The “Paper Deal—Real Deal” Gap

in Consumer Markets:

Evidence from A Field Experiment

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

# Introduction

Most of our everyday transactions are governed by standard form contracts. Virtually every firm selling goods or services uses boilerplate 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, and various other elements of the sales relationship. Although these standardized provisions considerably facilitate transactions,[[2]](#footnote-2) scholars and commentators have expressed concern that some of these terms might be harmful to consumers.[[3]](#footnote-3)

In response to these concerns, several scholars have suggested that, at least in competitive markets, sellers’ enforcement of their contractual provisions may be more lenient and flexible in practice.[[4]](#footnote-4) As these scholars argue, while the existence of rigid and unconditional terms on paper enables sellers to fend off buyers likely to exploit 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 customers with just claims.[[5]](#footnote-5) I refer to this approach as the gap theory, because it suggests that sellers sensitive to reputational consequences will depart from seemingly rigid 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.[[6]](#footnote-6)

While the potential discrepancies between consumer contracts on paper and in action have attracted some scholarly attention, academics continue to debate both their prevalence in consumer markets and their normative consequences. In particular, some scholars have questioned the actual ability of competitive forces, and reputation in particular, to deter sellers from adhering to the four corners of the agreement.[[7]](#footnote-7)

This debate has significant policy implications. If reputational constraints are sufficient to discipline sellers 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.[[8]](#footnote-8) 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.[[9]](#footnote-9) Yet, despite the significant practical implications of this debate, there has nonetheless been little systematic empirical research on how sellers actually enforce their contractual provisions.[[10]](#footnote-10)

This Article represents an important step towards mitigating this deficiency. Addressing the issue of product returns as a first test case, it reports the results of an original, large-scale field experiment covering ninety-five retail stores in Chicago. For the purposes of the experiment, six testers (all white and American females of similar age) were hired and trained to follow a uniform bargaining script before being sent to return clothing items without receipts to stores with a clear receipt requirement for returns. Each store was audited by two testers, resulting in 190 independent observations. This audit study examines whether, when, and to what extent retailers depart from seemingly rigid, unconditional return policy terms in favor of consumers.

Another key goal of the study was to test the effect of consumers’ bargaining behavior on sellers’ departure decisions in order to assess whether sellers routinely depart from the written terms, or do so mainly in those cases where consumers are insistent. The study hypothesized that sellers may be more lenient towards consumers once they complain, under the assumption that the reputational costs from leaving consumers who complain dissatisfied exceed the costs of extra leniency to the sellers. To test this hypothesis, testers were instructed to escalate their assertiveness throughout the bargaining process, ultimately asking to speak to the store’s manager. Testers then reported the outcomes of the attempted returns both at the initial stage, before complaining, and at the final, post-complaining stage.

The findings reveal that, across a wide variety of stores—both chain and local, luxurious and casual—a significant proportion of sellers (22%) departed from their formal contractual requirements in favor of consumers upon consumers’ initial requests, and that sellers were almost twice as likely to depart from their requirements once consumers complained.

The study also tested the interactions between retailers’ policies on paper and their behavior on the ground, by comparing sellers’ departure patterns across stores with varying degrees of formal policy leniency. The results reveal that stores with the harshest policies on paper were the least likely to depart from their formal requirements, plausibly due to their uniquely high costs of accepting returns.

Notably, some terms were “stickier” than others, with sellers very rarely departing from stickier terms to consumers’ advantage.[[11]](#footnote-11) Sellers were reluctant to apply pro-consumer gaps in contexts where gaps would arguably be expected to be the most prevalent: the cases of unequivocally rigid terms completely denying refunds.[[12]](#footnote-12) Importantly, these policies of completely denying refunds are also the contexts where gaps may be needed the most, because these harsher terms significantly limit consumers’ withdrawal rights.

Finally, stores’ characteristics significantly interacted with their departure decisions, with high-end, established, and chain stores significantly more likely to exercise tailored forgiveness than were more casual, less established, and local stores.

These findings raise important questions concerning the desirability of regulatory interventions in consumer markets. Many scholars have argued that the gap between the “paper deal” and the “real deal” renders legal intervention “less necessary.”[[13]](#footnote-13) 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.”[[14]](#footnote-14) Yet, the findings reveal that this practice is hardly universal, and that reputational constraints on businesses are not always sufficient to prevent sellers’ enforcement of their rigid contract terms to the letter against consumers.[[15]](#footnote-15) Consequently, the presumed gap between the “paper deal” and the “real deal” should not obviate the need to consider regulating the terms of consumer contracts when those seem to be excessively burdensome or harmful to consumers, at least in the absence of empirical data that show that pro-consumer gaps are prevalent in the particular context under investigation.

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 seemingly unconditional, standardized terms and their actual implementation might generate distortions and regressive distributive effects.

As the findings reveal, 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 controlling for store characteristics. This unexplained variation might harm consumers’ ability to distinguish between stores that strictly adhere to their formal return policies 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 likely 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.[[16]](#footnote-16) As the findings reveal, store employees are significantly more likely to exercise discretionary, tailored forgiveness once consumers complain.[[17]](#footnote-17) 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 many (if not most) good-faith consumers. Accumulating empirical evidence suggests that consumers tend to be contract formalists, with most believing that whatever the contract says is the final word.[[18]](#footnote-18) 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.[[19]](#footnote-19)

Furthermore, this complaint-based mechanism is likely to have regressive distributional effects.[[20]](#footnote-20) Quantitative evidence and original interviews with store clerks indicate that lower income consumers and minority group members are less likely to complain than are higher income consumers or those belonging to majority groups.[[21]](#footnote-21) 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.[[22]](#footnote-22) 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, by introducing the gap theory and discussing the importance of shedding empirical light on the contract in action. 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 as to 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

According to the gap theory, clear and unconditional standard form contract terms may allow sellers to exercise tailored discretion on the ground, while segmenting between opportunistic (or low-value) and non-opportunistic (or high-value) consumers.[[23]](#footnote-23)

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.[[24]](#footnote-24)

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, “pro-seller” terms against low-value or opportunistic consumers (defined as consumers whose benefits from the concession are lower than the cost to the seller of granting it), while departing from the formal agreement to the benefit of the remaining, high-value consumers (defined as consumers whose benefits from the concession exceed the cost to the seller of granting it).[[25]](#footnote-25)

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.[[26]](#footnote-26) 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 (and non-opportunistic) consumers.

This descriptive theory 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, seemingly “one-sided” terms, since these terms will allow them to behave efficiently (on account of their discretion to forgive).[[27]](#footnote-27)

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.[[28]](#footnote-28)

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 vis-à-vis consumers. These scholars rely mainly on the manifold evidence that informational flows in consumer markets are far from perfect, and that reputational information, in particular, is neither reliable nor accurate.[[29]](#footnote-29)

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.”[[30]](#footnote-30) Clayton Gillette has suggested that “we might be wary of systematically invalidating clauses in SFCs [standard form contracts] that permit sellers latitude in enforcement on the assumption that sellers use such clauses selfishly.”[[31]](#footnote-31) And Lucian Bebchuk and Richard Posner have similarly proposed 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.[[32]](#footnote-32) At the same time, those who believe that sellers might insist on adhering to the contract even when it is not socially desirable to do so call for “more serious and systematic consideration of the use of mandatory regulation of the content of transactions.”[[33]](#footnote-33)

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 its practical implications, empirical investigation into the gap is surprisingly lacking. This Article sheds light on the contract in action through the test case of product returns.

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 without any meaningful empirical investigation into sellers’ return practices on the ground.[[34]](#footnote-34) 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.[[35]](#footnote-35) In a significant number of cases, consumers who purchase goods or services end up regretting their purchases after the fact.[[36]](#footnote-36) In 2016, for example, Americans returned $260 billion in merchandise (or eight percent of all purchases made in the United States) to retailers.[[37]](#footnote-37)

The ability to withdraw from consumer transactions is important to consumers.[[38]](#footnote-38) It is also considered welfare-enhancing, as 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 ability to withdraw from consumer transactions, consumers are not in most cases[[39]](#footnote-39) entitled to return non-defective goods to the seller under current U.S. law.[[40]](#footnote-40) 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,[[41]](#footnote-41) others believe that statutory intervention may not be warranted due to market incentives already in place.[[42]](#footnote-42) 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.[[43]](#footnote-43) Yet, in recent years, stores have begun imposing more restrictions on consumer returns.[[44]](#footnote-44) These shifts in firms’ policies have spurred consumer outrage,[[45]](#footnote-45) reviving the debate over the need to regulate consumers’ withdrawal rights.[[46]](#footnote-46)

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.[[47]](#footnote-47) 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.[[48]](#footnote-48) The study uses an audit technique, in which testers were sent to return clothing items to 95 retail stores located in Chicago.[[49]](#footnote-49) 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.[[50]](#footnote-50) 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.[[51]](#footnote-51)

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 a store 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.[[52]](#footnote-52)

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

The stores’ mean annual revenues is $434 million (SD = $2.3 million). The average company age is 52 years (SD = 42 years).[[53]](#footnote-53) Seventy-one percent of the stores in the sample are national retail chains, and 29% are local stores (operated by independent retailers).[[54]](#footnote-54) The average median price of items among the sampled stores is $92 (SD = $208).

