Communication, Learning, and Bargaining Breakdown: An Empirical Analysis*

Matthew Backus† Thomas Blake‡
Jett Pettus§ Steven Tadelis¶

July 10, 2023

Abstract

Bargaining breakdown is common in bargaining in environments with incomplete information. We study whether, in these environments, permitting communication impacts bargaining outcomes. On May 23, 2016, eBay Germany's Best Offer platform introduced unstructured communication allowing desktop users, but not the mobile users, to accompany offers with a message. Using this natural experiment, our difference-in-differences approach documents a 14% decrease in the the rate of breakdown among compliers. Though adoption is immediate, the effect is not. We show, using text analysis, that the dynamics are consistent with repeat players learning how to use communication in bargaining. Finally, tying the two results together, we show that messages that emulate the text content of experienced sellers are more likely to be accepted. *JEL* classifications: C78, D82, D83, M21.

*We are grateful to numerous seminar and conference participants, as well as to Charles Angelucci, Anne Bartel, Wouter Dessein, Liran Einav, Guillaume Fréchette, Guillaume Haeringer, Alessandro Lizzeri, Bentley Macleod, Leslie Marx, Sarah Moshary, Michaela Pagel, Andrea Prat, Andrzej Skrzypacz, Yossi Feinberg, our anonymous reviewers and program committee member for ACM:EC 2021 and especially Jonah Rockoff for thoughtful comments, as well as many eBay employees for their input and support, and finally Brenden Eum and Alphonse Simon for excellent research assistance; all remaining errors are our own. Part of this research was supported by NSF Grant SES-1629060. Part of this research was conducted while Matthew Backus, Thomas Blake, and Steven Tadelis were affiliated with eBay Research Labs.

†UC Berkeley, NBER, and CEPR. Email: backus@berkeley.edu (corresponding author)
‡Amazon. Email: tomblake@gmail.com
§MIT. Email: jvpettus@mit.edu
¶UC Berkeley, NBER, CEPR, and CESifo. Email: stadelis@berkeley.edu

1 Introduction

Strategic bargaining is the mechanism of choice for many of our most important transactions: mergers and acquisitions, the sale of enterprise software, the purchase of a car or specialized machinery, climate change mitigation efforts, peace negotiations, and federal budgets, to name a few. Still, when parties to a negotiation do not know each other's outside options, each has a strategic incentive to overstate the strength of their position to get a better deal, sometimes resulting in costly delays and other times in a complete loss of socially beneficial trades (Myerson and Satterthwaite, 1983).

We begin with an open question in this domain: does unstructured communication improve the efficiency of bargaining outcomes in the field, or is it all "cheap talk?" Using large-scale bargaining data from the field, we document a sizable and significant positive, causal relationship between communication and the success of bargaining interactions. Exploring the initial introduction of the communication protocol, we find that the relationship grows over four weeks and then stabilizes. This feature motivated the development of our second question: do bargaining parties seem to learn how to communicate effectively? Applying text analysis techniques to users' messages, we document a convergent pattern which, we argue, represents repeat players learning how to communicate in the bargaining protocol. We show that this pattern is related to experience, maps well to the evolution of the treatment effect, and that the strategies adopted by experienced bargainers are correlated with successful negotiation. These results are consistent with players learning how to communicate in the spirit of "learning by doing" (Arrow, 1962).

Our findings inform online platforms, where bargaining is increasingly commonplace and the market design question of how and whether to regulate communication in bargaining is pertinent. eBay's Best Offer bargaining platform accounted for approximately 10% of eBay's trade volume, which was \$84 billion in 2016.[1] On the Chinese platform TaoBao, which hosted trade volumes of \$115 billion in 2016, bargaining is the standard for all transactions. And in 2014, Amazon.com introduced their own bargaining mechanism for the Amazon Marketplaces platform called "Make an Offer."

