Copyright Enforcement: Evidence from Two Field Experiments*

Hong Luo† Julie Holland Mortimer‡

February 19, 2016

Abstract

Effective dispute resolution is important for reducing private and social costs. We study how resolution responds to changes in price and communication using a new, extensive dataset of copyright infringement incidences by firms. The data cover two field experiments run by a large stock-photography agency. We find that substantially reducing the requested amount generates a small increase in the settlement rate. However, for the same reduced request, a message informing infringers of the price reduction and acknowledging possible unintentionality generates a large increase in settlement; including a deadline further increases the response. The small price effect, compared to the large message effect, can be explained by two countervailing effects of a lower price: an inducement to settle early, but a lower threat of escalation. Furthermore, acknowledging possible unintentionality may encourage settlement due to the typically inadvertent nature of these incidences. The resulting higher settlement rate prevents additional legal action and reduces social costs.

*The authors thank James Dana, Alberto Galasso, Shane Greenstein, Michael Grubb, William Hickman, Gastón Llanes, Mike Luca, Cynthia Montgomery, Felix Oberholzer-Gee, Jason Snyder, Scott Stern, Ricardo Perez-Truglia, Dennis Yao and seminar participants at the Canadian Conference on the Economics of Innovation and Entrepreneurship, Duke Strategy Conference, Harvard Business School, MIT Sloan, Microsoft Research, NBER Economics of Digitization Conference, NBER Productivity Lunch, Stern School of Business, the Strategic Research Forum, and Washington University in St. Louis for helpful comments. We thank the Agency for providing the proprietary data and for their collaboration on the experiments. We also thank Danielle Wedde and Esther Yan for their excellent research assistance. All errors and omissions are ours.

†Harvard Business School; email: hluo@hbs.edu.

‡Boston College and NBER; julie.mortimer.2@bc.edu.

1 Introduction

Disputes involving economic damages, such as breach of contract, overdue payment, and violation of property rights, are commonplace in the economy. Effectively resolving these disputes can save time and stress, and reduce legal expenses. The threat of legal sanction may be credible only in some circumstances (e.g., it is impractical when monetary claims are small), and may not always be desirable even if effective, because a litigious reputation may drive away customers. Encouraging collaborative outcomes in these situations may require alternative, creative methods.

Copyright enforcement is an interesting example of this problem, and is important in the management of intellectual property (IP). Monetary amounts are generally small, and many content producers want to avoid a litigious reputation because it may hurt their primary source of revenue from the legitimate market. Legal actions against file-sharing networks and infringing individuals, notably by trade associations in the music and movie industries, have been controversial and have met with limited success. While digital technologies have facilitated instances of copyright infringement in recent years, they have also substantially lowered the costs of detecting infringement, making enforcement an increasingly relevant option.[1] In this paper, we study the effectiveness of new approaches to enforcement on copyright infringement settlement outcomes. In particular, we focus on two aspects of enforcement methods: (1) the initial amount requested from an infringer, and (2) communication.

Our empirical application comes from the stock-photography industry.[2] Stock photos are pre-shot images that are readily available for licensing. On behalf of photographers, agencies such as Corbis Images and Getty Images manage, market, and license images to business customers. The extent of piracy is enormous in this market, as digital images are easily accessible online through search engines and social media sites, and users often have limited knowledge about the legal obligation of using an image. Furthermore, the frequent use of third-party advertisers and web designers results in many indirect instances of infringement, for which end-users are ultimately liable. We use data provided by one of the leading agencies in the industry (hereafter, the Agency), which monitors the online use of a small set of its most expensive images and pursues settlement for instances of infringement carried out for commercial purposes. The dataset contains detailed information on all infringement instances for which the Agency pursued settlement in the U.S. and Canada between October 2013 and March 2014, including the requested amount, the characteristics of

[1] Some recent examples of anti-piracy services that have been developed by private companies include Digimarc in the publishing industry and MarkMonitor (purchased by Thomson Reuters in 2012) in industries including movies, music, software, and publishing.

[2] Glückler and Panitz (2013) estimates that the global revenue of the stock-photography industry was $2.88 billion in 2011. In comparison, the revenues of ASCAP and BMI, the largest two performance rights organizations that collect royalties on behalf of copyright holders in the music industry for public use of their works, were $1B each in 2012.

infringing firms, and settlement outcomes.

Our data are novel for three reasons. First, previous empirical studies on settlement have depended primarily on court filings, which capture only a small percentage of disputes.[3] Our sample is 'unselected' in that it includes all instances of infringement for which the Agency pursues settlement, and the data capture the earliest stage of the enforcement process. This is important because the set of infringers (and their settlement behavior) at this stage are likely to differ from those in the cases that reach litigation. Second, it is rare for field data to contain exogenous variation in enforcement methods, and laboratory experiments may be hard to generalize to real-world settings. Our data are generated by two field experiments, and include exogenous variation in both pricing and communication. Third, infringing firms in our sample are largely representative of the U.S. firm population based on observable characteristics.[4] This is valuable because it documents the prevalence of infringement among businesses and provides insight into the resolution of disputes between firms, rather than between individuals, as typically studied in lab experiments.

Prior to either of the two field experiments, the Agency's settlement requests include two components: a baseline licensing fee for each image, and an additional $400 per-image to recover the Agency's enforcement expenses. The baseline licensing fee depends on the nature of infringing use (e.g., a high-resolution image posted on a website's home page), and is a weighted average of use-related list prices on the Agency's licensing site. The two price components are not broken out separately in the Agency's letters to the infringer. In the *pricing* experiment, the Agency exogenously varies the requested settlement amount by eliminating the $400 per-image enforcement cost in some cases but not in others. The wording of the letter is otherwise kept identical, and no information on the price reduction is provided. Baseline licensing fees range from $380 to $825 per image, so removing the enforcement cost is economically significant.

After the pricing experiment, the Agency drops the enforcement cost for all infringement cases. After this policy change, we worked with the Agency on the second *message* experiment. In this experiment, an extra statement is inserted in the letters sent to infringers in four treatment groups, with no extra statement added for the control group. The first two messages explain that the Agency has forgiven a $400 per-image amount, either as a waiver of enforcement costs or as a discount. To justify the reduction, the messages acknowledge that the infringement may have been unintentional. We cannot separately identify the effect of this acknowledgement from that of the price-reduction information. The last two treatment messages are

[3] Trubek et al. (1983) estimate that 90 percent of civil disputes are settled without filing suit. In our setting, according to Lex Machina (a legal analytics database), the Agency filed twelve cases in the U.S. courts between 2008 and 2013. In contrast, the Agency pursues settlement for about a thousand cases per month in the U.S., resulting in a likelihood of litigation of about 0.02 percent.

[4] According to the 2012 Statistics of U.S. Businesses, 62 percent of U.S. firms have 0-4 employees, 17 percent have 5-9, 10 percent have 10-19, 8.6 percent have 20-99, 1.4 percent have 100-499, and 0.3 percent have more than 500. The corresponding percentages in our data are 50, 17.6, 12, 15, 5.1, and 0.89. Firms in our analysis sample are slightly larger because the Agency stopped pursuing the smallest firms for settlement in January 2014.

the same as the first two, but with a deadline after which the forgiven amount is added back. We do not have a treatment message that explicitly increases the threat of legal sanction, because the Agency moved away from any enforcement approach that was explicitly threatening.

The pricing experiment shows that after 30 days, on average, removing the $400 per-image charge from the requested amount increases the settlement probability by three percentage points (from 12 to 15 percent). However, this increase is offset by the significantly lower revenue for those cases that do settle. As a result, the expected revenue per case is slightly lower for the lower-priced group (although the difference is not statistically significant). In contrast, the message experiment after 30 days shows that adding either of the first two messages increases settlement probability by 12 percentage points, and expected revenue by 80 percent. Adding a deadline to either message further improves the settlement rate by another six percentage points, and increases expected revenue by 130 percent relative to the control group.

The large positive effect from the messages is particularly salient when compared to the small effect observed in the pricing experiment, suggesting that the price of settlement is not the primary factor influencing infringer behavior. One plausible explanation for the small price effect is that infringers perceive the Agency as being less serious when the requested amount is lower. In other words, a lower request may not generate a higher settlement rate if it simultaneously sacrifices an implicit threat of escalation. The addition of a deadline, attached to the last two messages, appears to simultaneously leverage the inducement of a lower price today and signal a willingness to escalate in the future. Interestingly, while the deadline has a large and positive effect on settlement rate for relatively small firms, it has no effect on larger firms (those in the top 40 percent of the firm-size distribution). This is possibly because larger firms are less price sensitive and have more precise knowledge about legal risks and, thus, are not influenced by the implied threat.

Another explanation for the large message effect is that acknowledging possible unintentionality encourages collaborative outcomes by moving the infringers' frame of mind away from calculating legal risks (which are likely to be small in such circumstances). Relative to the control group, the effect of the first two messages (without a deadline) is large and positive both for small "mom-and-pop" businesses and for the largest firms. At least two mechanisms may be at work. First, the acknowledgement of possible unintentionality may encourage intrinsic motivation to correct a mistake, especially given that infringement in this market is likely to be inadvertent (or indirect through third-party web designers) and that monetary claims are small. This mechanism is consistent with findings in the tax compliance literature, and may be especially relevant for small firm.[5] Second, the acknowledgement of possible unintentionality may make the settlement request conform more closely to standard invoices, providing a smoother path for processing

[5]For example, Feld and Frey (2002a,b) argue that paying taxes is quasi-voluntary and show that courteous treatment by the tax authorities (e.g., a presumption of innocence) dampens tax evasion.

payment, instead of initiating an internal legal review. This mechanism may be especially pertinent for large firms.

Finally, the additional settlements induced by the treatment messages substantially reduce social costs. The probability that the Agency engages a professional law firm for enforcement decreases by roughly 5 to 12 percentage points (a 10-27 percent reduction) for the treatment groups relative to the control group. The estimated savings in legal costs (for the Agency and infringing firms combined) range from about 3 to 7 percent of the Agency's in-house settlement revenue. This calculation is likely to underestimate the savings in total social costs because it does not take into account the opportunity cost of time or the stress associated with prolonged disputes.

1.1 Related literature

This paper brings together two sets of literature that investigate the issue of disputes that fail to reach efficient settlement and instead go to trial. The first group of papers are from law and economics (see surveys by Cooter and Rubinfeld (1989) and Spier (2007)). Theory in this literature typically applies bargaining models and attributes settlement failures to asymmetric information (Bebchuk (1984) and P'ng (1983)), divergent expectations (Priest and Klein (1984)), or strategic bargaining (e.g., Cooter et al. (1982) and Spier (1992)). Empirical work in this area has studied determinants of settlement (versus trial), including the expected size of the verdict, litigation costs, asymmetric beliefs, and legal environments (e.g., Danzon and Lillard (1983), Viscusi (1988), Waldfogel (1995), Fournier and Zuehlke (1996), and Eisenberg and Farber (1997)). We contribute to this literature by introducing a novel dataset examining pre-filing settlements and two large-scale field experiments that provide exogenous variation in settlement methods.

The second set of studies on settlement is from behavioral law and economics, and uses lab experiments to study the effects of psychological factors on pre-trial bargaining (see a survey by Jolls et al. (1998)). For example, Loewenstein et al. (1993) and Babcock et al. (1995) study the roles of self-serving biases in bargaining outcomes. Rachlinski (1996) shows that, consistent with prospect theory (Kahneman and Tversky (1979)), disputants are more likely to accept a settlement offer if they view it as a gain relative to provided references. While we do not set out to test a specific behavioral theory and are limited in our ability to identify precise mechanisms, our field experiments are of particular interest in that they capture real-world disputes. While typical lab experiments study individuals, our data describe interactions between firms. Furthermore, our paper highlights the role of culpability in dispute resolution. To the extent that acknowledging possible unintentionality contributes to our results, we provide evidence that is consistent with Korobkin and Guthrie (1994). They find that victims are more likely to settle if there is a sympathetic explanation for the harm they have suffered, reflecting general concerns for fairness and reciprocity (e.g.,

Güth et al. (1982) and Kahneman et al. (1986)).

Recent studies of copyright enforcement in economics have examined the effects of stronger enforcement on file-sharing and sales in the music and movie industries (Bhattacharjee et al. (2006), Danaher et al. (2014), Peukert et al. (2015), and Reimers (2015)). Our study differs from these papers and from the broader consumer piracy literature (see survey by Waldfogel (2012)) in its investigation of infringement by businesses and its focus on settlement outcomes. A related literature in law (also on consumer piracy) studies legal strategies for copyright litigation. It generally argues that sanctions are effective in inducing settlement because statutory damages for willful infringement are high (Lantagne (2004) and DeBriyn (2012)). Fellner et al. (2013), for example, carries out a field experiment that randomizes messages mailed to Austrian households that evade TV licensing. They observe a significant positive effect on compliance as a result of threat, while moral appeal (emphasizing compliance as a matter of fairness) and social-norm information (highlighting the high level of compliance) have no effect. One salient difference between our paper and these studies is the nature of infringement. For example, Fellner et al. (2013) argues that moral persuasion may lack effect in their setting due to a particularly dishonest sample of infringers. In contrast, we are among the first to assess infringement incidences that are likely to be inadvertent.