## 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. The testers audited 95 retail stores in Chicago. They 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.[[55]](#footnote-55)

In each store, testers attempted to return a clothing item that had been purchased in advance[[56]](#footnote-56) in its original packaging and condition, with tags attached, but *without the receipt*.[[57]](#footnote-57) 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.[[58]](#footnote-58) 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. After leaving each store, all testers filled out a detailed report, in which they described the outcomes of the attempted returns.

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 an informal 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.[[59]](#footnote-59) Both testers in a pair visited the same store, usually within a few days of one another. The analysis includes only those stores in which both testers obtained identical return outcomes.[[60]](#footnote-60)

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

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.[[61]](#footnote-61) If asked for an identification card, testers would say that they did not have one with them.[[62]](#footnote-62)

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.[[63]](#footnote-63)

## 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.[[64]](#footnote-64) Stores were subsequently classified by two independent research assistants as having harsh, moderate or lenient paper policies.[[65]](#footnote-65)

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.[[66]](#footnote-66) Table 1 summarizes this classification method.

*Table 1. Paper Policy Types*

|  |  |  |
| --- | --- | --- |
|  | 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 |

Table 2 reports the basic company characteristics for each type of policy store. Age is defined as 2019 minus the year of establishment. I also report the proportion of chain stores and the average median prices in each category.

*Table 2. Summary Statistics*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean Age  (SD) | Mean of Median Prices  (SD) | Proportion of Chain |
| Harsh Policy Stores  (n =23) | 21  (14) | 198  (423) | 0.04 |
| Moderate Policy Stores  (n = 33) | 46  (32) | 70  (61) | 0.85 |
| Lenient Policy Stores  (n = 39) | 71  (48) | 59  (67) | 0.97 |

## 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 3 describes this analysis.

*Table 3. 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 Gaps

Figure 1 describes what happened when testers initially asked for a refund—overall and across the different types of stores.

*Figure 1. Return Outcomes at the Initial Stage.*

*Figure 1.* 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 1 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.[[67]](#footnote-67)

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 2 illustrates the prevalence of pro-consumer gaps in each type of store.

*Figure 2: Gaps across Stores*

*Figure 2. Gaps at the initial (pre-complaining) stage*. The upper bar represents the proportion of gaps overall (n = 95), while the lower three bars represent the proportion of gaps across stores with different policies.

As figure 2 shows, stores with moderate return policies were much more likely than stores with harsh return policies to exhibit a pro-consumer gap.[[68]](#footnote-68) 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.[[69]](#footnote-69) These local stores probably incur higher depreciation costs from accepting returns, in view of their lower ability to resell items or to return them to the supplier. They plausibly adopted harsh policies on paper in view of these uniquely high costs, and were consequently reluctant to depart from their policies in practice.

Unexpectedly, a small subset of the lenient policy stores (8%) 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. It is even more surprising that the testers in these stores were denied the returns even after asking to speak to the managers. I leave these findings for future investigation.[[70]](#footnote-70)

## Final Return Outcomes and the Effect of Complaining

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.[[71]](#footnote-71) 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 3 reports testers’ initial and final return outcomes across the different policy stores.

*Figure 3: Initial and Final Return Outcomes by Policy Type*

*Figure 3. 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 3 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,[[72]](#footnote-72) and moderate stores were significantly more likely to accept such returns than harsh policy stores.[[73]](#footnote-73)

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.[[74]](#footnote-74) 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.[[75]](#footnote-75)

Figure 4 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 4. Pro-Consumer Gaps: Before and After Complaining*

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

As Figure 4 shows, complaining significantly improved testers’ return outcomes among both the harsh and moderate policy stores.[[76]](#footnote-76) Complaining also operated in the expected direction of improving consumers’ outcomes among the lenient policy stores, yet the effect was not significant.[[77]](#footnote-77) 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 sheds light on sellers’ use of the complaint-based mechanism.[[78]](#footnote-78) 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. [[79]](#footnote-79)

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. 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.”[[80]](#footnote-80) 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.[[81]](#footnote-81)

More casual stores, in contrast, build their reputations by offering lower prices. Therefore, they will likely offer no discretionary benefits but charge lower prices.[[82]](#footnote-82)

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 $28). Mainstream stores were defined as all stores with median prices between the 25th and 75th percentiles (i.e., between $28 and $78), and luxury stores were defined as all stores with median prices in the upper 25th percentile (i.e., between $78 and $1850);
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 4 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 4. 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 | 95 | 95 | 95 | 95 | 56 | 56 |
| 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.[[83]](#footnote-83) 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.[[84]](#footnote-84) 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.

Why are more luxurious, established, and chain stores 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.[[85]](#footnote-85) 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. Some consumers care about their withdrawal rights, and are willing to pay more for a better deal (on paper and in practice).

# IV. Discussion & Implications

## Pro-Consumer Gaps Persist, although Practice is not Universal

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 proportion of retailers behaves more leniently towards consumers than their formal policies require.

What can explain this forgiveness on the ground? Some of the observed departures might be the result of store clerk error or a principal-agent problem. In particular, store clerks might depart from the store’s formal policy in an attempt to please consumers even if they do not believe that it is in the best interests of the store to provide the concession, in view of their desire to avoid disputes with consumers. Yet, qualitative interviews with store clerks, as well as the finding that substantial proportions of sellers’ departures occur only at the managerial level, suggest that retailers allow their employees to depart from their formal policy requirements in certain cases.

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.[[86]](#footnote-86)

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 did 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.

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 (64%) 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.[[87]](#footnote-87) 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.

While 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,[[88]](#footnote-88) 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.[[89]](#footnote-89)

Yet, this Article makes a more nuanced claim. It proposes that even when gaps persist, they might distort consumers’ and sellers’ decisions. The next section explains why distortions might occur, and surveys the types of *ex ante* and *ex post* distortions that “paper deal—real deal” disparities might generate.

## Gaps Might Generate Distortions

### *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 they 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.[[90]](#footnote-90) Sellers that systematically deviate from their policies when facing non-opportunistic consumers may be 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.”[[91]](#footnote-91)

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.[[92]](#footnote-92) 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, 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.[[93]](#footnote-93)

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. Hence, the complaint-based mechanism observed in this study might be harmful to consumers. In this study, 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.[[94]](#footnote-94)

This reaction can be attributed to the fact that laypeople are contract formalists.[[95]](#footnote-95) 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.[[96]](#footnote-96) Consumers are often demoralized by harsh and unconditional contractual language, and consequently refrain from bringing claims to the seller.[[97]](#footnote-97) In fact, previous research has shown that consumers rarely question the validity of contracts that disclaim their mandatory rights and remedies.[[98]](#footnote-98) Consumers are similarly unlikely to challenge contracts induced by fraud because they feel bound by contracts they signed.[[99]](#footnote-99)

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.[[100]](#footnote-100)

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.[[101]](#footnote-101) 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.[[102]](#footnote-102) 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 an informal policy, rather than one-time deviations, for the same reasons mentioned earlier. Furthermore, consumers may fail to search for information about sellers’ actual practices online, instead relying on the formal, written agreement or on the seller’s oral assertions.[[103]](#footnote-103)

### *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 sellers’ strategy of combining rigid standard terms with *ex post* tailored forgiveness is welfare enhancing. For example, Jason Scott Johnston has argued that firms’ practice of providing benefits to consumers 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.”[[104]](#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 as a group, 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 aggregate costs of accommodating buyers who complain, of discouraging consumers from buying at the store, and of leavingbuyers who fail to complain disappointed—are lower than the costs of accommodating the claims of all consumers (including the non-insistent and the opportunistic), sellers will adopt the strict term denying the accommodation and depart only selectively in favor of those consumers who complain.[[105]](#footnote-105) Importantly, as the above analysis suggests, sellers may adopt the strict term (combined with *ex post*, complaint-based forgiveness) even if a more lenient term reflects a more efficient risk allocation from a social welfare perspective.

