 Md. L. Rev. 147 (2018) (showing that “people with less education or lower household income are significantly more likely to have false beliefs about contract remedies than people with more education or higher household income”); Jessica M. Choplin, Debra Pogrund Stark, & Jasmine N. Ahmad, *A Psychological Investigation of Consumer Vulnerability to Fraud: Legal and Policy Implications*, 35 L. & Psychol. Rev. 61, 94 (2011) (presenting findings indicating that “those with lower status are more likely to agree and accept senseless explanations . . . . Those with higher status seem to be more vigilant, perhaps in an effort to protect their higher status”). This evidence supports assumptions made in prior contract law literature. *See, e.g.*, Jeffrey L. Harrison, *Class, Personality, Contract, and Unconscionability*, 35 Wm. & Mary L. Rev. 445 (1994) (explaining that xxxxxxx [↑](#footnote-ref-109)
96. Interview #13 with David (rug store) (recorded interview on file with the Author). [↑](#footnote-ref-110)
97. Preliminary evidence from a field experiment I administered supports this hypothesis. In the experiment, black and white male and female testers were sent to make non-receipted returns in 60 retail stores located in downtown Chicago. The results reveal a large and significant racial gap in return outcomes. Black customers were almost twice as likely to be denied a return as were white customers, and the difference was larger after asking to speak to the store’s manager. [↑](#footnote-ref-111)
98. 42 U.S.C. §§ 1981(a), 1981(b). *See also* Hill v. Kookies, Inc., 1999 WL 608713 (N.D. Ill. 1999) (applying Illinois law). [↑](#footnote-ref-112)
99. Outcomes were coded as “return denied” if the store clerk either refused to offer any concession when testers came without a receipt or only allowed testers to exchange the item for a different size or color. [↑](#footnote-ref-113)
100. These figures include cases where store clerks allowed testers only to exchange the item for a different size or color. [↑](#footnote-ref-114)