Most of the IP enforcement literature studies patents, for which ex-post licensing is a common aspect of settlement agreements. For example, by matching filed cases to a control group of similar patents, Lanjouw and Schankerman (2001) show that litigation is more likely when either the value of a patent or the benefit of a tough reputation is higher. Similarly, Cohen et al. (2014) show that non-practicing entities target companies based on the expected profitability of litigation (e.g., firms with more cash).[6] Our paper instead contributes empirical evidence to the literature on copyright enforcement and settlement, which is much less understood. It is important to note that the two contexts, though conceptually similar, differ in substantial ways. First, liability in our context is generally clear, given the accuracy of detection technology and the simpler nature of copyright infringement; liability is much harder to establish in patent disputes. Second, patent claims (at least for complaints filed in court) are typically much larger. These differences may limit the effectiveness of our particular interventions in patent contexts.

Finally, our paper joins the small but increasing body of work that uses large-scale field experiments to study the effects of small, low-cost interventions (especially communication). Other enforcement contexts include tax compliance and enforcement (Hallsworth et al. (2014) and Perez-Truglia and Troiano (2015)) and loan repayments (Cadena and Schoar (2011) and Karlan et al. (2015)).[7] DellaVigna and Gentzkow

[6] Examples of theoretical papers on patent litigation include Meurer (1989) and Bessen and Meurer (2006), and empirical studies include the role of cross-licensing (Lanjouw and Schankerman (2004) and Galasso et al. (2013)), non-practicing entities (Bessen et al. (2011) and Scott Morton and Shapiro (2014)), and effects on the value of patenting and innovation (Lanjouw and Lerner (1998), Hall and Ziedonis (2001), and Cockburn and MacGarvie (2009)).

[7] Studies in other settings include advertising (Bertrand et al. (2010)) and donations (Frey and Meier (2004), and Falk (2007)).

(2010) surveys empirical evidence on the effects of persuasive communication in several domains and defines a 'persuasion rate' that harmonizes results across studies.[8] Compared to the 17 previous studies surveyed in that paper, in which the persuasion rate ranges from 0.7 to 29.7 percent, our results show generally high persuasion rates of 14 to 20 percent for messaging within the first 30 days.

The rest of the paper proceeds as follows. Section 2 provides the background and introduces the two experiments. Section 3 describes the sample and the variables. Section 4 reports the results. Section 5 discusses plausible explanations and effects on social costs, and Section 6 concludes.

2 Background and Experiments

The Agency monitors the online use of images that satisfy the following criteria: (i) the image is governed by a 'rights-managed' license model, which allows a licensee a one-time use of the image for a fee that depends on the specified scope of use (e.g., a high-resolution image placed on a firm's home page for three months); and (ii) the Agency distributes the image exclusively, allowing the Agency to determine whether the use is unauthorized by checking its own licensing database. Images that satisfy these criteria are the most expensive ones; they account for less than five percent of the number of images in the Agency's complete portfolio. The Agency pursues enforcement only if infringement involves commercial use of an image, and it drops cases that involve small non-profit organizations or current customers.[9]

For cases that it pursues, the Agency sends a letter to each infringing firm, which (1) informs the firm of the detected infringement, (2) requests the removal of the image(s) from the firm's website, and (3) asks for a settlement to cover the (past) infringing use. Prior to both experiments, the requested amount was the baseline licensing fee plus an additional $400 charge per image to recoup enforcement costs incurred by the Agency. In the letter, a single amount is displayed instead of a breakdown between these two components.[10] Table 1 lists the six use categories for a single image, the corresponding baseline licensing fees for the sample period we study, and the percentage of cases from our combined experimental groups (described later) that corresponds to each use category. When multiple images are involved in a case, the licensing fee is the sum for all the images. The list prices posted on the Agency's website also vary across other dimensions, including image type (e.g., theme) and duration of use. The fee offered for an infringing use is a weighted average of a few listing prices from within one of the six categories listed in table 1.

[8] The notion of a 'persuasion rate' estimates the percentage of receivers that change behavior among those that receive a message and are not already persuaded. For a binary outcome with treatment group T and control group C, the persuasion rate defined by DellaVigna and Gentzkow (2010) is $f = 100 \times \frac{y_T - y_C}{e_T - e_C} \frac{1}{1-y_0}$, where e_i is the share of group i receiving the message; y_i is the share of group i adopting the behavior of interest; and y_0 is the share that adopts in the absence of a message. Where y_0 is not observed, it is approximated by y_C.

[9] Current customers may infringe if the use is beyond the scope of the current license.

[10] The letter comprises seven pages in total, with the invoiced amount displayed on the third page.

Table 1: Licensing fee

Image-use category	Licensing fee	% of cases
Low resolution & secondary page	$380	28
High resolution & secondary page	$475	31
Low resolution & home page	$565	8
Low resolution & repeat pages	$665	4
High resolution & home page	$775	12
High resolution & repeat pages	$825	7
Multi-image cases	Varies	10

Notes: This table describes the licensing-fee part of the total requested amount. The first six rows apply to single-image cases. The actual licensing fee may deviate slightly due to taxes that differ by state. The corresponding percentages are based on our combined experimental groups, described later.

Two experiments

2.1 The requested amount

Between October and December 2013, for cases in the U.S. and Canada, the Agency exogenously selected a subset of cases for which it removed the enforcement cost completely (the 'Low Price' group); for the rest of the cases, it continued to add $400 per image on top of the licensing fee (the 'High Price' group). The letters to the two groups are otherwise identical. That is, no explanation is given about the price reduction for the Low Price group. We refer to this as the *pricing* experiment. This exogenous price variation is valuable because, otherwise, it is not possible to separate the effect of a price change from the way in which the image is used, as the licensing fee is completely determined by the image use category.

2.2 The messages

In January 2014, the Agency implemented two changes to its enforcement policy. First, the Agency no longer requests any settlement amount from "less valuable" infringers, although it still requests removal of the images from the infringer's website. More precisely, infringers are categorized into five 'tiers' using a formula from Dun & Bradstreet, with tier 1 being the most likely to settle and tier 5 the least likely, and the Agency stops pursuing settlement from firms in tiers 4 and 5.[11] Second, it no longer adds any enforcement cost to cases it pursues. Following the two policy changes, we collaborated with the Agency to conduct a second experiment. This experiment was also conducted for infringement cases in the U.S. and Canada.

In the second, *message* experiment, we maintain the low price (i.e., with no additional $400 charge), while also providing information about the price reduction. Cases are allocated into four treatment groups

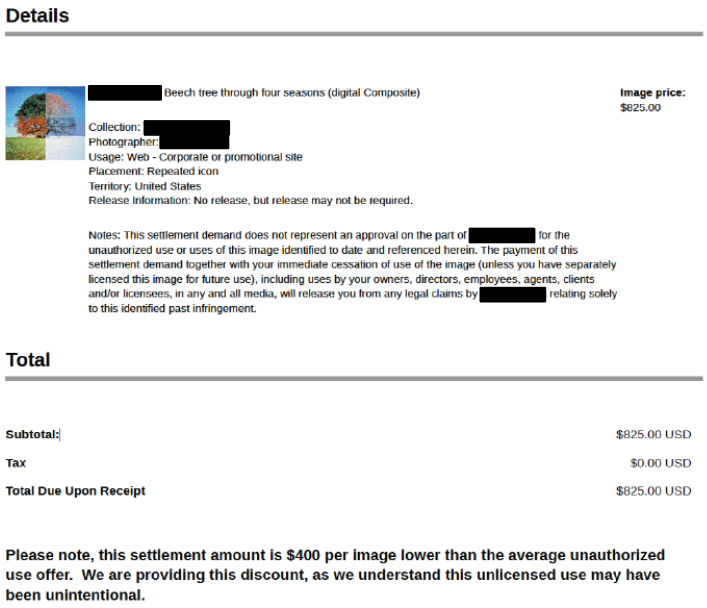
[11] We do not observe the formula that maps firm observables to tiers, but firms with lower annual revenues tend to have a higher tier number.

and one control group. All groups are requested to pay the same licensing fee (given the image use), although the letter's wording varies. The wording used for the control group is the same as in the pricing experiment; it provides no information about the price reduction. For the treatment groups, one of four extra messages (in boldface) is placed just below the invoiced amount. Figure 1 shows the positioning of the message in the letter. The four messages differ by how the forgiven amount is presented, and whether a 22-day deadline is imposed. As justification for the price reduction, all four messages acknowledge that the infringement may have been unintentional. It is, thus, not possible to separately identify the effect of this acknowledgement from that of the price-reduction information. The four messages are:

1. (Control group) no extra message.
2. (Waive the Cost) Please note, we are only charging the average licensing fee for commercial use of the rights-managed image(s) found on your website. The Agency has incurred additional costs of $400 per image related to the pursuit of this matter; we are currently waiving this cost, as we understand this unlicensed use may have been unintentional.
3. (Waive the Cost + Deadline) Please note, we are only charging the average licensing fee for commercial use of the rights-managed image(s) found on your website. The Agency has incurred additional costs of $400 per image related to the pursuit of this matter; we are currently waiving this cost, as we understand this unlicensed use may have been unintentional. However, cases not settled before xx-xx-xx will be charged the additional $400 cost per image related to the pursuit of this matter.
4. (Discount) Please note, this settlement amount is $400 per image lower than the average unauthorized use settlement offer. We are providing this discount, as we understand this unlicensed use may have been unintentional.
5. (Discount + Deadline) Please note, this settlement amount is $400 per image lower than the average unauthorized use settlement offer. We are providing this discount, as we understand this unlicensed use may have been unintentional. However, cases not settled before xx-xx-xx will no longer be eligible for this discount.

An infringer may contact the Agency to try to negotiate a lower price. Employees in the Agency's compliance department, "case handlers," respond to calls from infringers and have latitude to negotiate with them as they see fit. According to the Agency, case handlers are not given instructions to treat groups differently in either experimental period, mainly so as not to confuse the handlers. In principle, the handlers can observe the group to which a particular case is allocated. Our interviews with the handlers, however, suggest that they do not incorporate that information during conversations with infringers.

Figure 1: Settlement demand letter example



Details

[redacted] Beech tree through four seasons (digital Composite) — **Image price:** $825.00

Collection: [redacted]
Photographer: [redacted]
Usage: Web - Corporate or promotional site
Placement: Repeated icon
Territory: United States
Release Information: No release, but release may not be required.

Notes: This settlement demand does not represent an approval on the part of [redacted] for the unauthorized use or uses of this image identified to date and referenced herein. The payment of this settlement demand together with your immediate cessation of use of the image (unless you have separately licensed this image for future use), including uses by your owners, directors, employees, agents, clients and/or licensees, in any and all media, will release you from any legal claims by [redacted] relating solely to this identified past infringement.

Total

Subtotal:	$825.00 USD
Tax	$0.00 USD
Total Due Upon Receipt	$825.00 USD

Please note, this settlement amount is $400 per image lower than the average unauthorized use offer. We are providing this discount, as we understand this unlicensed use may have been unintentional.

2.3 Allocation process

The Agency allocated cases to different groups weekly during each experiment. For a given week, some cases were assigned as 'trial cases,' and all other cases were assigned to a "default" group. The "default" group is the High Price group in the pricing experiment, and it is the control group in the message experiment. After cases were allocated to the "trial" and "default" groups, cases in the "trial" group from each combination of infringer tier and price category were placed sequentially into each treatment arm.[12] The number of cases allocated as trial cases each week depended on the Agency's overall workload. In the pricing experiment, we use data from five weeks during which trials were conducted. In the message experiment, trials were conducted every week between the end of January and the end of March 2014.

It is not ideal that the allocation procedure did not follow a strict randomization protocol. In particular, we do not know exactly how cases were assigned to the trial and default groups in the first step. According to the Agency, trial cases were selected sequentially from a list of infringements generated each week by the detection software, which had been screened to eliminate authorized uses. The Agency is not aware of any dependence between the ordering of cases generated by this process and settlement likelihood. This is reasonable because neither the detection technology nor the Agency's internal authorized-use screening

[12] In the pricing experiment, there was also a small treatment arm in which a $100 enforcement cost was added to the licensing fee. The Agency discontinued this price point after a short time. We exclude these observations in the current results, but including them does not qualitatively change the basic conclusions.

team has information on the potential settlement outcomes. Nevertheless, to mitigate this issue, we conduct balance tests on the requested amount and infringer characteristics in order to verify that the treatment and control groups are consistent with random allocation (see table 3 below).