### Distributional Concerns

If consumer assertiveness is correlated with socio-economic status, sellers’ tailored forgiveness might have troubling distributive implications.[[106]](#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.[[107]](#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.[[108]](#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.[[109]](#footnote-109)

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.[[110]](#footnote-110)

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-boomers. […] 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.”[[111]](#footnote-111)

Sellers’ complaint-based segmentation of consumers is likely to exacerbate these distributional effects.[[112]](#footnote-112) In this particular study, stores were nearly twice as likely to depart from the paper contract once consumers complained. If lower income, less educated consumers, and consumers belonging to minority groups, generally feel less entitled, they might be more likely to acquiesce and less likely to complain, compared to consumers with higher socio-economic backgrounds and majority group members.[[113]](#footnote-113)

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.[[114]](#footnote-114) 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).[[115]](#footnote-115)

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

This Article uncovers substantial discrepancies between return policies on paper and in action and illustrates how these discrepancies shape the interactions between sellers and consumers and the resulting market outcomes.

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, the findings at the same time reveal 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; chain stores, luxury stores, and more experienced stores are more likely to exercise tailored forgiveness than local, more casual, and less experienced stores; and consumers who bargain and complain are more likely to obtain concessions than 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 unravel the terms of the “real deal” before concluding that regulatory intervention is unnecessary.

Yet, this Article does not only show 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 they will be able 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. Sellers, at the same time, will not be incentivized to offer more lenient terms in their standard form contracts, even if they 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 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 need to intervene in the contents of the “paper deal” after all, to ensure that it accurately reflects the “real deal.”

In addition, regulators could try to promote information-sharing about on-the-ground leniency.[[116]](#footnote-116) The gap strategy only harms consumers 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 fixed, or at least—their effect may be substantially mitigated.

Finally, this Article suggests that gaps might lead to regressive distributional outcomes, both because insistence and assertiveness are correlated with higher socio-economic status, gender, and race, and because store clerks may apply their on-the-ground discretion discriminatorily.[[117]](#footnote-117) If these concerns are corroborated in future research, it may be warranted to monitor sellers’ behavior through the use of audit techniques similar to the technique used in this study.[[118]](#footnote-118)

Before concluding, the study’s limitations should be acknowledged. 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 specific policy requirements in a specific type of contracts in a particular product market in one city. In future studies, it will be desirable to explore whether and when gaps persist in other types of consumer markets and contracts, and to further investigate the determinants of the gap, the effect of consumer complaints on sellers’ tailoring decisions, and the implications of this practice for consumer welfare.

While there is still much to be discovered about the gap and its normative implications, this Article makes an initial attempt to shift attention from looking almost exclusively at the terms of the “paper deal” to looking more critically at the terms of the “real deal” in consumer markets.

# Appendices

## A. Appendix A: Study Materials

### Interview Questionnaire

[*Interviews were conducted at the University of Chicago Law School in November 2018 in preparation for the field experiment. Fifteen interviewees—all current or former store clerks in Chicago—were recruited using online platforms (Craigslist and UChicago Marketplace). Interviews lasted 30 minutes and interviewees were rewarded $15 for their participation]*

1. Please tell me about your workplace—Where and when did you start working as a salesperson, and for how long?
2. Could you tell me about your training—Did you have any formal training when you started working as a salesperson? Who trained you? What did you learn during training? Were you provided with information about the stores’ return policy?
3. What is the store’s return policy? Did you receive instructions on how to implement the store’s return policy?
4. Do you remember if, during your time working at the store, customers asked to return items they had purchased from the store?
5. What would customers typically ask for when trying to make returns?
6. Did customers need a receipt to make a return?
7. Have customers ever asked you to return an item without a receipt?
8. If so, could you tell me what you would typically do in these cases? What were you supposed to do, according to the store’s management rules?
9. Did customers ever ask you to return an opened item? If so, could you tell me what you would typically do in these cases? What were you supposed to do, according to the store’s management rules?
10. Did customers ever ask you to return a damaged item? If so, could you tell me what you would typically do in these cases? What were you supposed to do, according to the store’s management rules?
11. Do you remember any situations in which you were told you can deviate from the store’s formal policy? If so, could you tell me about those situations?
12. Do you remember any other situations in which the store’s formal policy was different from the actual policy in terms of how it was applied to customers? If so, could you please explain?
13. Do you remember any situations in which customers complained or asked to talk to the store’s manager? What would the manager typically do under these situations?
14. Do you feel that different customers received different treatment when trying to return items to the store or in other situations?
15. Do you feel that some customers—for example African-Americans, Latinos, women, lower-class individuals—received different treatment? If so, could you explain how? And why do you think they received different treatment?
16. How important was it for the store to keep clients satisfied, and why?

### Summarized Instructions to Purchasers

Thank you for agreeing to participate in the study.

This is a study about retail stores’ return practices. You will be asked to go and buy items from stores located in downtown Chicago, while recording various details about the stores and the items purchased (please see the survey form).

You will be provided with a list of stores and their addresses. All stores are located in Chicago. You will be provided with money so that you could pay in cash for the items.

In each store, please buy two items for $20-$30. If there are no items within that price range in the store, please indicate so in the survey form and leave the store.

Please buy a clothing accessory if possible (i.e., a hat, gloves, socks, a scarf, a purse, or a bag). Otherwise, please buy a shirt, pants or another clothing item. Please do not buy underwear, swimwear, jewelry, electronic devices, “last act”/clearance/sale items, or any item that is not be eligible for returns (or will have stricter return requirements) according to the store’s return policy.

When making the purchase, please pay in cash and ask for separate bags and receipts for each item. Please keep the receipts—you will be required to bring the original receipts to the research team.

If the store clerk offers you to join a credit card/membership/sign up for something or asks for any personal details or contact information, please politely refuse and do not provide any such information.

If the store clerk asks if you would like a bag, please say yes and put each item in a separate bag.

While at the store, please see if there is a return policy sign anywhere. Please answer the related questions on the survey form and take a photo of the return policy as it appears on the sign. Please make sure the photographed policy is readable and scan the photo and email it to the project manager, Laura Ash Smith at: [lsmith@jd20.law.harvard.edu](mailto:lsmith@jd20.law.harvard.edu).

### Summarized Instructions to Testers (Auditors)

*Before attending the training session, testers were provided with the following summarized instructions*:

Thank you for agreeing to participate in the study. This is a study about retail stores’ return practices. You will be asked to go and make non-receipted returns to stores located in Chicago, while following a memorized script, and to record the outcomes of your attempted returns as explained in detail below.

Preparation (before going to stores):

Please memorize the script below and follow it to the letter. We will practice the script during our simulation sessions. You will be provided with a list of stores and their addresses. All stores are located in Chicago. You will also be provided with items that you will need to return to the stores. Each item will be in its original package with tags attached. You will not have the receipt. Please come to the stores dressed in casual clothes: jeans, t-shirt, and minimal make-up.

At the store:

When you enter into the store, please record the following:

The date and time in which you entered the store.

Whether you see a sign with the return policy in the store.

How easy/difficult it was to find the sign in the store.

The font size of the in-store sign.

How crowded the store is

Please approach the first available store clerk at the service/check-out counter. If no clerk is available, please wait in line until a store clerk becomes available. If more than one store clerk is available, please choose the store clerk standing to the right. Please record the perceived race and gender of the store clerk. When it is your turn, and after the store clerk asks you: “what can I do for you?” please say: “I want to return this” and put the item you bought, unopened and in its original packaging, on the counter.

If the store clerk asks you why you want to return the item, please say: “Oh, I just realized I don’t need it.” At this stage, the store clerk will probably try to see if there is a receipt in the bag, and will ask you if you have your receipt. Please answer: “No, I thought I had it with me, but I looked for it, and it seems I lost it.”

If the store clerk asks you for your name, please say your name is Emily Baker / Allison O’Brian [you will receive a list of stores with the name you need to use in each]. If the store clerk asks for your ID, please say “Oh, I don’t have an ID on me at the moment.”

If they ask you when you bought the item, please say that you bought it a few days ago. If they ask you if you remember when exactly, please say: “I don’t remember when exactly.”

If the store clerk gives you a full cash refund, say “thank you” and leave the store. Please record this outcome as “full cash refund” under “outcome 1” in the survey form.

If the store clerk offers anything other than a full cash refund, please say “Are you sure I can’t receive a cash refund? I just bought this [\_\_\_] a few days ago and then realized I don’t need it.” If the store clerk still refuses to give you a full cash refund, please say: “Can you please call the manager? Maybe the manager could help…” If the manager comes, please tell him/her the following: “I bought this [\_\_\_] a few days ago, and I realized I don’t need it. For some reason, I can’t find my receipt. Is it possible to receive a refund?”

If the store clerk tells you “I’m the manager”, please say the following: “Oh…and are you sure there is nothing you can do?” If the answer is still no, or if the manager refuses to provide a refund, say “Ok, thank you, I’ll go back to search for the receipt, then” and leave the store.

If the store clerk refuses to call a manager, saying that he/she already knew what the manager was going to say, please say “Ok, thank you, I’ll go back to search for the receipt, then” and leave the store.