Our setting is the Best Offer bargaining platform on eBay.de, the German counterpart of eBay.com. Sellers who create a listing on the website may enable Best Offer, a

[1]See Backus et al. (2020) and eBay's 2016 10k SEC filing.

free feature that allows buyers to make offers below the listing price. An offer may be countered, and the counter-offer countered, etc., yielding a protocol very similar to sequential, alternating-offers bargaining (Rubinstein, 1982). Originally, eBay.de did not allow communication, and only numerical offers were relayed between the parties. However, in a policy change that took effect May 23, 2016, offers could be accompanied by a 250-character message when made from the eBay.de website. Importantly, the change did not affect buyers who made offers using their mobile devices. The sudden nature of the change and its incomplete coverage affords us a natural experiment to study the role of communication in bargaining.

Our core results are presented in Sections 5, 6, and 7. In Section 5 we find that, while the typical bargaining session in our sample was successful forty-four percent of the time, bargainers who used messages were eight percentage points more likely to transact than they would have been in the absence of communication. This corresponds to a fourteen percent reduction in the rate of bargaining breakdown among bargainers who used messages. Additionally, examining week-specific effects, we find that the effect of communication grows gradually over the four weeks after introduction. One mechanism that may drive this pattern is that it takes time for players to learn how to best make use of the new feature after the change.

Inspired by this observation, in Section 6 we present descriptive analyses of the textual content of communications. Most notably, we offer evidence that over the weeks following the introduction of communication, sellers' communication strategies evolve as they gain experience. In contrast, the content of buyers' communications is largely random, which is consistent with them being short-run players on eBay, who do not learn. We show that not only are repeat sellers' strategies changing over time, but that they are convergent, supporting the hypothesis that they are learning.

Combined, these findings offer suggestive evidence that the evolution of the treatment effects are driven by sellers learning how to communicate. We complete this story in Section 7 with evidence that, throughout the ten-week sample, messages that are similar in content to the messages sent by experienced sellers are more likely to succeed, with a magnitude that is remarkably close to our estimates of the treatment effect on the treated from the introduction of messaging. We also, in Appendix **??**, borrow tools from text analysis to learn about what sellers are converging to; though this analysis is entirely descriptive, the findings are consistent with behavioral and experimental work which shows that cost-based rationales are effective, and overly effusive communication strategies can be counterproductive.

Our paper is the first, to the best of our knowledge, to study the relationship between communication and bargaining breakdown using real-world bargaining interactions. Perhaps the closest paper to ours is a recent contribution by Green and Plunkett (2022). Where we use tools from text analysis to study the messages sent by bargainers, they use tools from reinforcement learning to study optimal responses to bargaining on the platform. This allows them to offer insights on the mistakes made by (merely) human bargainers. Where we use data from Best Offer in Germany, they use the Best Offer dataset from the US-based eBay.com, which was made available by Backus et al. (2020). Their paper, like ours, contributes to a vibrant and growing literature on bargaining that uses data from real bargaining environments.

There are prior theoretical and experimental literatures on communication in bargaining, despite the fact that communication is generally unmodeled by standard "tacit" models of bargaining (Crawford, 1990). Many theoretical treatments append a cheap-talk "pre-game" to an exiting bargaining model, as in Farrell and Gibbons (1989) and Cabral and Sákovics (1995). There are mixed theoretical conclusions on the efficiency of allowing cheap talk: most predict that, among rational actors, it is at best irrelevant, and possibly detrimental to bargaining efficiency.[2] In contrast, experimental work in the lab has found potential for communication to improve bargaining outcomes.[3] Radner and Schotter (1989) and Valley et al. (2002) are two pioneering studies that manipulated the availability of communication in bilateral negotiations. Both offered evidence that communication may permit bargaining efficiency that exceeds the theoretical upper bounds outlined by Myerson and Satterthwaite (1983).[4]

[2]See Goltsman et al. (2009) for a treatment and review of communication and bargaining in the cheap-talk setting described by Crawford and Sobel (1982).