3 Sample and variables

Our dataset includes all the infringement instances for which the Agency pursued settlement during the two experiments. The information includes the initial requested amount, characteristics of infringing use, and settlement outcomes. The data on infringer characteristics come from Dun & Bradstreet and include annual sales, number of employees, age, industry, and firm location.

The final sample excludes cases as defined here. First, because the Agency stops pursuing settlement from infringers in tiers 4 and 5 during the message experiment, we also remove these cases in the pricing experiment, in order to make the populations of the two experiments comparable. This excludes 30 percent of all cases, leaving 5,660 cases. We further exclude cases for which (1) the mail is undeliverable (10 percent of the 5,660 cases); (2) the case is later closed because the infringer is able to show a valid license or permission from the photographer (6.3 percent); (3) there is no information on the firm's annual sales (8.5 percent); or (4) more than six images are involved, ensuring that the maximum number of images in all treatment and control groups is the same (19 observations). The final sample includes a total of 4,278 cases, of which 1,983 cases come from the pricing experiment and 2,295 cases from the message experiment.[13]

Panel 1 in table 2 summarizes the requested amount, inclusive of the additional $400 charge when applicable, and the number of images involved in each case. The median requested amount is $699. Ninety percent of all cases involve a single image, and nearly 60 percent are in the two lowest-priced use categories (as shown in table 1). For cases involving multiple images, the median number of images is two.

Panel 2 in table 2 summarizes infringer size and age. The median firm has $320K in annual sales; the median number of employees is five and the median firm age is 11 years. The distribution of several firm characteristics (e.g., annual sales) is highly skewed. Annual sales and the number of employees are strongly correlated (the correlation is 0.92), so we use only the former to measure firm size. The top ten industries (defined by two-digit SIC codes and not shown in table 2) are responsible for 63 percent of all cases, and they are largely service industries. The three states with the largest number of infringers are California, Florida, and New York.

Table 3 compares key pre-treatment variables for different treatment and control groups for the two ex-

[13]Table A7 in the Online Appendix reports results using unfiltered data. Our basic conclusions do not change because of any exclusions.

Table 2: Summary statistics of case and firm characteristics

Variables	N	Mean	SD	Min	Median	Max
Panel 1: Price and image use						
Requested amount	4,278	748.0	438.3	354.9	699.0	5040.0
Single-image case	4,278	0.90	0.30	0	1	1
Image count (if a multi-image case)	437	2.49	0.85	2	2	6
Panel 2: Firm characteristics						
Annual sales ($000's)	4,278	23165.7	423542.4	0.3	320.0	23.7M
Employees	4,278	112.9	1348.5	0	5	67800
Age	4,155	17.8	21.1	0	11	271

Notes: This table summarizes the requested amount, the number of images, and the characteristics of the firms using cases in both experiments. Firm characteristics are provided by Dun & Bradstreet based on the name and address of the infringer. Industry and state (province) dummies are omitted.

Table 3: Balance check

(a) Pricing experiment

Group	N	Baseline price	Sales, $1,000		log(Sales, $)	Age		log(Age+1)
			mean	median	mean	mean	median	mean
Low Price	707	642.6	15533	324	13.1	18	11	2.5
High Price	1,276	623.2	13797	281	13.0	17	11	2.5
(p-value)		(0.22)	(0.84)	(0.44)	(0.41)	(0.42)	(0.58)	(0.92)

(b) Message experiment

Group	N	Baseline price	Sales, $1,000		log(Sales, $)	Age		log(Age+1)
(p-value, diff. to control)		(=Requested amount)	mean	median	mean	mean	median	mean
Control	598	590.7	30614	390	13.4	19	11	2.6
Waive the Cost	378	636.8	50682	385	13.4	19	11	2.5
		(0.03)	(0.40)	(0.91)	(0.93)	(0.86)	(0.83)	(0.86)
Waive the Cost + Deadline	414	595.3	7951	290	13.1	16	9	2.5
		(0.82)	(0.16)	(0.13)	(0.03)	(0.04)	(0.17)	(0.06)
Discount	448	588.4	10112	305	13.2	18	12	2.5
		(0.90)	(0.19)	(0.13)	(0.13)	(0.48)	(0.73)	(0.85)
Discount + Deadline	457	580.2	55206	360	13.0	17	11	2.5
		(0.59)	(0.61)	(0.47)	(0.01)	(0.18)	(0.59)	(0.37)

Notes: Panel (a) summarizes the baseline price (i.e., the licensing-fee portion of the total requested amount), the number of images, and firms' basic characteristics between the two groups in the pricing experiment. Two-sided t-test p-values are reported in parentheses. Panel (b) summarizes the same information by group for the message experiment. Two-sided t-test (or median test) p-values are between each treatment group and the control group.

periments. Because sales and age are skewed, we summarize both the mean and the median of the variables and the mean of their log transformation. Panel (a) shows that, in the pricing experiment, there is no significant difference in the average baseline price (i.e., the licensing-fee portion of the requested amount) or the summary statistics of the distributions of firm size and age between the treatment and control groups.

Panel (b) summarizes the same information for the control and four treatment groups of the message experiment, along with p-values of two-sided t-tests (or median tests) comparing each treatment group to the control group. The variables are similar between the treatment groups and the control, with the following exceptions: compared to the control, (1) the average requested amount in the 'Waive the Cost' group is significantly higher; (2) firms in the two deadline message groups are smaller in log(Sales), even though the mean and the median of the level of sales are not statistically different from the control; and (3) firms in the 'Waive the Cost + Deadline' group are younger.

Consultation with the compliance team at the Agency does not identify any sources of these incidences of imbalance. If the suggestions of imbalance here are not false positives, they are likely to make our estimates of the treatment effects more conservative. This is because our results consistently show that settlement is less likely for smaller and younger firms and for cases with a higher requested amount. We further verify that the regression results without using controls (reported in table A1 in the Online Appendix) are similar to those using controls. Both of these are also consistent with the raw data, indicating that the allocation is not systematically correlated with observable variables.

About 30 days after the first letter, the Agency sends follow-up letters unless the case has already been settled. Eventually, cases that do not settle are sent to external law firms.[14] In the main analysis, we focus on the 30-day outcomes as a clean assessment of the interventions. Two factors confound assessments made at later time periods. First, more follow-up letters are sent to groups that have lower initial settlement rates. These endogenously-determined follow-up actions are designed to encourage settlement and should cause the worst-performing groups to 'catch up' over the longer horizon. Second, software constraints at the Agency accidentally resulted in substantially more intense follow-up for groups designated as the 'default' group in each experiment.[15] In the Discussion section, we provide estimates that bound the longer-term outcomes, including the 120-day settlement outcomes and the effects of the interventions on the probability of escalation to an external law firm.[16]

Table 4 summarizes 30-day outcomes. Taking the settlement probability into account, the average revenue per case is $104. For settled cases, the average revenue collected is $555. Infringers negotiate a lower settlement amount for 38 percent of settled cases, and for those cases, the median reduction is 36 percent. The overall settlement rate is 19 percent. Broken down by firms' responses to the initial letter, seven percent settle directly without contacting the Agency; 27 percent contact the Agency (43 percent of which settle by

[14] In our sample, 90 percent of cases sent to external law firms are sent between 58 and 159 days after the first letter.

[15] Recall from section 2.3 that the 'default' group was the High Price group in the pricing experiment and the control group in the message experiment. We provide a detailed description of the Agency's follow-up actions after 30 days in Section C of the Online Appendix.

[16] All cases in our sample are older than five months. For cases that settle within our sample period, 51 percent settle by 30 days, and 98 percent settle by 120 days.

30 days); and 66 percent ignore the letter. Our analysis pools the two types of settlement together (that is, direct settlement without contacting the Agency and settlement after making contact).

Table 4: 30-day outcomes

(a) 30-day revenues

Variables	N	Mean	SD	Min	Median	Max
Revenue per case	4,278	104.2	255.8	0.0	0.0	2992.3
Revenue (if settled)	803	555.2	313.4	91.0	475.0	2992.3
Paid a reduced amount? (if settled)	803	0.38	0.49	0	0	1
Negotiated reduction % (if reduced)	305	0.36	0.17	0.02	0.35	0.88

(b) 30-day settlement rate and infringer response

Contact/Outcome	Settle (19%)	Not Settle (81%)
Contact Agency (27%)	12%	15%
Do not contact Agency (73%)	7%	66%

Notes: Panel (a) summarizes the 30-day revenues using cases in both experiments. Revenue is coded as $0 if the case does not settle within 30 days. Panel (b) summarizes the 30-day settlement rate and detailed responses by infringers.

4 Results

For each of the two experiments, we present the baseline results for an average case. We then report the heterogeneity (or lack thereof) in the effects across two dimensions: the size of the infringing firm, and the extent of the price variation (in percentage terms) that is either induced by the pricing experiment, or implied by the message experiment.

4.1 Pricing experiment

Table 5 shows that, on average, the 30-day settlement rate of the Low Price group is three percentage points higher than that of the High Price group (15.6 vs. 12.5 percent with a p-value of 0.054). Although this represents a 25 percent increase, the difference is only marginally significant. For settled cases, the average revenue collected is $211 lower in the Low Price group. This suggests that more extensively negotiated reductions for the higher-priced group offset roughly half of the difference in the requested amount. Despite the higher settlement probability for the Low Price group, the expected 30-day revenue of a case is $10 lower for the Low Price group, but the difference is not statistically significant.

Table 6 provides OLS and logit regression analyses of the effect of charging the High Price on revenue and the likelihood of settlement. The marginal effect of the High Price dummy is consistent with the raw

Table 5: Price effects on settlement outcomes (raw data)

	Low Price	High Price	Difference	p-value
All cases:				
Settlement probability	0.156	0.125	-0.031	0.054
Revenue	82.6	92.4	9.8	0.432
N	707	1,276		
Settled cases:				
Revenue	530.9	741.5	210.6	0.000
Whether discounted	0.36	0.64	0.28	0.000
Discount	81.9	284.4	202.5	0.000
N	110	159		

Notes: This table summarizes the 30-day outcomes of the High Price and the Low Price groups in the pricing experiment.

data. The control variables in the regressions include the baseline licensing fee; characteristics of the infringing use, including image-use category indicators and number of images; characteristics of the infringer, including log(annual sales), log(age + 1), ten 'top' industry indicators (defined by two-digit SIC codes), a U.S. dummy, and ten 'top' state/province indicators; and week dummies indicating when the settlement letter was mailed.[17] For all regressions throughout the paper, we use the same set of control variables and report robust standard errors, clustered by industry.

On average, firms appear to be not price-elastic, in the sense that the substantial reduction in the requested amount does not result in higher expected revenue. Firms are more likely to settle if they are offered a lower price. However, the increase in the settlement probability is not large enough to offset the lower revenue for the cases that do settle. The price elasticity of the expected revenue is the sum of the elasticity of the settlement probability and the elasticity of the expected revenue conditional on settling.[18] The two elasticities based on the regression results are, respectively, -0.28 and 0.71, implying a positive elasticity of the expected revenue.

Table 6 also shows that larger firms are significantly more likely to settle and less likely to negotiate a discount conditional on settling; thus, they are associated with a significantly higher expected revenue. Older firms are also significantly more likely to settle, but generate only a slightly higher revenue (p-value is 0.15). The coefficient for log(Baseline price) is not always significant, because it captures only residual differences in the licensing after controlling for the complete set of image-use dummies (e.g., taxes that

[17] 76 firms (four percent) do not have age information. To preserve the sample size, a dummy variable indicating that the age information is missing is included in the regressions.

[18] The expected revenue of a case is the settlement probability multiplied by the expected revenue, conditional on settling. We denote expected revenue as $R(p)$, where $R(p) = \Pr(p)\hat{R}(p)$; p is the requested amount; and $\hat{R}(p) \leq p$ because negotiation is allowed. The decomposition of the price elasticities is then $\frac{\partial R(p)}{\partial p}\frac{p}{R(p)} = \frac{\partial \Pr(p)}{\partial p}\frac{p}{\Pr(p)} + \frac{\partial \hat{R}(p)}{\partial p}\frac{p}{\hat{R}}$.