In some stores, the store clerk might tell you that the store does not accept returns (with or without a receipt). If this happens, please write it down in the comments.

### Post-Audit Surveys: Purchasers

Survey Forms for Store Purchasers

**Purchaser’s Name:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Store number** (how many stores have you visited before this one): \_\_\_\_\_

**Please complete this form while at the store or immediately after leaving the store:**

*Background:*

1. Name of store: \_\_\_\_\_\_\_\_\_\_\_\_\_
2. Date of visit: \_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. Time of visit: \_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. Item Purchased (please check one):
   1. Hat \_\_\_\_\_
   2. Gloves \_\_\_\_\_
   3. Scarf \_\_\_\_\_\_
   4. Purse/Bag \_\_\_\_
   5. Shirt \_\_\_\_
   6. Pants \_\_\_\_
   7. Other: \_\_\_\_\_\_\_
5. Item Price: \_\_\_\_\_\_\_
6. If the store doesn’t offer any item within the $20-$30 price range, please check here: \_\_\_\_\_. Please don’t buy any item and leave the store.

*In-Store Sign:*

1. Is there a sign in the store that describes the return policy? Yes \_\_\_\_ No \_\_\_
   1. If there is a sign, how easy was it to find? Please check one:
      1. Very easy: \_\_\_\_
      2. Easy: \_\_\_\_
      3. Neither difficult nor easy: \_\_\_\_
      4. Difficult: \_\_\_\_
      5. Very difficult: \_\_\_\_
   2. If there is a sign, how big is its font (in your opinion)? Please check one:
      1. Very small (very difficult to read): \_\_\_\_
      2. Small (difficult to read): \_\_\_\_
      3. Neither small nor large: \_\_\_\_
      4. Large (easy to read): \_\_\_\_
      5. Very large (very easy to read): \_\_\_\_

*Inside the Store:*

1. How crowded is the store? Please check one:
   1. Not crowded at all: \_\_\_\_
   2. A little crowded: \_\_\_\_
   3. Moderately crowded: \_\_\_\_
   4. Crowded: \_\_\_\_
   5. Very crowded: \_\_\_\_
2. How would you describe the store? Please check one:
   1. Luxury (Upscale): \_\_\_\_\_\_
   2. Mainstream (Casual): \_\_\_\_\_\_
   3. Discount: \_\_\_\_\_\_
   4. Other: \_\_\_\_\_\_
3. Based on looking at items in the store, how would you describe their prices (on average, and based on your overall impression)?
   1. Very expensive: \_\_\_\_
   2. Quite expensive: \_\_\_\_
   3. Moderately priced: \_\_\_\_
   4. Quite cheap: \_\_\_\_
   5. Very cheap: \_\_\_\_
4. Was there a line at the service counter? Yes \_\_\_\_ No \_\_\_\_
   1. How many people were in front of you in the line? \_\_\_\_\_\_\_\_\_\_\_\_
   2. How many people were behind you in the line? \_\_\_\_\_\_\_\_\_\_\_\_\_\_
   3. How much time did it take you to reach the front of the line? \_\_\_\_ minutes.
5. Did the store clerk say anything about the store’s return policy when you made the purchase? If so—Please describe: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
6. Did the store clerk ask you for your name? \_\_\_\_\_

*Other:*

1. Do you have any other comments or issues that arose from your visit? If so, please explain: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### Post-Audit Surveys: Testers

Survey Forms for Store Auditors

**Auditor’s Name:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Store audit number** (how many stores have you audited before this one):\_\_

**Please complete this form while at the store:**

*Background:*

1. Name of store: \_\_\_\_\_\_\_\_\_\_\_\_\_
2. Date of visit: \_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. Time of visit: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Item Type: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Item Price: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

*In-Store Sign:*

1. Is there a sign in the store that describes the return policy? Yes \_\_\_\_ No \_\_\_
   1. If there is a sign, how easy was it to find? Please check one:
      1. Very easy: \_\_\_\_
      2. Easy: \_\_\_\_
      3. Neither difficult nor easy: \_\_\_\_
      4. Difficult: \_\_\_\_
      5. Very difficult: \_\_\_\_
   2. If there is a sign, how big is its font (in your opinion)? Please check one:
      1. Very small (very difficult to read): \_\_\_\_
      2. Small (difficult to read): \_\_\_\_
      3. Neither small nor large: \_\_\_\_
      4. Large (easy to read): \_\_\_\_
      5. Very large (very easy to read): \_\_\_\_

*Inside the Store:*

1. How crowded is the store? Please check one:
   1. Not crowded at all: \_\_\_\_
   2. A little crowded: \_\_\_\_
   3. Moderately crowded: \_\_\_\_
   4. Crowded: \_\_\_\_
   5. Very crowded: \_\_\_\_
2. How would you describe the store? Please check one:
   1. Luxury (Upscale): \_\_\_\_
   2. Mainstream (Casual): \_\_\_\_
   3. Discount: \_\_\_\_
   4. Other: \_\_\_\_\_\_
3. Based on looking at items in the store, how would you describe their prices (on average, and based on your overall impression)? Please check one:
   1. Very expensive: \_\_\_\_
   2. Quite expensive: \_\_\_\_
   3. Moderately priced: \_\_\_\_
   4. Quite cheap: \_\_\_\_
   5. Very cheap: \_\_\_\_
4. **Please complete this form upon leaving the store:**

*Background:*

1. Was there a line at the service counter? Yes \_\_\_\_ No \_\_\_\_
   1. How many people were in front of you in the line? \_\_\_\_\_\_\_\_\_\_\_\_
   2. How many people were behind you in the line? \_\_\_\_\_\_\_\_\_\_\_\_\_\_
   3. How much time did it take you to reach the front of the line? \_\_\_\_ minutes.
2. For the store clerk you spoke with:
   1. What do you think their race was? Please check one:
      1. Black: \_\_\_\_
      2. White: \_\_\_\_
      3. Asian American: \_\_\_\_
      4. Hispanic: \_\_\_\_
      5. Native American: \_\_\_\_
      6. Mixed: \_\_\_\_
      7. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   2. What do you think their gender was? Please check one:
      1. Female: \_\_\_\_
      2. Male: \_\_\_\_
      3. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   3. What do you think their age was? Please check one:
      1. Younger than 18: \_\_\_\_
      2. 18—30 years old: \_\_\_\_
      3. 31—50 years old: \_\_\_\_
      4. 51—70 years old: \_\_\_\_
      5. Older than 70: \_\_\_\_
3. Did the store clerk ask for:
   1. Your name? Yes \_\_\_\_ No \_\_\_\_
   2. Your ID? Yes \_\_\_\_ No \_\_\_\_
   3. Other information?
      1. Yes (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
      2. No \_\_\_\_
4. Did the store clerk ask you any other question?
   1. Yes (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   2. No \_\_\_\_
5. At first, what did the store clerk offer you? Please check one:
   1. Refund \_\_\_\_
   2. Store Credit \_\_\_\_
   3. Exchange \_\_\_\_
   4. Exchange for a different size/color only \_\_\_\_
   5. Nothing offered \_\_\_\_
   6. Other (please explain):\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
6. When you asked if the store clerk was sure that you could not receive a full cash refund, what did they say (please check all that applies)?
   1. That unfortunately that is all they could do without a receipt
   2. That they could ask a manager
   3. They called/asked a manager
   4. They called/asked another clerk at the store
   5. They offered:
      1. Refund \_\_\_\_
      2. Store Credit \_\_\_\_
      3. Exchange \_\_\_\_
      4. Exchange for a different size/color only \_\_\_\_
      5. Nothing \_\_\_\_\_\_
   6. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   7. Any additional comments? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
7. When you asked to see a manager, what did the store clerk say/do?
   1. Said that he/she was the manager
   2. Said that the manager is not available
   3. Refused to call manager, saying he/she already knew what the manager was going to say
   4. Called the manager
   5. Went and spoke to manager
   6. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
8. Did you see a manager? Yes \_\_\_\_ No \_\_\_\_
   1. What do you think their race was? Please check one:
      1. Black: \_\_\_\_
      2. White: \_\_\_\_
      3. Asian American: \_\_\_\_
      4. Hispanic: \_\_\_\_
      5. Native American: \_\_\_\_
      6. Mixed: \_\_\_\_
      7. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   2. What do you think their gender was? Please check one:
      1. Female: \_\_\_\_
      2. Male: \_\_\_\_
      3. Other (please explain): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   3. What do you think their age was? Please check one:
      1. Younger than 18: \_\_\_\_
      2. 18—30 years old: \_\_\_\_
      3. 31—50 years old: \_\_\_\_
      4. 51—70 years old: \_\_\_\_
      5. Older than 70: \_\_\_\_
9. When you talked to the manager, what did they offer you? Please check one:
   1. Refund: \_\_\_\_
   2. Store Credit: \_\_\_\_
   3. Exchange: \_\_\_\_
   4. Exchange for a different size/color only: \_\_\_\_
   5. Nothing offered: \_\_\_\_