[3]In experiments mirroring Crawford and Sobel (1982), Cai and Wang (2006) find that subjects consistently reveal "too much" information to be rationalized by equilibrium behavior. This is consistent with lie aversion (Gneezy, 2005) and guilt aversion (Charness and Dufwenberg, 2006), as well as communication fostering other-regarding preferences (Coffman and Niehaus, 2020). Crawford et al. (2013) argue that it is consistent with level-K reasoning, even in the absence of preferences for truthfulness, as L0 types anchor on the truth. In a recent paper, Bochet et al. (2023) study multi-issue bargaining in a lab experiment and show that access to more information increases agreement rates in small-surplus negotiations but not so in large-surplus negotiations.

[4]Subsequent extensions, however, were not uniformly positive on communication: Moore et al. (1999) stresses the limitations of electronic communication in particular, highlighting a role for group affiliation and rapport-building; Ert et al. (2014) construct an experimental scenario with misalignment of incentives, similar to bargaining over lemons, where communication elicits skepticism

These unstructured bargaining experiments, although few, are an important frontier in understanding real-world bargaining (Karagözoğlu, 2019). The richness and real-world character of our field data complements these experimental lab studies and contributes to the scope of our understanding of real-world bargaining. Where lab negotiations are typically low-stakes, asking prices in our data range from one to one thousand dollars, with an average of $100.60. Further, while lab negotiations span a single session, we study the evolution of communication by repeat users over a ten-week horizon. Finally, subjects in our study are real market participants engaging in real transactions, and the results have immediate practical implications for the design of the marketplace.[5]

Our work also contributes directly to the literature on learning in strategic settings. Empirical work in this area is particularly scarce because there are few opportunities to observe the introduction of a novel mechanism in a continuously operating marketplace. In this respect, our work is related to Doraszelski et al. (2018), who document bidding behavior in a newly opened electricity auction market. Our exercise builds on the recent introduction of text analysis and natural language processing; see Gentzkow et al. (2019a) for a survey. For the same reason that we need new tools to study text—that messages live in a high-dimensional space—we conjecture that it is a natural environment to study learning and experience.[6]

Last, communication and bargaining is also of interest in antitrust economics, where colluding parties have to bargain over the division of surplus. Perhaps the most closely related paper in this space is pioneering work by Aryal et al. (2021), who empirically tie communication in analyst calls in the airline industry to coordination of firm conduct. Also related, Cooper and Kühn (2014) experimentally study the role of communication in repeated Prisoners' Dilemma games, and Harrington and Ye (2019) show how cheap-talk communication among sellers in an upstream market can facilitate collusion when the downstream market is characterized by negotiations, as in many intermediate goods markets. Clark and Houde (2014) offer an empirical

and increases breakdown; Bolton et al. (2003) highlight drawbacks of communication in bargaining games with more than two players.

[5]In a previous study using a field experiment on eBay's Germany site, Bolton and Ockenfels (2014) show that consistent with lab results, buyer behavior exhibits signs of social preferences.

[6]Though no prior work isolates the interaction of communication and experience in bargaining, there is evidence of the importance of experience in bargaining: see, e.g. Card and Dahl (2012), who documented the role of expertise in bargaining agents in arbitration, and Backus et al. (2020), who found effects of experience on bargaining outcomes on eBay.com.

example of the role of communication in collusion, studying the collapse of a cartel resulting from antitrust enforcement equipped with a wire tap in Quebec. Finally, Byrne et al. (2019) conduct an audit study to understand how scripted, buyer-side communication affects bargaining outcomes with an oligopolist.

2 Our Empirical Setting: Bargaining on eBay.de

2.1 Best Offer Bargaining

eBay's online platform matches buyers to sellers who sell products ranging from art and collectibles to mobile phones, with over $95 billion dollars in gross merchandise volume worldwide in 2018. The platform operates in many markets throughout the world, the largest by revenue are the US (eBay.com) and Germany (eBay.de).[7]

Sellers can list items either as an auction or a fixed-price ("Buy-it-Now"). We focus on a subset of fixed-price listings for which the seller enables the "Best Offer" bargaining feature as shown in Figure 1. A buyer considering this listing can either purchase at the posted price (here, $55), or they can offer to purchase at a lower price. Following a buyer's offer, the seller is notified and may then accept, decline, or make a counter-offer. Following a counter-offer, the buyer may accept, decline, or make a counter-offer, and so on, in the spirit of Rubenstein-Ståhl alternating, sequential-offers bargaining. Offers by either party expire automatically after 48 hours.[8]