Table 6: Price effects on settlement outcomes

Dependent variable	Settle	Revenue for settled cases	Revenue
	Logit	OLS	OLS
	(1)	(2)	(3)
High Price dummy	-0.03*	255.84***	13.72
	(0.02)	(34.79)	(11.45)
Use cat 2: High res. + secondary page	0.02	-131.03	3.65
	(0.02)	(80.16)	(27.71)
Use cat 3: Low res. + home page	-0.03	-325.96**	-24.43
	(0.05)	(143.73)	(49.89)
Use cat 4: Low res. + repeat pages	0.04	-285.46	30.10
	(0.06)	(178.13)	(74.11)
Use cat 5: High res. + home page	-0.00	-371.27	-12.47
	(0.07)	(249.95)	(94.20)
Use cat 6: High res. + repeat pages	0.00	-366.29	-4.53
	(0.07)	(263.55)	(93.50)
Use cat 7: Multi-image cases	0.03	-557.20***	20.58
	(0.08)	(190.23)	(99.43)
log(Baseline price)	-0.04	893.66***	41.77
	(0.09)	(319.06)	(124.66)
Number of images	-0.01	346.52**	17.09
	(0.04)	(156.38)	(75.92)
log(Annual sales)	0.01***	23.07*	17.20***
	(0.00)	(12.63)	(3.38)
log(Age+1)	0.02**	-7.09	10.03
	(0.01)	(13.25)	(6.75)
Missing-age dummy	0.10**	-97.38	46.87
	(0.04)	(99.54)	(36.44)
Top 10 industry dummies	Y	Y	Y
Top 10 state (province) dummies	Y	Y	Y
U.S. dummy	Y	Y	Y
Mail week dummies	Y	Y	Y
Price elasticity	-0.28	0.71	0.24
Mean(dependent variable) for the Low Price group	0.155	530.9	82.6
Adj R-squared		0.645	0.031
N	1983	269	1983

Notes: This table reports the results (marginal effects) from the following regressions using observations from the pricing experiment. The dependent variables are (1) whether the case settles within 30 days; (2) the revenue for settled cases; and (3) the revenue of a case (coded as $0 if not settled in 30 days). The independent variable is a dummy indicating the High Price group (the Low Price group is the omitted group). Baseline price is the licensing-fee portion of the requested amount. Column (2) uses cases that settle in 30 days. The corresponding price elasticities are calculated based on the formula $\frac{\Delta DV}{\Delta p}\frac{p}{DV}$, in which ΔDV is the estimated marginal effect of the High Price dummy, Δp is the difference between the mean requested amounts of the two groups, and p and DV are the respective sample averages for the Low Price group. All models report robust standard errors, clustered by industry (defined by the two-digit SIC codes). ***, **, and * are, respectively, significance levels of 1%, 5%, and 10%.

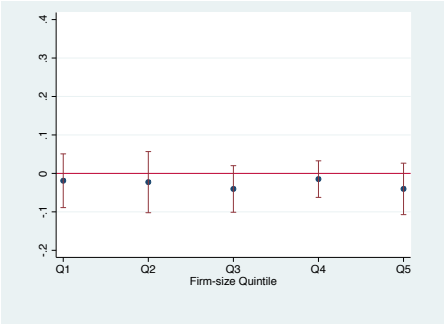
differ by state).

Price effects for different subsamples

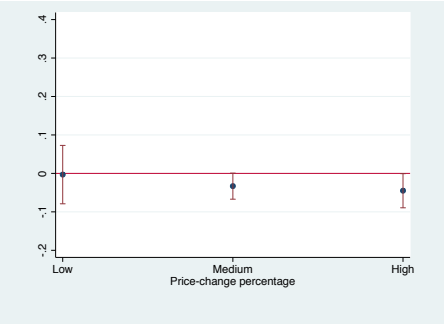
In this section, we analyze subsamples of the data based on (i) firm size and (ii) the size of the price reduction relative to the licensing fee(s) associated with a case. Firm size, measured by annual sales, varies substantially in our data.[19] Since the reduced amount is uniform per image, the relative price reduction is determined by the baseline licensing fees associated with a case, and it ranges from 43 to 107 percent.

Figure 2: Price effects on settlement outcomes for different subsamples

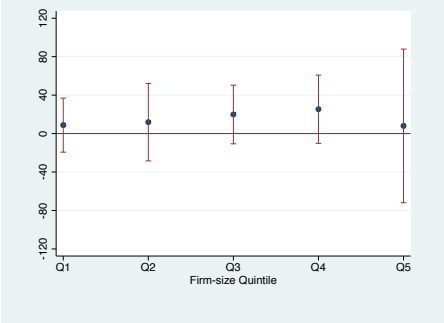
(a) Settlement probability (by firm size)



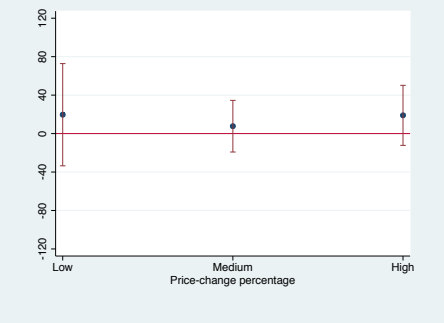
(b) Settlement probability (by price-change %)



(c) Revenue per case (by firm size)



(d) Revenue per case (by price-change %)



Notes: These figures report marginal effects of the High Price dummy on 30-day settlement outcomes: settlement probability and expected revenue per case. Q1-Q5 indicate firm-size quintiles based on annual sales, with Q1 including the smallest firms and Q5 the largest. Price-change percentage is ($400×number of images)/licensing fee. Bands indicate 95 percent confidence intervals.

The results for an average case hold across these different subsamples (i.e., when stratifying either by firm size or by price-change percentage). Figure 2a plots the marginal effects of the High Price dummy on the settlement probability by firm-size quintiles, with Q1 being the smallest firms in terms of annual sales and Q5 the largest (this and all other regression results for different subsamples are reported in the Online Appendix). For all size quintiles, the settlement probabilities are not statistically different between the two different price groups. Consequently, for all quintiles, the expected revenue is higher for the higher-priced

[19] Firms at the 20th, 40th, 60th, and 80th percentiles have $111K, $190K, $480K and $1.7m in annual sales, respectively.

group, though none of the differences are statistically significant (see figure 2c). Figure 2b plots the marginal effects of the High Price dummy on the settlement probability by the price-change percentage.[20] The effect of the price change on settlement probability increases monotonically in magnitude with the percentage and reaches statistical significance (at the .05 level) when the percentage is medium or high (above 0.70). However, even in those ranges, the effect of the price change on expected revenue per case is statistically indistinguishable from zero (see figure 2d).

4.2 Message experiment

We analyze three comparisons within the message experiment: (1) comparing either non-deadline message to the control group; (2) comparing a deadline message to its corresponding non-deadline version; and (3) comparing the two non-deadline messages to each other.

Table 7: Message effects on settlement probability (raw data)

(a) Comparing the treatment groups to the control

Group	N	Settlement probability	Difference to Control	(p-value) diff. to control
Control	598	0.14		
Waive the Cost	378	0.24	0.09	(0.00)
Waive the Cost + Deadline	414	0.29	0.14	(0.00)
Discount	448	0.24	0.10	(0.00)
Discount + Deadline	457	0.30	0.16	(0.00)

(b) Comparing between different treatment groups

	No deadline	With deadline	Effect of a deadline	(p-value)
Waive the Cost	0.24	0.29	0.05	(0.113)
Discount	0.24	0.3	0.06	(0.046)
Difference in presenting the forgiven amount	-0.003	-0.01		
(p-value)	(0.91)	(0.68)		

Notes: (A) Summary of 30-day settlement probabilities for the four treatment groups and the control group in the message experiment. Two-sided t-test p-values are between each treatment group and the control group. (B) Comparison of the settlement probabilities between different treatment groups and p-values for these differences. For example, for the waive-the-cost explanation, the difference in 30-day settlement probability between the deadline and the non-deadline versions is 0.05, and the p-value of the difference is 0.113.

Table 7a shows that, for an average case, relative to the control group, adding either of the two non-deadline messages increases the settlement probability by about nine percentage points (a 64 percent increase); adding either of the two deadline messages improves the 30-day settlement probability by about 15

[20]The requested amount is grouped into three categories due to the presence of six discrete baseline prices. We categorize 30 percent of the cases as low, 40 percent as medium, and 30 percent as high. The separating thresholds are 0.70 and 0.90.

percentage points (a 107 percent increase relative to the control group); and all the improvements are statistically significant.[21] For each of the two different ways of presenting the forgiven amount (i.e., "Waive-the-Cost" and "Discount"), the difference between the deadline and non-deadline versions is about five or six percentage points (p-values are 0.11 and 0.05, table 7b). However, given that a deadline is added (or not), there is no significant difference between the two different presentations of the forgiven amount. Regression results on settlement probability including control variables are shown in the first two columns of table 8. The estimated marginal effects are slightly larger than that from the raw data.

Table 8: Message effects on settlement outcomes

Dependent variable	Settle		Revenue for settled cases		Revenue	
	Logit		OLS		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Waive the Cost	0.12***		-1.89		65.54***	
	(0.03)		(17.50)		(20.54)	
Waive the Cost + Deadline	0.18***		7.31		86.08***	
	(0.03)		(17.05)		(15.53)	
Discount	0.13***		17.11		66.67***	
	(0.03)		(20.21)		(16.93)	
Discount + Deadline	0.19***		3.72		88.54***	
	(0.02)		(14.72)		(15.60)	
Non-deadline message		0.12***		4.89		65.17***
		(0.03)		(17.21)		(19.80)
Deadline message		0.18***		1.90		86.42***
		(0.03)		(16.99)		(15.45)
Control variables	Y	Y	Y	Y	Y	Y
Mean(dependent variable) for the control group	0.14	0.14	485.2	485.2	68.1	68.1
Adj R-squared			0.81	0.81	0.07	0.07
N	2291	2291	534	534	2295	2295

Notes: This table reports the results (marginal effects) from three sets of regressions using the message-experiment sample. The dependent variables are, respectively, whether the case settles in 30 days; the revenue for settled cases; and the revenue of a case (coded as $0 if not settled in 30 days). In columns (1), (3) and (5), the independent variables are four dummies indicating the treatment messages (the default is the control group). In the other columns, Non-deadline message equals one if the case belongs to either of the two non-deadline message groups; and Deadline message equals one if the case belongs to either of the two deadline message groups. Columns (3) and (4) use cases that settle in 30 days. In all regressions, the control variables are log(Requested amount), image count, log(Annual sales), log(Age+1), missing-age dummy, dummies indicating the top ten industries, the top ten states (provinces), the U.S., and the mail date. All models report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.

For settled cases, the conditional settlement revenue is similar across all groups (column (3) in table 8). In terms of the expected revenue per case (taking the settlement probability into account), the inclusion of

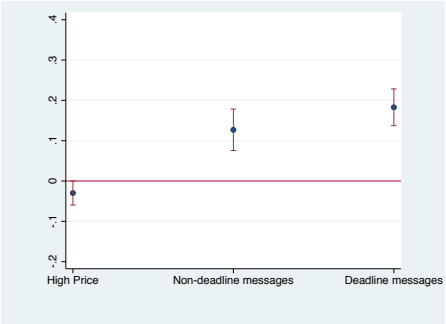
[21] The deadline is set for the 22nd day. Pre-deadline outcomes are reported in a later section and, in summary, show that the two deadline messages improve both the settlement rate and the expected revenue by about 150 percent relative to the control group (0.25 versus 0.10, and $120 versus $47).

a non-deadline message increases expected revenue by $65, while a deadline message increases expected revenue by $86 (columns (5) and (6) in table 8).

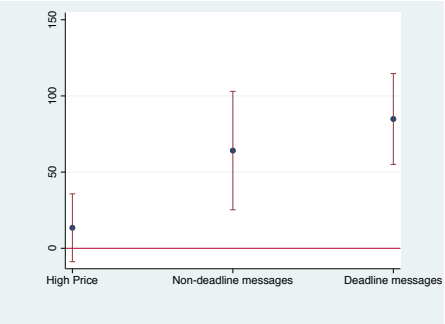
Figure 3 contrasts the large positive effects of the treatment messages with the small price effect from the pricing experiment. The comparison pools the two non-deadline messages together and the two deadline messages together. The baseline group in both experiments (normalized at 0) adds no additional charge to recover enforcement costs, and does not include any extra message. As shown previously, Figure 3a illustrates that the effect of a substantial reduction in the requested amount results in a small difference in the settlement probability, while the effect of including these messages is large and positive. Figure 3b provides the comparison for expected revenue. Although the estimates are obtained from different samples, the two sample periods are close in time, and the settlement rates of the baseline group in the two experiments are not statistically different from each other (0.156 vs. 0.141 with a p-value of 0.44).

Figure 3: Comparing the price and message effects

(a) Settlement probability



(b) Expected revenue per case



Notes: This figure compares the marginal effects on the 30-day settlement probability and the expected revenue of the High Price dummy estimated from observations in the pricing experiment to that of the non-deadline message dummy and the deadline message dummy estimated from the message-experiment sample. The baseline group in both samples (normalized at 0) adds no additional charge to recover enforcement cost, and does not include any extra message. Refer to tables 6 and 8 for the regression results.