*Other:*

1. Do you have any other comments or issues that arose from your visit? If so, please explain: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## B. Appendix B: Supplemental Analyses

### Simple Linear Regressions of Store Characteristics on the Gap

*Table 1. Simple Linear Regression of Policy Type on the Gap*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Moderate Policy | 0.277\*\* | 0.271\*\* | -0.0132 | 0.0777 | 0.246\*\* | 0.194 |
|  | (0.014) | (0.032) | (0.867) | (0.406) | (0.032) | (0.146) |
|  |  |  |  |  |  |  |
| Lenient Policy | 0.0925 | -0.0992 | 0.136\* | 0.162\* |  |  |
|  | (0.389) | (0.412) | (0.076) | (0.076) |  |  |
|  |  |  |  |  |  |  |
| Constant | 0.0870 | 0.304\*\*\* | 0.0435 | 0.0435 | 0.0870 | 0.261\*\* |
|  | (0.308) | (0.002) | (0.472) | (0.545) | (0.316) | (0.012) |
|  |  |  |  |  |  |  |
| Observations | 95 | 95 | 95 | 95 | 56 | 56 |
| R-squared | 0.070 | 0.116 | 0.059 | 0.035 | 0.082 | 0.039 |
|  |  |  |  |  |  |  |

*Notes*. The first column (model 1) is a simple 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. The second column (model 2) is a simple linear regression of pro-consumer gaps at the final stage on store policy type. The third column (model 3) is a simple linear regression of refund outcomes at the initial stage on store policy type. The fourth column (model 4) is a simple linear regression of non-receipted returns at the initial stage on store policy type (out of the stores that require a receipt for any return or exchange), and the fifth column (model 5) is a simple linear regression of non-receipted returns at the final stage on policy type (out of the stores that require a receipt for any return or exchange). The reference category for policy type is “harsh.” P-values are reported in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

The results show that the store’s policy type significantly affected return outcomes. Moderate policy stores were significantly more likely than harsh policy stores to apply a pro-consumer gap, both at the initial and final stages. Lenient policy stores were marginally significantly more likely than harsh policy stores to provide non-receipted refunds, both at the initial and final stages. Finally, moderate policy stores were significantly more likely than harsh policy stores to offer non-receipted exchanges or store credits.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| Mainstream | 0.0278 | 0.0167 | 0.0269 | 0.0915 | 0.262 | 0.292 |
|  | (0.691) | (0.595) | (0.778) | (0.401) | (0.215) | (0.177) |
|  |  |  |  |  |  |  |
| Luxury | 0.183\*\* | 0.317\*\* | 0.191\* | 0.319\*\* | 0.200 | 0.200 |
|  | (0.039) | (0.040) | (0.092) | (0.015) | (0.370) | (0.380) |
|  |  |  |  |  |  |  |
| Constant | 0.15 | 0.25\*\* | 0.0588 | 0.0556 | 0.200 | 0.400\*\* |
|  | (0.104) | (0.017) | (0.452) | (0.528) | (0.208) | (0.017) |
|  |  |  |  |  |  |  |
| Observations | 95 | 95 | 95 | 95 | 56 | 56 |
| R-squared | 0.077 | 0.071 | 0.053 | 0.093 | 0.053 | 0.061 |

*Table 2. Simple Linear Regression of Store Prestige on the Gap*

*Notes*. The first column (model 1) is a simple 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 prestige. The second column (model 2) is a simple linear regression of pro-consumer gaps at the final stage on store prestige. The third column (model 3) is a simple linear regression of refund outcomes at the initial stage on store prestige. The fourth column (model 4) is a simple linear regression of non-receipted returns at the initial stage on store prestige (out of the stores that require a receipt for any return or exchange), and the fifth column (model 5) is a simple linear regression of non-receipted returns at the final stage on store prestige (out of the stores that require a receipt for any return or exchange). The reference category for store prestige is “discount.” P-values are reported in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

The results show that the store’s prestige significantly affected return outcomes. Luxury stores were significantly more likely than discount stores to apply a pro-consumer gap, both at the initial and final stages. Luxury stores were also significantly more likely than discount stores to provide non-receipted refunds.

*Table 3. Simple Linear Regression of Store Experience on the Gap*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| Experience | 0.0863 | 0.0901 | 0.123\*\* | 0.0888 | 0.0853 | 0.218\* |
|  | (0.233) | (0.264) | (0.021) | (0.163) | (0.449) | (0.069) |
|  |  |  |  |  |  |  |
| Constant | -0.0795 | 0.0235 | -0.351\* | -0.162 | 0.000533 | -0.257 |
|  | (0.777) | (0.940) | (0.086) | (0.509) | (0.999) | (0.544) |
|  |  |  |  |  |  |  |
| Observations | 95 | 95 | 95 | 59 | 56 | 56 |
| R-squared | 0.021 | 0.019 | 0.079 | 0.029 | 0.019 | 0.103 |

*Notes*. The first column (model 1) is a simple 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 experience (measured by years since establishment). The second column (model 2) is a simple linear regression of pro-consumer gaps at the final stage on store experience. The third column (model 3) is a simple linear regression of refund outcomes at the initial stage on store experience. The fourth column (model 4) is a simple linear regression of non-receipted returns at the initial stage on store experience (out of the stores that require a receipt for any return or exchange), and the fifth column (model 5) is a simple linear regression of non-receipted returns at the final stage on store experience (out of the stores that require a receipt for any return or exchange). P-values are reported in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

*Table 4. Simple Linear Regression of Store Size on the Gap*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| Revenues (Log) | 0.00652 | 0.0122 | -0.00827 | -0.00659 | 0.0267 | 0.0398\* |
|  | (0.705) | (0.513) | (0.518) | (0.627) | (0.221) | (0.084) |
|  |  |  |  |  |  |  |
| Constant | 0.255\*\*\* | 0.339\*\*\* | 0.125\*\*\* | 0.141\*\*\* | 0.322\*\*\* | 0.500\*\*\* |
|  | (0.000) | (0.000) | (0.006) | (0.004) | (0.001) | (0.000) |
|  |  |  |  |  |  |  |
| Observations | 95 | 95 | 95 | 95 | 56 | 56 |
| R-squared | 0.003 | 0.008 | 0.008 | 0.004 | 0.057 | 0.111 |

*Notes*. The first column (model 1) is a simple 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 size (measured logarithm of annual revenues from the year of 2018). The second column (model 2) is a simple linear regression of pro-consumer gaps at the final stage on store experience. The third column (model 3) is a simple linear regression of refund outcomes at the initial stage on store experience. The fourth column (model 4) is a simple linear regression of non-receipted returns at the initial stage on store experience (out of the stores that require a receipt for any return or exchange), and the fifth column (model 5) is a simple linear regression of non-receipted returns at the final stage on store experience (out of the stores that require a receipt for any return or exchange). P-values are reported in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

*Table 5. Simple Linear Regression of Store Type (Chain or Local) on the Gap*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Chain | 0.212\*\* | 0.153 | 0.0837 | 0.0928 | 0.340\*\*\* | 0.396\*\*\* |
|  | (0.023) | (0.159) | (0.208) | (0.235) | (0.003) | (0.002) |
|  |  |  |  |  |  |  |
| Constant | 0.0714 | 0.250\*\*\* | 0.0357 | 0.0714 | 0.0741 | 0.259\*\*\* |
|  | (0.356) | (0.007) | (0.521) | (0.276) | (0.347) | (0.006) |
|  |  |  |  |  |  |  |
| Observations | 95 | 95 | 95 | 95 | 56 | 56 |
| R-squared | 0.054 | 0.021 | 0.017 | 0.015 | 0.154 | 0.157 |

*Notes*. The first column (model 1) is a simple 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 type—chain or local (defined as a store that has no more than two locations, both in Illinois). The second column (model 2) is a simple linear regression of pro-consumer gaps at the final stage on store type. The third column (model 3) is a simple linear regression of refund outcomes at the initial stage on store type. The fourth column (model 4) is a simple linear regression of non-receipted returns at the initial stage on store type (out of the stores that require a receipt for any return or exchange), and the fifth column (model 5) is a simple linear regression of non-receipted returns at the final stage on store type (out of the stores that require a receipt for any return or exchange). P-values are reported in parentheses (*\* p* < 0.1, \*\* *p* < 0.05, \*\*\* *p* < 0.01).

### Tables of Return Outcomes across Stores: Initial and Final Stages

Return Outcomes at the Initial Stage:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Store Type | Return Denied[[119]](#footnote-119) | 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 \*.

Return Outcomes at the Initial and Final Stages:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Store Type | Stage | Return Denied[[120]](#footnote-120) | 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% |

*Note:* When the differences in return outcomes between the initial and final stages (within policy category) 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 \*.

1. \* Olin Law & Economics Teaching Fellow and Lecturer in Law, University of Chicago Law School. E-mail: mfurth@uchicago.edu. This research was generously supported by the Coase-Sandor Institute at the University of Chicago and the Institute for Quantitative Social Science at Harvard University. [Acknowledgements to be added.] [↑](#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”); Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 Harv. L. Rev. 1173, 1221 (1983).  [↑](#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 are often overly harsh); 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) (noting that adhering parties tend to neither read nor negotiate the standard-form contracts they sign); 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.”); 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); 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 rampant use of clear and unconditional boilerplate 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); 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.*, Rakoff, *supra* note 1, at 1221 (noting that standardized agreements might omit “obligations that the firm recognizes in its actual practice” so as to leave “room to maneuver”); 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, 436 (2002) (suggesting that “the business seeks to establish and maintain a good reputation with the purchasing public and generally will stand behind its product”); 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 underenforce 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, 2080 (2014) (noting “the gap between contract law on the books and in practice”); 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) (explaining that “the terms of the paper deal” often differ from “the terms of the real deal—that is, the way sellers actually behave in the shadow of both written contracts and the wide variety of other forces that may constrain or influence their behavior”). [↑](#footnote-ref-4)
5. *See, e.g.*,Gillette 2005, *supra* note 3, at 977; Bebchuk & Posner, *supra* note 3, at 828. [↑](#footnote-ref-5)
6. I borrow the terms “paper deal” and “real deal” 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, almost no empirical attention has been given to the “paper deal—real deal” divergences in consumer markets. On the need to shift scholarly attention from the “paper deal” to the “real deal” in consumer settings, *see, e.g.*, Bernstein and Volvovsky, *supra* note 3, at 129. [↑](#footnote-ref-6)
7. *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”); Rory Van Loo, The Corporation as Courthouse, 33 Yale J. on Reg. (2016) (noting the challenges with reputation-based sanctions); 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 their contract 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-7)
8. *See, e.g.*, Gillette 2004, *supra* note 3, at 712 (“we might be wary of systematically invalidating clauses in SFCs [standard form contracts] that permit sellers latitude in enforcement on the assumption that sellers use such clauses selfishly.”); Bebchuk & Posner, *supra* note 3, at 834 (suggesting 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). [↑](#footnote-ref-8)
9. *See, e.g.*, Arbel, *supra* note6; Becher & Zarsky, *supra* note 6, at 14; *see also* Eyal Zamir & Doron Teichman, Behavioral Law & Economics 304 (2018) (explaining that legal intervention may be necessary to protect consumers in cases of behavioral market failure); Van Loo, supra note 7 at 595-97 (arguing for regulatory supervision of corporations as a policy proposal for addressing “the pitfalls of the private consumer legal system,” id. at 595); 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”); Yonathan A. Arbel & Roy Shapira, Theory of the Nudnik: The Future of Consumer Activism and What We Can Do to Stop it, Vanderbilt Law Rev. 40-41 (forthcoming 2020) (arguing legal intervention is necessary for reputational constraints to even function properly). [↑](#footnote-ref-9)
10. Several scholars have acknowledged this deficiency, calling for future empirical work on these issues. *See, e.g.*,Becher & Zarsky, *supra* note 6, 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-10)
11. Importantly, I use the term “sticky” here to describe sellers’ adherence to the formal agreement. This is distinct from the use of the term “stickiness” to describe drafting parties’ reluctance to change the formal terms of their agreements over time (see, e.g., Anna Gelpern, Mitu Gulati & Jeromin Zettelmeyer, *If Boilerplate Could Talk* (Duke Law School Public Law & Legal Theories Series 2017-45 (2018), <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2984293>). [↑](#footnote-ref-11)
12. *See, e.g.*, Gillette 2004, *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 6, 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-12)
13. 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-13)
14. Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-14)
15. For a discussion of the shortcomings of reputational considerations as an enforcement mechanism, see Van Loo, supra note 7 at 584, 592-93. [↑](#footnote-ref-15)
16. By providing support for these predictions, this Article contributes to a long line of research on the interactions between sellers and consumes 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); Christian Homburg & Andreas Furst, *See No Evil, Hear No Evil, Speak No Evil: A Study of Defensive Organizational Behavior Towards Customer Complaints*, 35 J. Academy Marketing Sci. 523 (2007) (investigating why firms engage in defensive organizational behavior towards customer complaints); Laura Nader, *Disputing without the Force of Law*, 88 Yale L.J. 998, 1012-14 (1979) (describing the consumer complaint process as manipulating consumers but generally giving persistent consumers what they request); Amy J. Schmitz, *Access to Consumer Remedies in the Squeaky Wheel System,* 39 Pepp. L. Rev. 279 (2012) (concluding that corporations who violate the law use their customer service departments to appease consumers who complain for the strategic purpose of being able to continue violating other consumers’ rights). [↑](#footnote-ref-16)
17. 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-17)
18. *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 feel generally 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, *On the Unexpected Use of Unenforceable Contract Terms: Evidence from the Residential Rental Market*, 9 J. Legal Anal. 1 (2017) (finding, based on a survey of residential tenants, that tenants often rely on the written lease agreements when disputes with the landlord arise); Meirav Furth-Matzkin, *The Harmful Effects of Unenforceable Contract Terms: Experimental Evidence*, 4 Ala. L. Rev. 1032 (2019) (providing experimental evidence 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) (providing experimental evidence that consumers are demoralized by fraudulent fine print even when it contradicts what they were promised at the pre-contractual stage). [↑](#footnote-ref-18)
19. For an exploration of why many consumers are passive and what may lead a minority of consumers to complain, see Yonathan A. Arbel & Roy Shapira, Theory of the Nudnik: The Future of Consumer Activism and What We Can Do to Stop it, Vanderbilt Law Rev. 7-15 (forthcoming 2020). [↑](#footnote-ref-19)
20. See Van Loo, supra note 7, at 579-80 (noting that “[s]tudies have for decades found that wealthier and better-educated consumers are more likely to complain to corporations and are more successful when they do than are low-income consumers,” id. at 579). [↑](#footnote-ref-20)
21. *See infra* notes 98–101. [↑](#footnote-ref-21)
22. Cf. Arbel & Shapira, supra note 9, at 32-34. [↑](#footnote-ref-22)
23. *See* sources cited *supra* note 3. [↑](#footnote-ref-23)
24. *See, e.g.*, Gillette 2004, *supra* note 3, at 704–07 (explaining that reputational constraints impact sellers but not consumer buyers); Bebchuk & Posner, *supra* note 3, at 827 (similarly discussing this asymmetry). [↑](#footnote-ref-24)
25. *See, e.g.*, 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-25)
26. *See, e.g.*,Gillette 2004, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 831 (noting that such party-specific information may be “available to the parties but not easily and accurately observable by [a] court”). [↑](#footnote-ref-26)
27. *See, e.g.*,Gillette 2004, *supra* note 3, at 704; Bebchuk & Posner, *supra* note 3, at 834; Johnston, *supra* note 3, at 882. [↑](#footnote-ref-27)
28. 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 consumer.”). [↑](#footnote-ref-28)
29. Zamir & Teichman, *supra* note 8, at 311 (arguing that “the flow of information in consumer markets, even in our information era, is far from perfect, and so, too, is the use of information by customers” who may, for instance, improperly assess the reputational information they receive); Becher & Zarsky, *supra* note 6, 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 6, at 5 (suggesting that we have reached a point of “reputational failure”). [↑](#footnote-ref-29)
30. Johnston, *supra* note 3, at 887. [↑](#footnote-ref-30)
31. Gillette 2004, *supra* note 3, at 712. [↑](#footnote-ref-31)
32. Bebchuk & Posner, *supra* note 3, at 834. [↑](#footnote-ref-32)
33. Zamir & Ayres, *supra* note 8, at 3. *See also* Arbel, *supra* note 6, at 1 (arguing that the current state of “reputational failure” highlights “the centrality of the law to the future of the marketplace”). [↑](#footnote-ref-33)
34. *See, e.g.*,Gillette 2004, *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”); Bebchuk & Posner, *supra* note 3 (using the case of return policies in their model); Johnston, *supra* note 3, at 873 (providing the example of product returns); Becher & Zarsky, *supra* note 6, 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”). So far, these commentators have only relied 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-34)
35. 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-35)
36. *See, e.g.*,Zamir & Teichman, *supra* note 8, at 290–91; Shmuel Becher & Tal Zarsky, *Open Doors, Trap Doors and the Law*, 74 L. & Contemp. Probs. 63, 72–73 (2011) (discussing how sellers often use “open door policies” to allow consumers to return purchases). [↑](#footnote-ref-36)