The rest of this section examines the effects of the messages for different subsamples. We focus on the settlement probability because the requested amount is the same across different groups given the same use category and the revenue collected per settled case does not differ significantly across groups. Thus, the effects on expected revenue remain consistent with those for the settlement probability.

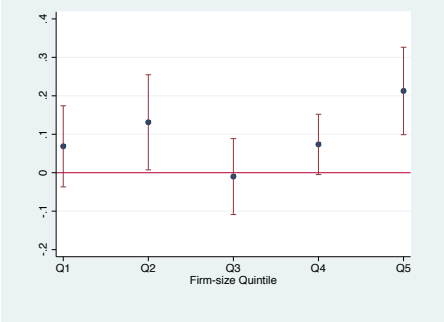
Comparing the non-deadline messages to the control for different subsamples

Overall, the results show that the effect of a non-deadline message is relatively flat across different subsamples, both across firm size and the relative size of the forgiven amount. Figure 4a shows that, relative to the control, adding a non-deadline message has an effect of at least seven percentage points for all firms,

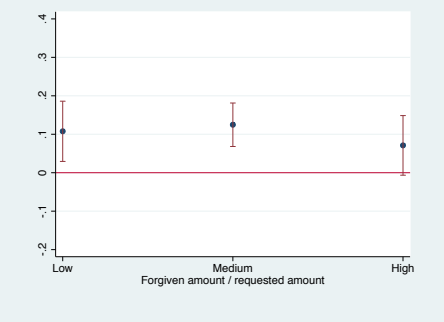
except those in the third quintile of the firm size distribution.[22] Compared to the firms in the first quintile, the message effect for larger firms is not statistically different, except for the firms in the top quintile, where it is 14 percentage points larger (p-value of the difference is 0.008).

Figure 4: Effects of a non-deadline message on settlement probability for different subsamples

(a) By firm size



(b) By $\frac{\text{forgiven amount}}{\text{requested amount}}$



Notes: This figure plots the marginal effect of the dummy variable that equals one if the case belongs to either of the two non-deadline message groups (the baseline is the control group) on the 30-day settlement probability. Q1 are the smallest firms and Q5 the largest. Reported bands reflect 95% confidence intervals.

Recall that the forgiven amount is a uniform $400 per image, while the licensing fee varies from $380 to $825 per image. Figure 4b shows that the effect of a non-deadline message is large and statistically significant, regardless of the proportion of the forgiven amount relative to the requested amount. The differences in the effects between different ranges are small and not statistically different from zero.

Comparing messages with and without a deadline for different subsamples

We evaluate the impact of adding a deadline after which the price-reduction offer expires by comparing outcomes for the groups with and without a deadline, holding the rest of the message fixed.[23] The deadline is set for the 22nd day after the first letter is mailed. In practice, the Agency did not take action immediately after the deadline; nor did it increase the requested amount in subsequent letters. Thus, we focus on the 22-day outcome to provide the cleanest comparison.

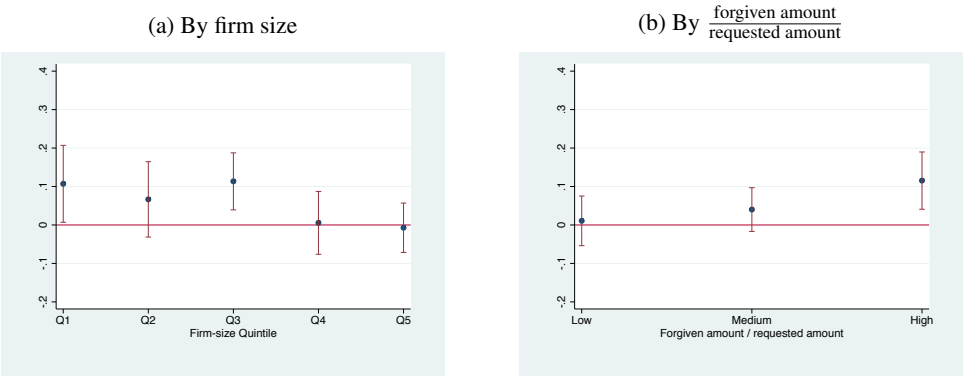
For an average case, the addition of a deadline increases the 22-day settlement likelihood by 5.4 per-

[22]The figure is based on logit regression results that use cases in the control group and the two non-deadline message treatment groups. The regression includes the interaction terms between the treatment indicator (which equals one if the case belongs to either of the two non-deadline message groups and zero if it belongs to the control group) and firm-size quintile indicators. The results are reported in table A3 in the Online Appendix. A linear probability model with the interaction terms produces similar results.

[23]Note that this may underestimate the stand-alone effect of adding a deadline because infringers may be already motivated by the rest of the message to settle. One could alternatively create a second control group in which a deadline alone is added to the control letter.

centage points (the p-value is 0.007).[24] This effect, however, differs among subsamples. The effect of a deadline on pre-deadline settlement probability is positive (roughly ten percentage points) for smaller firms in the lowest three quintiles, but not for larger firms in the top two quintiles (figure 5a). The effects of including a deadline for firms in the second and the third quintiles are statistically indistinguishable from the first quintile, while effects for the fourth and fifth quintiles are significantly smaller (differences of ten and eleven percentage points, p-values 0.069 and 0.051).

Figure 5: Effect of a deadline on pre-deadline (22-day) settlement probability for different subsamples



Notes: This figure plots the marginal effects of a dummy that equals one if the case belongs to either of the two deadline message groups on the pre-deadline (22-day) settlement probability, holding the rest of the message constant. The baseline are cases belonging to either of the two non-deadline message groups. Refer to table A4 in the Online Appendix for the regression results. Q1 are the smallest firms and Q5 the largest.

Adding a deadline is significantly more effective when the potential future price increase is larger relative to the requested amount (figure 5b). For cases with the highest ratios, the pre-deadline settlement rate increases by more than ten percentage points (p-value of 0.0024) relative to the non-deadline messages; and the effect of a deadline is significantly different from that for cases with the lowest ratio (p-value of the difference is 0.002). To summarize, the addition of a deadline is not effective for the largest firms, and firms, on average, pay attention to the relative size of the potential price increase.[25]

[24] The regressions use cases from the four treatment groups of the message experiment. A 'deadline' indicator is used for cases from either of the two deadline message groups versus those from either of the two non-deadline groups. Refer to table A4 in the Online Appendix for the regression results. A linear probability model with the interaction terms produces similar results.

[25] An empirical regularity in the bargaining literature with a deadline is that the agreement is delayed until the very last minute before the deadline (e.g.,Roth et al. (1988)). Two types of strategic incentives may give rise to such a clustering-around-the-deadline phenomenon. One is that, by delaying until the last minute, bargainers try to get into the position of the proposer in an ultimatum game (e.g., Ma and Manove (1993)); and the other is that, with incomplete information, bargainers in a stronger position would want to delay to signal their type (e.g., Spier (1992)). In our data, however, we do not observe that settlement clusters around the 22nd day, possibly because, unlike sharp-termination games in lab experiments or a fixed trial date, the deadline in our setup is not a hard stop after which the opportunity of settlement disappears. Thus, the aforementioned strategic incentives may be too small to generate the clustering-around-the-deadline phenomenon.

Comparing the two non-deadline messages for different subsamples

The two different messages present the price reduction as either 'Waiving the Cost' intended to cover enforcement costs, or as a 'Discount' on the average settlement offer. As shown previously, for an average case, the two messages do not perform differently from each other in the first 30 days. Regression results for different subsamples also show that the 30-day outcomes of the two groups are statistically similar, regardless of firm size or the ratio between the forgiven and requested amounts.[26]

5 Discussion

5.1 Plausible explanations

These results show a sharp contrast between the small effect of reducing the settlement offer and the large effect from changing the communication messages, which is isolated due to experimental separation of the two changes. We propose two mechanisms that, taken together, help to explain our findings. First, we propose a simple rational model in which the infringer begins with limited information about what may happen if he does not settle immediately. These beliefs are exogenously influenced by the experiments' price reduction or messaging information. This simple model is sufficient to capture the average results of both experiments, and to potentially explain some of the results by different subsamples. The second mechanism examines the treatment messages' acknowledgment of possible unintentionality. This acknowledgment may change the mindset of some infringers away from treating the infringement as a legal matter. This, in turn, may encourage more collaborative outcomes, if the monetary stakes are not large.

Belief-updating mechanism

We first present a single-agent decision model in which the infringer has three choices: settle now, settle later, or do not settle. Consider the infringer's utility as:

$$U = \begin{cases} -p_1 - \varepsilon_1 & \text{if settle now} \\ -p_2 - \varepsilon_2 & \text{if settle later} \\ -C(p_2) & \text{if not settle} \end{cases}, \tag{1}$$

where p_1 is the payment if the infringer settles now; p_2 is what the infringer expects to pay for delayed settlement; and $C(p_2)$ is the expected cost from not settling.[27] Importantly, we assume that $C(p_2)$ is a

[26] Refer to table A5 in the Online Appendix for the regression results. We focus on the non-deadline messages because a deadline may interact differently with the different messages, and we do not have a good prior for the direction of these interaction effects. A comparison between the two deadline messages yields similar results for the 30-day outcome.

[27] For simplicity, we abstract from negotiation, which is not essential to illustrate the mechanism.

positive function of p_2 because the Agency is more likely to escalate higher-valued disputes.[28] Intuitively, the commercial value of the goods is positively correlated with price, and especially with the price associated with delayed settlement. ε_1 and ε_2 are independently distributed, mean zero exogenous shocks that affect the infringer's willingness to settle in the two periods (with cumulative distribution F and density f).

This defines the probability of settling now as

$$\Pr(\text{Settle now}) = \Pr(-p_1 - \varepsilon_1 \geq -p_2 - \varepsilon_2, -p_1 - \varepsilon_1 \geq -C(p_2)). \tag{2}$$

In the pricing experiment, there is only one price for infringers in each group. To obtain the expected treatment effect of a higher settlement offer, let $p_1 = p_2$, and take the derivative of Pr(Settle now) with respect to p_1; that is,

$$\frac{\partial \Pr(\text{Settle now})}{\partial p_1} = (1 - F(C(p_1) - p_1)) f(C(p_1) - p_1)(C'(p_1) - 1). \tag{3}$$

By this definition (equation (3)), a higher settlement offer generates two effects that respectively decrease and increase settlement probability. These effects offset each other and, thus, may explain why a substantial reduction in price generates only a small difference in the observed settlement rate. On one hand, a 'direct price' effect may increase the settlement probability when the settlement offer is lower, due to the disutility of paying (i.e., the demand for settling is downward-sloping). On the other hand, a higher price may suggest a greater 'threat' of escalating enforcement (that is, $C'(p_1) > 0$), discouraging settlement when the requested amount is lower.

The message experiment can be represented in this model by assuming that all four treatment groups have the same p_1 as the control group, and that knowledge of the Agency's $400 higher recoupable revenue (informed by the treatment messages) can be encoded as a higher p_2.[29] We can then derive the effect of the messages by taking the derivative of Pr(Settle now) with respect to p_2:

$$\begin{aligned} \frac{\partial \Pr(\text{Settle now})}{\partial p_2} &= \int_{-\infty}^{C(p_2)-p_1} f(-p_2 + p_1 + \varepsilon_1) f(\varepsilon_1) d\varepsilon_1 \\ &\quad + (1 - F(C(p_2) - p_2)) f(C(p_2) - p_1) C'(p_2). \end{aligned} \tag{4}$$

Both terms in equation (4) are positive. Thus, in contrast to the pricing experiment, the direct price effect and

[28] In patent litigation, Lanjouw and Schankerman (2001) show that the likelihood of litigation is higher when the stakes are higher, either because the commercial value of the patent is higher or because there is a greater incentive to establish a tougher reputation in a particular technology area.

[29] Findings in the marketing literature show that placing a sale sign on an item is sufficient to increase demand for the item even without changing the actual price because consumers expect the prices to go up after the promotion period (e.g., Inman et al. (1990) and Anderson and Simester (1998)).

the threat effect in the message experiment are synergistic (rather than oppositional) in making 'settling now' a more attractive option. Relative to the messages without a deadline, the two messages with a deadline yield a stronger effect by making both the inter-temporal price trade-off and the increased threat more explicit.