37. 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-37)
38. According a recent consumer poll, 91% of consumers consider return policies as very important to 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* note 8, at 290–91 (arguing for the importance of the right to withdraw given that consumers may be subject to sellers’ deceptive “low-ball[ing]” and “bait-and-switch” tactics). [↑](#footnote-ref-38)
39. There are few notable exceptions to this general rule. Federal law provides for a cooling-off period in a handful of door-to-door transactions. *See* 16 C.F.R. § 429 (stating 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-39)
40. At the state level, some jurisdictions provide little additional protections beyond those laid out federally, and some (e.g., Illinois, Connecticut, Washington, Pennsylvania, and Michigan) have even expanded the federal three-day right to cancel transactions to extend to certain purchases not covered under federal law (e.g., gym memberships). 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-40)
41. *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 8, 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-41)
42. *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-42)
43. Johnston, *supra* note 3, at 873–74 (discussing the commonality of liberal return policies as standard practice in, e.g., consumer electronics and other retail areas). Others have relied on the observed commercial norms to justify more comprehensive regulation of consumers’ withdrawal rights. See, e.g., Teichman & Zamir, *supra* note 8, at 291 (advocating the importance of lenient return policies with cooling-off periods to protect consumers); Ben-Shahar & Posner, *supra* note 34, at 120–21 (noting that “nearly all retail stores in the United States permit customers to return merchandise for a refund,” and calling for the adoption of a default right to withdraw). [↑](#footnote-ref-43)
44. *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-44)
45. *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-45)
46. *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-46)
47. 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 meta-analytic reviews of this literature, see generally 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); Scott Davis, Michael Hagerty, & Eitan Gerstner, *Return Policies and the Optimal Level of Hassle*, 50 J. Econ. & Bus. 445 (1998). [↑](#footnote-ref-47)
48. Field experiments are increasingly used in legal scholarship, and are considered as “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-done studies); Michael Abramowicz, Ian Ayres, & Yair Listokin, *Randomizing Law*, 159 U. Pa. L. Rev. 929 (2011); D. James Greiner & Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 Annual Rev. L. & Soc. Sci. 295, 305–08 (2016) (describing, and explaining, the general paucity of randomized controlled experiments in legal scholarship). [↑](#footnote-ref-48)
49. 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 Car Negotiations*, 104 Harv. L. Rev. 817, 822–27 (1991) (employing audit methodology to study discrimination in car dealerships); Ian Ayres and Peter Siegelman, *Race and Gender Discrimination in Bargaining for a New Car*, 85 Am. Econ. Rev. 304, 305–06 (1995). These 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 41; Abramowicz et al., *supra* note 43. 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-49)
50. In a larger sample of 192 retail stores (established 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 42. [↑](#footnote-ref-50)
51. 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 limited light on the return practices of the most luxurious designer stores that do not offer items for $30 or less. I leave this issue for future research. [↑](#footnote-ref-51)
52. 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-52)
53. Age refers to the number of years of operation since incorporation. Data on firms’ revenues and age were obtained primarily from Bloomberg Law and Hoover’s Company Directories’ databases. [↑](#footnote-ref-53)
54. This generally reflects the market share division between local and chain stores in Chicago. *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-54)
55. For consistency, testers were instructed to audit the stores on weekdays in the afternoons. [↑](#footnote-ref-55)
56. 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 concern 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-56)
57. 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-57)
58. 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 exchange or store credit. Unlike returning a gift or exchanging an item for a different size, explicitly saying that 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-58)
59. The composition of pairs varied from audit to audit: 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-59)
60. Testers obtained inconsistent return outcomes in 15% of the stores sampled (n = 132). [↑](#footnote-ref-60)
61. In order to keep testers’ fake 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, 911 (2004); Benjamin Edelman, Michael Luca, & Dan Svirsky, *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, 9 Am. Econ. J. 1, 7 (2017) (experimentally manipulating “names that signaled African-American males, African-American females, White males, and White females” in a field experiment). [↑](#footnote-ref-61)
62. In 17% of the stores, store clerks asked testers to show an ID to process the return. In these cases, a second pair of testers came with IDs and their return outcomes were used for the analysis. [↑](#footnote-ref-62)
63. Testers were accompanied by project coordinators to the stores. The coordinators ensured that testers were following the script and accurately reporting the results. [↑](#footnote-ref-63)
64. 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-64)
65. 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. [↑](#footnote-ref-65)
66. Importantly, the sample did not include any stores that did not allow for refunds while at the same time allowing for non-receipted 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-66)
67. Under a simple regression of initial return outcomes (defined as a categorical variable taking 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-67)
68. 60 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. [↑](#footnote-ref-68)
69. 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, compared to 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-69)
70. In the script, consumers were instructed to complain and ask to speak to the manager if denied the return, but they were instructed not to mention the contractual language. In future research, it is desirable to explore how sellers react when testers/consumers point them to the language that entitles them to the denied benefit. [↑](#footnote-ref-70)
71. This prediction is consistent with previous writings 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). [↑](#footnote-ref-71)
72. *b* = 0.35, SE = 0.09, *p* < 0.001. [↑](#footnote-ref-72)
73. *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-73)
74. 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-74)
75. 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-75)
76. 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-76)
77. *b =* 0.02, SE = 0.09, *p* = 0.8. [↑](#footnote-ref-77)
78. Before embarking on the field experiment, I had 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 UChicago Marketplace. The interviews were semi-structured, and mainly focused 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 1. [↑](#footnote-ref-78)
79. Interview #12 with Abercombie store clerk (interview script on file with the Author). [↑](#footnote-ref-79)
80. Interview #7 with Saks Fifth Avenue store clerk (recorded interview on file with the Author). [↑](#footnote-ref-80)
81. Interview #4 with Bally store clerk (recorded interview on file with the Author). [↑](#footnote-ref-81)
82. 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 dismiss their claims. [↑](#footnote-ref-82)
83. 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-83)
84. This effect is driven by older stores’ greater likelihood to provide refunds notwithstanding testers’ failure to show a receipt. Older stores were not significantly more likely to allow for non-receipted returns more generally, as Models Five and Six show. [↑](#footnote-ref-84)
85. In an ongoing work, I study the variation in return policy leniency across stores. For this purpose, I established a database of return policies, as they appeared on the stores’ websites, in-store signs, and receipts. The policies belong to 192 randomly selected clothing retail stores operating in Chicago. The return policies were then analyzed according to a policy leniency index I constructed. The results show that luxury stores offer significantly better terms than other, more casual stores. In the particular context of return policy length, older and more luxurious stores offered significantly lengthier return periods than younger, more casual stores. [↑](#footnote-ref-85)
86. Note, however, that sellers may authorize their representatives to operate a different set of rules, rather than using standards. For example, some airlines have an informal “flat tire” rule, stating that customers who miss their flight due to unforeseen circumstances beyond their control—like a flat tire—are able to get on the next flight to their destination at no charge, as long as they arrive within two hours of the missed flight. *See, e.g.,* Claire Nowak, *This Little-Known Airplane Rule Can Help When You Miss Your Flight*, Reader’s Digest, available at https://www.rd.com/advice/travel/missed-flight-flat-tire-rule/; Amanda Harding, *This Surprising Airline Rule Can Help If You Miss Your Flight*, Shobiz Cheatsheet (May 5, 2018), available at https://www.cheatsheet.com/culture/this-surprising-airline-rule-can-help-if-you-miss-your-flight.html/. [↑](#footnote-ref-86)
87. Yet, note that biased demand generated by imperfectly rational consumers might lead sellers to adopt terms that are both inefficient and harmful to consumers. If consumers, for example, underestimate the benefits from a more lenient return policy (or the likelihood that they will need to return the good), sellers might inefficiently use more rigid contractual provisions, while lowering the price of the product to entice consumers. On the interaction between consumers’ bounded rationality and market forces, see, e.g., Bar-Gill, *supra* note 2 (suggesting that competitive forces drive sellers to respond to consumer misperceptions, thereby generating a “behavioral market failure”); Korobkin, *supra* note 2; Tom Baker & Peter Siegelman, *Behavioral Economics and Insurance Law: The Importance of Equilibrium Analysis, in* The Oxford Handbook of Behavioral Economics and the Law (2014); David Gilo & Ariel Porat, *The Hidden Roles of Boilerplate and Standard-Form Contracts: Strategic Imposition of Transaction Costs, Segmentation of* *Consumers, and Anticompetitive Effects*, 104 Mich. L. Rev. 983, 985 (2006) (observing that “there is a risk that the supplier will extract payment from the consumer without the latter being aware of the fact that the payment does not reflect the reduction of value due to the harsh clause”). In the employment context, seeRachel Arnow-Richman, *Cubewrap Contracts and Worker Mobility: The Dilution of Employee Bargaining Power via Standard Form Noncompetes*, 2006 Mich. St. L. Rev. 963, 981 (2006) (noting that scholars have increasingly “questioned the extent to which certain contracting parties, in particular consumers, may be ‘boundedly rational’ in assessing contract terms,” and that “workers assess[ing] standardized terms like noncompete agreements” are likely susceptible to such “cognitive failures”). [↑](#footnote-ref-87)
88. 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 of their lower ability to resell used items or return them to the supplier. [↑](#footnote-ref-88)
89. Indeed, the question of whether a certain contractual arrangement is desirable or welfare-enhancing is undoubtedly a difficult one, and there may be reasons to refrain from intervening in the contents of standardized agreements. In particular, in competitive markets *without informational asymmetries*, firms—rather than policymakers or courts—may be better equipped to determine whether specific terms are socially optimal, because they are better able to estimate both the benefits of these terms to consumers and the costs of offering them. Yet in reality, in most consumer transactions, substantial informational asymmetries often persist, casting doubt on the ability of market forces of reputation and competition to adequately incentivize sellers to offer favorable terms. [↑](#footnote-ref-89)
90. And in fact, firms that try to sparingly depart from their terms may be incentivized to do so quietly. Cf. Arbel & Shapira, supra note 9, at 32-34. [↑](#footnote-ref-90)
91. *See* George A. Akerlof, *The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*, 84 Q. J. Econ. 488 (1970). [↑](#footnote-ref-91)
92. Cf. Arbel & Shapira, supra note 19, at 32-34. [↑](#footnote-ref-92)
93. *See, e.g.*, Gillette 2004, *supra* note 3, at 706 (“Sellers may use contract terms in an *in terrorem* effort to deter requests for redress, or as an initial response to buyer complaints”). [↑](#footnote-ref-93)
94. *See, e.g.*, 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); Furth-Matzkin 2019, *supra* note 16 (finding that most tenants in an experimental survey acquiesced to whatever their lease terms said, even when they contained unenforceable liability disclaimers); Furth-Matzkin & Sommers, s*upra* note 16 (finding that consumers are unlikely to take action against deceptive sellers after reading contractual terms that conflict with the sellers’ prior representations). [↑](#footnote-ref-94)
95. *See supra* note 15. [↑](#footnote-ref-95)
96. *See, e.g.*, Furth-Matzkin 2017, *supra* note 16 (finding, based on a survey of residential tenants, that tenants often rely on the written lease agreements when disputes with the landlord arise); Furth-Matzkin 2019, *supra* note 16 (providing experimental evidence that tenants perceive their lease terms as enforceable and binding, even when they contain unenforceable contract terms); Furth-Matzkin & Sommers, *supra* note 16 (finding that consumers are demoralized by fraudulent fine print even when it contradicts what they were promised at the pre-contractual stage). [↑](#footnote-ref-96)
97. *See generally, e.g.*, Wilkinson-Ryan 2017, *supra* note 16; Furth-Matzkin 2019, *supra* note 16; Furth-Matzkin & Sommers, *supra* note 16. [↑](#footnote-ref-97)
98. Furth-Matzkin 2019, *supra* note 16. 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-98)
99. Furth-Matzkin & Sommers, *supra* note 16. [↑](#footnote-ref-99)
100. 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 8 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) (arguing in favor of the efficacy of legal entitlements, in the context of redistributive legal rules). [↑](#footnote-ref-100)
101. See Arbel & Shapira, supra note 9, at 19 (arguing that a small subset of consumers do air their grievances with companies online and that consumers “often do read and engage with consumers’ reviews.”). [↑](#footnote-ref-101)
102. See id. at 19-20. [↑](#footnote-ref-102)
103. *See, e.g.*, Furth-Matzkin 2019, *supra* note 16 (finding that consumers are discouraged by unenforceable fine print from searching for online information about their rights and often rely on the contractual language instead). [↑](#footnote-ref-103)
104. Johnston, *supra* note 3, at 858. [↑](#footnote-ref-104)
105. For a similar assertion, *see* Gillette 2004*, supra* note 3, at 706 (suggesting that sellers might choose an inefficient risk allocation if the “cost of absorbing the costs of all defects” exceeds the costs of “accepting the risk for insistent buyers and any reputational loss from alienating disappointed but noninsistent buyers”). The analysis here adds two other components to sellers’ cost-benefit calculation: the costs of scaring away consumers *ex ante* and the costs of accommodating opportunistic consumers. [↑](#footnote-ref-105)
106. Regressive distributional concerns might yield different policy prescriptions than an analysis that only addresses overall efficiency (or welfare). *See, e.g.*, Levinson-Zamir, *supra* note 90 at 396–97 (explaining the benefits of redistributive private law rules as a matter of public policy); Lee Fennell & Richard McAdams, *The Distributive Deficit in Law and Economics*, 100 Minn. L. Rev. 1051, 1053 (2016) (explaining that “both efficiency and distribution matter to welfare,” with different associated policy prescriptions). [↑](#footnote-ref-106)
107. *See generally, 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, 92 (2001) (demonstrating that women exhibit “depressed entitlement by paying themselves less than men pay themselves for the same work,” due to variations in their respective *perceptions* of self-worth and performance); Paul K. Piff, *Wealth and the Inflated Self: Class, Entitlement and Narcissism*, 40 Personality & Soc. Psych. Bulletin 34 (2014) (demonstrating that higher socioeconomic class status is associated with higher levels of entitlement); 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) (arguing that objectively disadvantaged groups experience lesser senses of personal entitlement); 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) (demonstrating that a lower position within the social hierarchy is associated with a lesser inclination to believe that one is being treated unfairly); 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 Pol. Psych.: Soc. & Psych. Bases Ideology & System Justification(2009) (explaining that one’s sense of personal entitlement is impacted by both system justification and social comparison processes, by which low-status groups ultimately experience lesser senses of entitlement); 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 to get what they want later in life, while childrearing strategies among the lower classes 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)
108. *See generally, 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) (finding lower levels of perceived pay entitlement among women in comparison to men); Patricia J. Williams, *Alchemical Notes: Reconstructing Ideals from Deconstructed Rights*, 22 Harv. C.R.-C.L. L. Rev. 401 (1987) (arguing for racial differences in perceptions of rights-entitlements). Black and lower-income people were also found to be less likely to seek legal help when encountering civil legal problems (see, e.g., Sara Sternberg Greene, *Race, Class, and Access to Civil Justice*, 101 Iowa L. Rev. 1263 (2016)). [↑](#footnote-ref-108)
109. *See, e.g.*, Furth-Matzkin & Sommers, *supra* note 16 (finding that nonwhite participants were inclined to see the consumer as more bound by the fine print than 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) (arguing that social class is an important determinant of one’s sense of entitlement, and that people with a higher sense of entitlement are more likely to bargain and to require more from their contracting parties during contract negotiations). [↑](#footnote-ref-109)
110. Cf. Van Loon, supra note 7, at 579-80. [↑](#footnote-ref-110)
111. Interview #13 with a local rug store clerk (recorded interview on file with the Author). [↑](#footnote-ref-111)
112. See Van Loon, supra note 7, at 579-80. [↑](#footnote-ref-112)
113. See id. [↑](#footnote-ref-113)
114. 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 white customers, and the difference was larger after asking to speak to the store’s manager. [↑](#footnote-ref-114)
115. Cf. Van Loon, supra note 7 at 579-80. [↑](#footnote-ref-115)
116. For a similar proposal, see Arbel & Shapira, supra note 9, at 44-46. [↑](#footnote-ref-116)
117. See Van Loon, supra note 7, at 579-80. [↑](#footnote-ref-117)
118. The information revolution may also be used to combat market discrimination. For example, unlawful discrimination could be reduced if sellers have more information on consumers (such as their purchasing history and past return behavior). *See, e.g.*, Lior Jacob Strahilevitz, *Reputation Nation: Law in an Era of Ubiquitous Personal Information*, Nw. U. L. Rev. 102 (2008); Lior Jacob Strahilevitz, *Less Regulation, More Reputation, in* The Reputation Society 71 (2012) (suggesting that “an important potential upside of new regulation tracking technologies is their potential to displace statistical discrimination on the basis of race, gender, age, appearance, and other easily observable characteristics”). For a similar proposal, see Van Loo, supra note 7 at 595-97. [↑](#footnote-ref-118)
119. Outcomes were coded as “return denied” if the store clerk either refused to offer any concession when testers came without receipt or only allowed testers to exchange the item for different size or color. [↑](#footnote-ref-119)
120. These figures include cases where store clerks only allowed testers to exchange the item for different size or color. [↑](#footnote-ref-120)