In addition to capturing average-case results, the model may also help to explain why the pure effect of including a deadline is greater when the infringing firms are smaller, and when the forgiven amount is larger relative to the requested amount (see figure 5). If larger firms are more price-inelastic than smaller firms, or do not change their beliefs about the potential risk of delaying payment, the difference resulted from including an explicit deadline will be small (as observed).[30] The monotonicity of the effect of a deadline with respect to the relative size of p_1 and p_2 is also intuitive, because the larger is the explicit increase in p_2, the greater is the price inducement to settle now, and the greater the potential threat. Recall that comparing the two non-deadline messages to the control group, however, shows that the effect is relatively flat across different subsamples (figure 4). This is consistent with (i) the price tradeoff implied by these non-deadline messages is non-explicit; and (ii) compared to the control group, these messages also contain the acknowledgement of possible unintentionality.

Collaborative mindset mechanism

The prevalence of unintentional infringement makes it plausible that the message's acknowledgement of possible unintentionality plays an important role in encouraging the infringers to collaborate. Table 9 summarizes the (self-reported) reasons for infringement, which are collected from correspondence records between the Agency and firms that made contact during the message experiment. The infringer claims to be unaware of any infringement (or to have obtained the image online thinking it free to use) in 43 percent of the cases. In 42 percent of the cases, an outside party is involved in designing the website.[31] In some cases, the firm uses a previously-licensed image outside the specified scope (4%) or uses an image without a required watermark or as a placeholder for the website (7%). Recognizing that the incentive to claim unawareness is strong, the extent of unintentional infringement still seems substantial. Table 9 also shows that this phenomenon is similar across infringers of all sizes. Thus, firms may feel confused and frustrated when faced with enforcement because they were not aware of the mistake.

At least two different sub-mechanisms may work to shift the infringer's frame of mind away from the calculation of legal risk as a result of the acknowledgement of possible unintentionality. First, this

[30] For example, correspondence data from the Agency show that the larger is the infringing firm, the more likely it is that the person who contacts the Agency is an internal lawyer.

[31] We define third-party involvement to include cases for which either: (1) the party who contacts the Agency is a third-party designer, family member or friend (17%) or (2) the end-user claims that a third party designed the website (25%). It is possible that some family or friends make contact because they have a legal background. However, we group these calls with third-party designers because anecdotal evidence from the Agency suggests that friends and family who contact the Agency are usually involved with the website design.

Table 9: (Self-reported) reason(s) for infringement

Infringer-size quintile	N	Third-party involved	Unaware/obtain from the Internet	Improper use of license	Other reasons	Deny infringement
1	154	0.41	0.47	0.01	0.06	0.05
2	168	0.42	0.43	0.01	0.05	0.08
3	181	0.43	0.44	0.02	0.07	0.04
4	143	0.46	0.40	0.03	0.06	0.05
5	176	0.37	0.45	0.09	0.07	0.02
Total	909	0.42	0.43	0.04	0.07	0.05
Corr. with log(sales)		-0.04	0.02	0.17	-0.02	-0.06
p-value		(0.24)	(0.67)	0.00	(0.65)	(0.07)

Notes: The data are taken from correspondence records between the Agency and firms in the message-experiment sample that contact the Agency. We include infringers that make contact at *any* time during the sample period, not just in the first 30 days. There is no significant difference in annual sales, age or the number of employees between infringers that make contact (about half of the cases) and those that either ignore the letters or settle without making contact. Twenty percent of the cases for which firms make contact do not contain enough information and, thus, are not used in the tabulation. Size quintiles are based on annual sales, with Q1 being the smallest firms and Q5 the largest.

acknowledgement may uphold the infringers' intrinsic motivation to pay, which is especially effective when the amount of the claim is small. The literature on tax compliance has similar findings. For example, Feld and Frey (2002a,b) argue that paying tax is quasi-voluntary and show that courteous treatment (i.e., a presumption of innocence) dampens tax evasion. Second, a less behavioral mechanism may exist in which the acknowledgement of unintentionality makes the settlement request letter conform more closely to other standard invoices. When firms have routines in place for paying bills or resolving disputes, this could provide a smoother path for processing payment instead of invoking a costly internal legal review.

5.2 Longer-term (120-day) outcomes

As explained previously, outcomes beyond 30 days may be influenced by follow-up letters sent to different groups of infringers by the Agency. The difference in the follow-up intensity is relatively small between the two price groups in the pricing experiment, and, thus, we evaluate the 120-day outcomes of the pricing experiment using the actual data without further adjustment. The results are consistent with the 30-day results. Removing the $400 per-image charge results in a seven-percentage-point increase in the 120-day settlement probability, but the increase in the settlement probability is not enough to outweigh the decrease in revenue conditional on settlement ($211). As a result, the expected revenue per case is statistically similar between the two groups (the lower-priced group generates $11 less, but the p-value of the difference is 0.49).

The confounding problem of the Agency's follow-up letters, however, is likely to be severe for the message experiment because substantially more letters are sent for cases in the control group than in the treatment groups. For example, 42 percent of the control group received a FedEx letter after about 30 days,

while only 15 percent of the treatment groups did.[32] The settlement rate of the control group does catch up after 120 days. The timing of the start of the catchup (there is a discrete change in the slope of the cumulative settlement probability for the control group but not for the treatment groups) is after about 40 days, consistent with the firms responding to follow-up letters.

Given the available information, we bound the 120-day effect of the messages, using the actual difference as the lower bound and the 30-day difference as the upper bound, assuming that the short-run effect is persistent. The results (reported in table A8b in the Online Appendix) show that the two messages that do not impose a deadline perform similarly to the control group after 120 days, while the two messages that impose a deadline yield a settlement rate that is 3.4 percentage points higher than the control (p-value is 0.20). Combining these actual differences with the 30-day estimates, the 120-day effect of the non-deadline messages is estimated to be between 0 and ten percentage points, representing a 0-25 percent increase relative to the control group. The 120-day effect of the two deadline messages is bounded between 3.4 and 15 percentage points (an increase of 8.5-37.4 percent relative to the control).

5.3 Effects on legal costs

Apart from increasing revenue for photographers, a higher settlement rate may also reduce the social costs of resolving disputes. We provide a back-of-the-envelope calculation of the difference in legal costs associated with escalating a case to an external law firm, which is attributable to the treatment messages. Very few cases end up in court, and we do not consider the potential savings from reduced litigation. In the message experiment, over a 120-day horizon, 50.5 percent of the cases in the control group and 45.8 percent in the treatment groups (pooled together) are sent to external law firms, and the difference is statistically significant (p-value is 0.052). Similar to the 120-day outcomes, the observed difference in the probability of using external law firms is likely to substantially under-estimate the effect of the messages because significantly more follow-up letters are sent to the control group. We use the actual difference of 4.6 percentage points as a lower-bound estimate and the raw-data difference in the 30-day settlement rate (12.4 percentage points) as an upper-bound estimate. Thus, the messages reduce the probability of escalating to external enforcement by roughly 9.1-25 percent.

The main law firm with which the Agency contracts in the U.S. uses the following fee structure (reported with all details disguised for confidentiality). The Agency pays a fixed fee ($\$a$) for each case sent. The law firm typically adds an attorney fee of $\$c$ in the requested amount from infringers and deducts this amount

[32] In addition, table A9 in the Online Appendix reports that 82 percent of the control group received a standard second letter (also about 30 days after the first letter), while 64-72 percent in the treatment groups did. Finally, 14 percent of the control group received an additional email reminder (about two months after the first letter), while none of the cases in the four treatment groups did.

from the collected revenue for a settled case. Furthermore, the Agency also pays a contingency fee of b percent of the remaining collected revenue for a settled case.[33] Using this fee structure to approximate legal costs, we first calculate a settlement rate for cases sent to external law firms (d percent) based on historical data, consisting of infringement cases in 2011-2013, provided to us by the Agency.[34] Next, we assume that the average revenue (after deducting the attorney fee \$$c$) for settled case is the same as for cases settled within 30 days of the message experiment (\$505). Finally, we assume that infringers hire legal counsels for twenty percent of the cases and that their attorney fee is also \$$c$ per case.[35]

Table 10: A rough calculation of savings in legal costs

	Lower bound	Upper bound
Panel 1: Assumptions		
Number of cases	1,000	1,000
Difference in the number of cases sent to external law firms between the control and the treatment groups	46	124
Contractual terms by the Agency's law firm:		
Fixed fee per case sent	\$$a$	\$$a$
Attorney fee deducted from the collected revenue of a settled case	\$$c$	\$$c$
Contingency fee based on remaining revenue of a settled case (after deducting \$$c$)	b%	b%
Settlement rate for cases sent to external law firms	d%	d%
Average revenue per settled case after deducting the attorney fee \$$c$	505	505
Percentage of infringers obtaining legal service	0.2	0.2
Attorney fee per case charged by the infringers' lawyers	\$$c$	\$$c$
Panel 2: Results		
Savings for the Agency (fixed and contingency fees)	\$866.5	\$2,335
Savings for the infringing firms	\$4,347	\$11,718
Total savings in legal costs	\$5,214	\$14,053

Notes: This table presents a calculation of savings in the legal costs, relative to the control group, due to the treatment messages for the population of cases in the message experiment. Take the upper-bound estimate for example. The savings for the Agency consist of (1) fixed fees ($124 \times \$a$), and (2) contingency fees ($124 \times d\% \times \$505 \times b\%$). The savings for the infringing firms consist of (1) the attorney fees for firms that employ legal counsel ($124 \times 20\% \times \$c$) and (2) attorney fees requested by the Agency's law firm for firms that settle ($124 \times d\% \times \$c$).

For each 1,000 cases in the U.S. and Canada, the savings in legal costs are approximately \$866-\$2,335 for the Agency and \$4,347-\$11,718 for infringing firms (table 10). Total in-house enforcement revenue (revenue obtained without the use of external law firms) per thousand cases for the control group in the message experiment is \$192,000. Thus, savings in total legal costs are between 2.7 and 7.3 percent of the Agency's in-house enforcement revenue. This may, in fact, underestimate the total savings in social costs due to additional factors, such as the opportunity cost of time and the psychological costs of stress for both

[33] We observe the actual numbers and use them for all calculations in table 10, but are not allowed to report them separately.

[34] We use historical data for these statistics because enforcement through external law firms typically takes a long time, and the time frame of our experimental sample is not long enough to capture all associated outcomes.

[35] We do not have information about the percentage of infringers that employ legal counsel after the Agency moves a case to an external law firm. The correspondence records between the Agency and firms that made contact show that, at the in-house enforcement stage, infringing firms are already employing external legal counsel eight percent of the time. Once the Agency escalates the case to outside law firms, we expect the percentage to be higher.

parties from being entangled in prolonged disputes.

6 Conclusion

We study the effects of different enforcement methods on settlement outcomes for instances of copyright infringement by businesses. We use a novel, proprietary dataset generated from two large-scale field experiments run by a leading stock-photography agency. The experimental design avoids the confounding effects often encountered during real price changes, and we find that, on average, a substantial reduction in the requested amount alone generates only a small increase in the settlement probability. In contrast, given the same lower requested amount, the addition of a message that informs infringers of the price reduction and acknowledges the possibility of unintentional infringement has a large positive effect on settlement. In addition, including a deadline further improves the settlement rate. Beyond generating higher revenues for copyright owners, the higher settlement rate resulting from the messages also substantially reduces the social costs of dispute resolution.

Our results suggest that, when disputes are likely to be the result of inadvertent mistakes, proactive acknowledgment of this possibility may encourage collaborative outcomes (at least when the monetary stakes are small). This result may be generalized to other similar contexts. Given the representativeness of the firms in our data, the effects that we document for digital images are likely to generalize to other types of copyrighted materials used by businesses, such as videos, music, and designs. For consumer piracy, many contexts studied by the existing literature typically involve willful infringement (e.g., repeat file-sharers of music and movies). There are, however, areas in which potential infringers are unfamiliar with their legal obligations. An interesting example is the increasing use of 3D printing: many people download designs from the Internet without being aware of its illegality (Depoorter (2014)). Beyond copyright, inadvertent mistakes due to forgetfulness, disorganization, and limited awareness could also result in conflicts in a wide variety of contexts that may be mitigable using such strategies.

More generally, our results also highlight the fact that communication can have an important impact on economic outcomes in some situations. Existing evidence from the lab and field are primarily based on individuals (see, e.g., Roth and Murnighan (1982) on the role of information in bargaining games; DellaVigna and Gentzkow (2010) for a survey of studies on persuasive communication in a variety of contexts; and Thaler and Sunstein (2008)). Our study shows that these interventions can also influence firms. In this sense, it provides a small but important step towards deepening our understanding of these types of mechanisms beyond individual consumers in the broader economy.

References

Anderson, E. and D. Simester (1998). The role of sale signs. *Marketing Science 17*, 139–155.

Babcock, L., G. Loewenstein, S. Issacharoff, and C. Camerer (1995). Biased judgments of fairness in bargaining. *American Economic Review 85*, 1337–1343.

Bebchuk, L. (1984). Litigation and settlement under imperfect information. *RAND Journal of Economics 15*, 404–415.

Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir, and J. Zinman (2010). What's advertising content worth? evidence from a consumer credit marketing field experiment. *Quarterly Journal of Economics 125*, 263–306.

Bessen, J. E., J. Ford, and M. J. Meurer (2011). The private and social costs of patent trolls. *Regulation 34*, 26–35.

Bessen, J. E. and M. J. Meurer (2006, May). Patent Litigation with Endogenous Disputes. *American Economic Review 96*(2), 77–81.

Bhattacharjee, S., R. D. Gopal, K. Lertwachara, and J. Marsden (2006). Impact of legal threats on individual behavior: An analysis of music industry actions and online music sharing. *Journal of Law and Economics 49*, 91–114.

Cadena, X. and A. Schoar (2011). Remembering to pay? reminders vs. financial incentives for loan payments. Working paper.

Cockburn, I. and M. MacGarvie (2009). Patents, thickets and the financing of early-stage firms: Evidence from the software industry. *Journal of Economics and Management Strategy 18*, 729–773.

Cohen, L., U. Gurun, and S. D. Kominers (2014, August). Patent trolls: Evidence from targeted firms. HBS Working Paper 15-002.

Cooter, R. and D. Rubinfeld (1989). Economic analysis of legal disputes and their resolution. *Journal of Economic Literature 27*, 1067–1097.

Cooter, R. D., S. P. Marks, and R. Mnookin (1982). Bargaining in the shadow of the law: A testable model of strategic behavior. *Journal of Legal Studies 11*, 225–251.

Danaher, B., M. Smith, R. Telang, and S. Chen (2014). The effect of graduated response anti-piracy laws on music sales: Evidence from a natural experiment in france. forthcoming in Journal of Industrial Economics.

Danzon, P. M. and L. A. Lillard (1983). Settlement out of court: The disposition of medical malpractice claims. *Journal of Legal Studies 12*, 345–377.

DeBriyn, J. (2012). Shedding light on copyright trolls: An analysis of mass copyright litigation in the age of statutory damages. *UCLA Entertainment Law Review 19*.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: Empirical evidence. *Annual Review of Economics 2*, 643–669.

Depoorter, B. (2014). Intellectual property infringements & 3d printing: Decentralized piracy. *HASTINGS LAW JOURNAL 65*, 1483–1503.

Eisenberg, T. and H. S. Farber (1997). The litigious plaintiff hypothesis: Case selection and resolution. *RAND Journal of Economics 28*, 92–112.

Falk, A. (2007). Gift exchange in the field. *Econometrica 75*, 1501–1511.

Feld, L. P. and B. S. Frey (2002a). The tax authority and the taxpayer: An explotary analysis.

Feld, L. P. and B. S. Frey (2002b). Trust breeds trust: How taxpayers are treated. *Economics of Governance 3*, 87–99.

Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association 11*, 634–660.

Fournier, G. M. and T. W. Zuehlke (1996). The timing of out-of-court settlements. *Rand Journal of Economics 27*, 310–321.

Frey, B. S. and S. Meier (2004). Social comparison and pro-social behavior: Testing 'conditional cooperation' in a field experiments. *American Economic Review 94*, 1717–1722.

Galasso, A., M. Schankerman, and C. Serrano (2013). Trading and enforcing patent rights. *RAND Journal of Economics 44*, 275–312.

Glückler, J. and R. Panitz (2013). Survey of the global stock image market 2012. part i: Players, products, business. Technical report, Heidelberg: GSIM Research Group.

Güth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization 3*, 367–388.

Hall, B. and R. Ziedonis (2001). The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995. *RAND Journal of Economics 32*, 101–128.

Hallsworth, M., J. List, R. Metcalfe, and I. Vlaev (2014, March). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Working Paper 20007, National Bureau of Economic Research.

Inman, J. J., L. McAlister, and W. D. Hoyer (1990). Promotion signal: proxy for a price cut? *Journal of Consumer Research 17*, 74–81.

Jolls, C., C. R. Sunstein, , and R. Thaler (1998). A behavioral approach to law and economics. *Stanford Law Review 50*, 1471–1550.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review 76*, 728–741.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*, 263–291.

Karlan, D., M. Morten, and J. Zinman (2015). A personal touch: Text messaging for loan repayment. NBER Working Paper No. 17952.

Korobkin, R. and C. Guthrie (1994). Psychological barriers to litigation settlement: An experimental approach. *Michigan Law Review 93*, 107–192.

Lanjouw, J. and M. Schankerman (2001). Characteristics of patent litigation: A window on competition. *RAND Journal of Economics 32*, 129–151.

Lanjouw, J. O. and J. Lerner (1998). The enforcement of intellectual property rights: A survey of the empirical literature. *Annales d'conomie et de Statistique* (49/50), 223–246.

Lanjouw, J. O. and M. Schankerman (2004). Protecting intellectual property rights: are small firms handicapped? *Journal of Law and Economics 47*, 45–74.

Lantagne, S. (2004). The morality of mp3s: The failure of the recording industry's plan of attack. *Harvard Journal of Law & Technology 18*, 269–293.

Loewenstein, G., S. Issacharoff, C. Camerer, and L. Babcock (1993). Self-serving assessments of fairness and pretrial bargaining. *The Journal of Legal Studies 22*, 135–159.

Ma, C. and M. Manove (1993). Bargaining with deadlines and imperfect player control. *Econometrica 61*, 13131339.

Meurer, M. (1989). The settlement of patent litigation. *RAND Journal of Economics 20*, 77–91.

Perez-Truglia, R. and U. Troiano (2015, March). Tax debt enforcement: Theory and evidence from a field experiment in the united states.

Peukert, C., J. Claussen, and T. Kretscher (2015). Piracy and box office movie revenues: Evidence from megaupload.

P'ng, I. (1983). Strategic behavior in suit, settlement, and trial. *RAND Journal of Economics 14*, 539550.

Priest, G. and B. Klein (1984). The selection of disputes for litigation. *Journal of Legal Studies 13*, 1–55.

Rachlinski, J. J. (1996). Gains, losses, and the psychology of litigation. *Southern California Law Review 70*, 113–185.

Reimers, I. (2015, February). Can private copyright protection be effective? evidence from book publishing.

Roth, A., K. Murnighan, and F. Schoumaker (1988). The deadline effect in bargaining: Some experimental evidence. *American Economic Review 78*, 806–823.

Roth, A. E. and J. K. Murnighan (1982, September). The Role of Information in Bargaining: An Experimental Study. *Econometrica 50*(5), 1123–42.

Scott Morton, F. and C. Shapiro (2014). Strategic patent acquisitions. *Antitrust Law Journal 79*, 463–499.

Spier, K. E. (1992). The dynamics of pretrial negotiations. *Review of Economic Studies 59*, 93–108.

Spier, K. E. (2007). Litigation. In A. M. Polinsky and S. Shavell (Eds.), *Handbook of Law and Economics*, Chapter 4. North-Holland.

Thaler, R. H. and C. R. Sunstein (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven & London: Yale University Press.

Trubek, D. M., A. Sarat, W. L. Felstiner, H. M. Kritzer, and J. B. Grossman (1983). The costs of ordinary litigation. *UCLA Law Review 31*, 72–127.

Viscusi, W. K. (1988). Product liability litigation with risk aversion. *Journal of Legal Studies 17*, 101–121.

Waldfogel, J. (1995). The selection hypothesis and the relationship between trial and plaintiff victory. *Journal of Political Economy 103*, 229–260.

Waldfogel, J. (2012). Copyright protection, technological change, and the quality of new products: Evidence from recorded music since napster. *Journal of Law & Economics 55*, 715–40.

Online Appendix

Appendix A. Additional Tables and Figures

Table A1: Regression results on the settlement probability with and without using controls

	Logit model: DV = Settle			
	Pricing experiment		Message experiment	
	(1)	(2)	(3)	(4)
High Price dummy	-0.03*	-0.03*		
	(0.02)	(0.02)		
Waive the Cost			0.12***	0.11***
			(0.03)	(0.03)
Waive the Cost + Deadline			0.18***	0.16***
			(0.03)	(0.03)
Discount			0.13***	0.11***
			(0.03)	(0.03)
Discount + Deadline			0.19***	0.17***
			(0.02)	(0.02)
Control variables	Y	N	Y	N
Adj R-squared	0.06	0.002	0.06	0.02
N	1983	1983	2291	2295

Notes: This table reports the logit regression results on the 30-day settlement probability without using any controls. For comparison, the regression results using controls are replicated in the first and the third columns (that is, the results reported in column (1) in tables 6 and 8). All models report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant at levels of 1%, 5%, and 10%.

Table A2: Price effects on settlement outcomes for different subsamples

Dependent variable	Settle		Revenue for settled cases		Revenue	
	Logit		OLS		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
High Price dummy	-0.02	-0.00	155.77	324.20***	8.77	19.63
	(0.04)	(0.04)	(98.81)	(52.12)	(14.34)	(27.13)
Size quintile 2 × High Price	-0.00		30.39		3.11	
	(0.06)		(111.78)		(28.35)	
Size quintile 3 × High Price	-0.02		188.53		11.08	
	(0.06)		(115.43)		(26.16)	
Size quintile 4 × High Price	0.00		66.22		16.55	
	(0.04)		(115.06)		(22.53)	
Size quintile 5 × High Price	-0.02		170.03		-0.80	
	(0.05)		(106.42)		(45.26)	
Size quintile 2	0.01		80.13		13.33	
	(0.06)		(91.75)		(31.68)	
Size quintile 3	0.04		7.06		8.78	
	(0.06)		(83.24)		(24.63)	
Size quintile 4	0.04		108.86		41.95	
	(0.05)		(90.01)		(32.48)	
Size quintile 5	0.06		130.29		92.71	
	(0.08)		(123.70)		(60.51)	
Medium price-change % × High Price		-0.03		-89.41		-11.95
		(0.04)		(63.42)		(29.47)
High price-change % × High Price		-0.04		-92.44		-0.66
		(0.04)		(63.01)		(32.65)
Medium price-change %		0.05		-104.84		-1.90
		(0.05)		(164.30)		(61.53)
High price-change %		0.13*		-27.85		57.20
		(0.07)		(245.03)		(102.86)
Control variables	Y	Y	Y	Y	Y	Y
Adj R-squared			0.653	0.646	0.030	0.031
N	1983	1983	269	269	1983	1983

Notes: This table reports the marginal effects of three sets of regressions using the observations in the pricing experiment, in which the dependent variables are (1) whether the case settles in 30 days; (2) the revenue for settled cases; and (3) the revenue of a case (coded as $0 if not settled in 30 days). In all regressions, the independent variable is a dummy indicating the High Price group (the Low Price group is the default). Size quintiles are based on annual sales, with quintile 1 including the smallest firms. Price-change percentage is defined by ($400*number of images)/licensing fee. The controls in all regressions include log(Baseline price), image count, log(Annual sales), log(Age+1), a missing-age dummy, and dummies indicating the top ten industries (defined by the two-digit SIC codes), the top ten states (provinces), a U.S. location, and the mail week. All models report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant at levels of 1%, 5%, and 10%.

Table A3: Comparing the two non-deadline messages and the control for different subsamples

	Logit model: DV = Settle		
	(1)	(2)	(3)
Non-deadline message	0.10***	0.07	0.11***
	(0.02)	(0.05)	(0.04)
Size quintile 2 × Non-deadline message		0.06	
		(0.09)	
Size quintile 3 × Non-deadline message		-0.08	
		(0.08)	
Size quintile 4 × Non-deadline message		0.00	
		(0.07)	
Size quintile 5 × Non-deadline message		0.14***	
		(0.05)	
Size quintile 2		-0.04	
		(0.06)	
Size quintile 3		0.08	
		(0.05)	
Size quintile 4		0.05	
		(0.05)	
Size quintile 5		0.00	
		(0.07)	
Medium ratio × Non-deadline message			0.02
			(0.05)
High ratio × Non-deadline message			-0.04
			(0.04)
Medium ratio			-0.03
			(0.06)
High ratio			0.01
			(0.08)
Control variables	Y	Y	Y
N	1420	1420	1420

Notes: This table reports logit regression results (marginal effects) using cases in the two non-deadline messages and the control groups in the message experiment. The dependent variable is whether the case settles in 30 days. Non-deadline message equals one if the case belongs to either of the two non-deadline message groups (the default is the control group). Size quintiles are based on annual sales, with quintile 1 including the smallest firms. Medium ratio equals one if $0.77 < \frac{\text{forgiven amount}}{\text{requested amount}} < 0.98$; and high ratio equals one if the ratio is higher than 0.98. In all regressions, the control variables are log(Requested amount), image count, log(Annual sales), log(Age+1), missing-age dummy, dummies indicating the top ten industries (defined by the two-digit SIC codes), the top ten states (provinces), the U.S., and the mail date. All models report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant at levels of 1%, 5%, and 10%.

Table A4: Comparing messages with and without a deadline for different subsamples

	Logit model: DV = Settle (in 22 days)		
	(1)	(2)	(3)
Deadline message	0.05***	0.11**	0.01
	(0.02)	(0.05)	(0.03)
Size quintile 2 × Deadline message		-0.04	
		(0.06)	
Size quintile 3 × Deadline message		0.01	
		(0.06)	
Size quintile 4 × Deadline message		-0.10*	
		(0.06)	
Size quintile 5 × Deadline message		-0.11*	
		(0.06)	
Size quintile 2		0.08	
		(0.06)	
Size quintile 3		0.05	
		(0.05)	
Size quintile 4		0.13**	
		(0.05)	
Size quintile 5		0.20***	
		(0.07)	
Medium ratio × Deadline message			0.03
			(0.05)
High ratio × Deadline message			0.10***
			(0.03)
Medium ratio			0.02
			(0.05)
High ratio			0.01
			(0.05)
Control variables	Y	Y	Y
N	1696	1696	1696

Notes: This table reports logit regression results (marginal effects) using cases in all four treatment groups in the message experiment. The dependent variable is whether the case settles in the first 22 days. Deadline message equals one if the case belongs to either of the two deadline message groups (the default group are cases that belong to either of the two non-deadline messages). Size quintiles are based on annual sales, with quintile 1 including the smallest firms. Medium ratio equals one if $0.77 < \frac{\text{forgiven amount}}{\text{requested amount}} < 0.98$; and high ratio equals one if the ratio is higher than 0.98. In all regressions, the control variables are a dummy indicating the discount framing, log(Requested amount), image count, log(Annual sales), log(Age+1), missing-age dummy, dummies indicating the top ten industries (defined by the two-digit SIC codes), the top ten states (provinces), the U.S., and the mail date. All models report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant at levels of 1%, 5%, and 10%.

Table A5: Comparing the two non-deadline messages for different subsamples

	Logit model: DV = Settle		
	(1)	(2)	(3)
Discount	0.02	0.08	0.02
	(0.04)	(0.11)	(0.05)
Size quintile 2 × Discount		-0.10	
		(0.08)	
Size quintile 3 × Discount		-0.04	
		(0.13)	
Size quintile 4 × Discount		-0.10	
		(0.13)	
Size quintile 5 × Discount		-0.06	
		(0.13)	
Size quintile 2		0.08	
		(0.08)	
Size quintile 3		0.02	
		(0.08)	
Size quintile 4		0.13	
		(0.11)	
Size quintile 5		0.24**	
		(0.11)	
Medium ratio × Discount			0.00
			(0.06)
High ratio × Discount			-0.03
			(0.05)
Medium ratio			-0.03
			(0.07)
High ratio			-0.05
			(0.10)
Control variables	Y	Y	Y
N	825	825	825

Notes: This table reports logit regression results (marginal effects) using cases in the two non-deadline message groups in the message experiment. For all regressions, the dependent variable is whether the case settles in 30 days. Discount equals one if the case belongs to the Discount message group. Size quintiles are based on annual sales, with quintile 1 (i.e., Q1) including the smallest firms. Medium ratio equals one if $0.77 < \frac{\text{forgiven amount}}{\text{requested amount}} < 0.98$; and high ratio equals one if the ratio is higher than 0.98. In all regressions, the control variables are a dummy indicating the discount framing, log(Requested amount), image count, log(Annual sales), log(Age+1), missing-age dummy, dummies indicating the top ten industries (defined by the two-digit SIC codes), top ten states (provinces), and the U.S. All regressions report robust standard errors, clustered by industry. ***, **, and * are, respectively, significant at levels of 1%, 5%, and 10%.

Table A6: Balance check for the 'unfiltered' sample

(a) Pricing experiment

Group	N	Baseline price	Sales				Age			
				Sales, $1,000		log(Sales, $)		Age		log(Age+ 1)
			N	mean	median	mean	N	mean	median	mean
Low Price	945	662.3	896	25749	330	13.1	910	19.5	11	2.5
High Price	1,800	658.6	1,649	13020	290	13.0	1,593	18.1	11	2.5
		(0.85)		(0.20)	(0.41)	(0.15)		(0.16)	(0.71)	(0.70)

(b) Message experiment

Group	N	Baseline price	Sales				Age			
		(= Requested		Sales, $1,000		log(Sales, $)		Age		log(Age+ 1)
(p-value, diff. to control)		amount)	N	mean	median	mean	N	mean	median	mean
Control	866	639.3	754	25420	383	13.3	756	19.2	11	2.6
Waive the Cost	549	675.3	458	44562	348	13.4	455	19.5	11	2.6
		(0.15)		(0.32)	(0.64)	(0.86)		(0.82)	(0.88)	(0.91)
Waive the Cost + Deadline	564	612.0	493	11023	300	13.1	494	16.9	9	2.5
		(0.24)		(0.28)	(0.13)	(0.07)		(0.08)	(0.16)	(0.09)
Discount	571	609.7	526	10071	320	13.2	528	18.2	12	2.5
		(0.19)		(0.23)	(0.26)	(0.21)		(0.42)	(0.80)	(0.75)
Discount + Deadline	577	652.6	526	48841	378	13.1	533	17.6	10	2.5
		(0.79)		(0.55)	(0.89)	(0.01)		(0.19)	(0.30)	(0.32)

Notes: This table reports the balance tests using the 'unfiltered' sample (that is, without dropping cases for which the mail is returned or sales information is missing, etc). Note, however, tier-4 and tier-5 firms in the pricing experiment are excluded to ensure that firms in the two experiments are comparable in size. The test statistics are consistent with what we report in the paper after dropping some types of cases.

Table A7: 30-day settlement outcomes for the 'unfiltered' sample

(a) Pricing experiment

	Low Price	High Price	Difference	p-value
Settlement probability	0.12	0.10	-0.02	(0.06)
Revenue for settled cases	528.09	1080.1	552.1	(0.02)
Revenue per case	65.94	109.8	43.9	(0.12)
N	945	1,800		

(b) Message experiment

Group	N	Settlement probability	Difference (compared to control)	p-value
Control	866	0.12		
Waive the Cost	549	0.19	0.07	(0.00)
Waive the Cost + Deadline	564	0.25	0.13	(0.00)
Discount	571	0.21	0.09	(0.00)
Discount + Deadline	577	0.27	0.15	(0.00)

Notes: This table reports the settlement outcomes for different groups in the two experiments using the 'unfiltered' sample (defined in table A6). The results are consistent with what we report in the paper after dropping some types of cases.

Table A8: Price and message effects on the 120-day outcomes

(a) Pricing experiment

	Low Price group	High Price group	Difference	(p-value)
Settlement probability	0.32	0.25	-0.07	(0.00)
Revenue for settled cases	530.9	741.5	210.6	(0.00)
Revenue per case	167.9	178.9	11.0	(0.49)

(b) Message experiment

(p-value is relative to the control)	120-day settlement probability	120-day difference (lower bound)		30-day difference (upper bound)	
Control	0.401				
Non-deadline messages	0.393	-0.007	(0.76)	0.10	(0.00)
Deadline messages	0.435	0.034	(0.20)	0.15	(0.00)

Notes: This table reports the 120-day settlement outcomes for different groups in both experiments. For the message experiment, we report only the effects on the settlement probability because the requested amount is similar across groups given the same image use, and, thus, the effects on the expected revenue per case are consistent with those on the settlement probability. In the message experiment, analyses of long-term outcomes can be substantially confounded by systematically more follow-up letters for the control group. For example, as reported in table A9, 42 percent of the control group received a FedEx letter after around 30 days, while only 15 percent of the treatment groups did. 82 percent of the control group also received a standard second letter, while 64-72 percent in treatment groups did. Finally, 14 percent of the control group received an additional email reminder, while none of the cases in the four treatment groups did. As reported in this table, the control group in the message experiment does catch up after 120 days. The timing of the start of the catchup (there is a discrete change in the slope in the cumulative settlement probability for the control group but not for the other groups) is after about 40 days, consistent with the firms responding to follow-up letters.

Appendix B: The stock photography industry

The stock photography industry provides images that are readily available for licensing. On behalf of individual photographers or organizations (such as National Geographic), stock photography agencies manage, market and license images to business customers, such as advertising and graphic design agencies, publishers and corporations.

Rights-managed (RM) and royalty-free (RF) are two typical license types. A RM license allows the licensee a one-time use of the image (e.g., 100 billboards in Germany for three months). For a given image, the licensing fee depends on the specified use. A licensee who wants to use the image beyond the specified scope must purchase a new license. In contrast, an RF license gives the licensee the right to use the image with few restrictions after making a one-time payment. Under the RF model, some agencies price images on an à la carte basis, while others use subscription pricing. According to a survey by Graphic Design USA, 32, 98, and 54 percent of responding designers had licensed images under the RM, à la carte RF, and subscription RF models, respectively, in 2011.

Stock images are categorized as either editorial or creative, depending on whether commercial use is allowed. Editorial images portray specific people, places, things and events that provide context for newspaper and magazine articles, blog posts, and other non-commercial presentations. Editorial images cannot be licensed for advertising or promotional purposes because there are typically no permissions from the subjects of these photos. In contrast, creative images are produced and staged by the photographer. Creative images can be licensed for both non-commercial and commercial purposes because they feature generic objects, or subjects who have consented to commercial use.

The industry is divided into traditional (premium) and micro-stock segments. Leading agencies in the traditional segment include Corbis Images and Getty Images; leading agencies in the micro-stock segment are Fotolia, iStockphoto and Shutterstock. Suppliers of the traditional segment are typically professional photographers who target customers and projects that require high-quality images. Images in this segment are licensed under the RM or the à-la-carte RF model. Depending on the quality of the image and the specified use, the licensing fee ranges from a couple of hundred to tens of thousands of dollars.[36] In addition to facilitating transactions, agencies in this segment also provide complementary services, including searching for images with specific themes, large-project management, rights clearance and protection against copyright liabilities.

Agencies in the micro-stock segment source their images from a wide range of photographers, including professionals and hobbyists. Images in this segment are licensed under the RF model. The price level is significantly lower than in the traditional segment. For example, the vast majority of images on iStockphoto (à la carte pricing) are priced between $2 and $100 per image, depending on the size of the digital file; and, at Shutterstock (subscription pricing), the user can download up to 25 images per day for a monthly fee of $249.

[36] The most expensive cases usually occur because the licensed use is exclusive.

Appendix C. Follow-up letters

The Agency sends three types of follow-up letters. First, a second settlement request letter is sent after about 30 days unless the case has been settled by then. Second, also after 30 days or so, an extra FedEx letter is sent to a set of relatively large firms that have not yet responded. (The first and the second letters are sent by USPS standard mail). Third, after 60 days, the Agency sends a final notice by email if the firm has ignored the previous letters. Cases that do not settle (either because the firms ignore all the letters or do not settle after making contact) are eventually sent to external law firms. In our sample, about half of all cases are sent to external law firms, with 90 percent sent between 58 and 159 days after the initial letter.

The proportion of cases receiving the second letter is determined by the initial settlement outcomes of a particular group of infringers. Normally, the extra FedEx letter and the last email reminder are sent only if the infringer ignores the settlement request letter. In these cases, a 'case-handler' field is coded as empty. During the trial period, the compliance team at the Agency records the group to which a case belongs also in the 'case-handler' field. However, this group information is *not* recorded for the group designated as the 'default' group in each experiment (i.e., the High Price group in the pricing experiment and the control group in the message experiment). Thus, during the trial periods, the non-default groups do not get these two types of follow-up letters because the 'case-handler' field is not empty, even for infringers that have not made contact. This is sometimes corrected, but not always.

Table A9 reports, by group, the percentage of cases that receive the three types of follow-up letters. The differences suggest that the potential upward bias for the 'default' group is likely to be substantial for the message experiment. These incidences reflect the complexity and constraints that firms face when conducting field experiments. It is, at least, fortunate that we know the exact data-generating process, so we can discover any confounding factors.

Table A9: Follow-up letters by group

	Pricing experiment		Message experiment				
	High Price	Low Price	Control	Waive the Cost	Waive the Cost + Deadline	Discount	Discount + Deadline
Second letter	82.6%	79.6%	82.8%	72.8%	66.7%	71.7%	64.3%
FedEx letter	5.2%	0%	41.9%	16.3%	15.6%	14.5%	14.2%
Final email	7.9%	0%	14.1%	0%	0%	0%	0%

Notes: This table reports, by group, the proportion of cases that receive different types of follow-up letters.