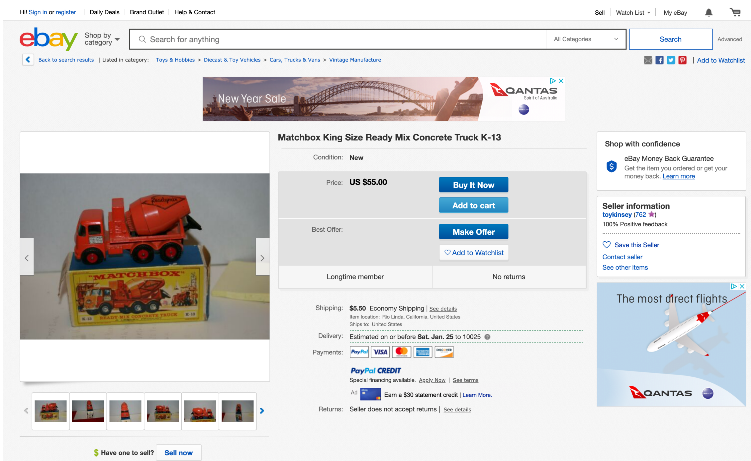
The Best Offer bargaining mechanisms on eBay.com and eBay.de are mostly identical except for one peculiarity that was unique to eBay.de up until the policy change we study: bargainers weren't allowed to communicate like on eBay.com. Figure 2(a) depicts the "Make Offer" interface through which buyers submit their offer, for eBay.com. If a buyer clicks the "add message to seller" button, they may include a free-form text message of up to 250 characters which will accompany their offer. Figure 2(b) depicts the parallel "Preis Vorschlagen" interface, for eBay.de. The red rectangle highlights where the missing option to send a message might have been.[9]

[7]Figure and revenue ranking based on eBay's 2018 Annual Report.

[8]For a thorough description of the Best Offer environment, see Backus et al. (2020).

[9]The reason for the difference seems to be a historical artifact: eBay.de is the successor of Alando.de, a clone of eBay.com created by the company Rocket Internet in 1999. 100 days after its creation, it was sold to eBay for $43 million. Thanks to Ariel Stern of HBS for pointing this out.

Figure 1: Example of a Best Offer Listing



Notes: This is an example View Item page for a Best-Offer enabled listing. A buyer may purchase at the asking price of $55 by clicking the Buy-It-Now button, or they may engage in bargaining by clicking the Make Offer button.

2.2 The Policy Change: May 23, 2016

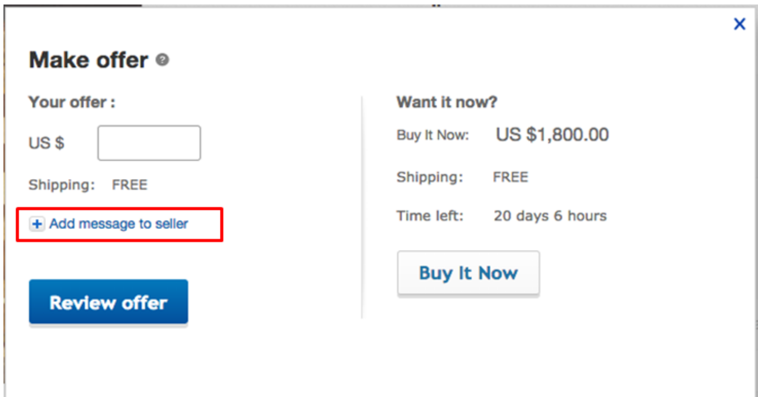
On May 23, 2016, the “add a message to seller” feature was added to eBay.de’s Best Offer bargaining platform. However, the rollout was not site-wide. The website eBay.de was updated, but the app for mobile users was not.

The policy change affords the main source of variation that we will exploit in this paper. Unlike many changes to the eBay platform, the introduction of messaging on eBay.de was not accompanied by a far-reaching “seller update”, and no other major changes to the eBay.de experience happened around this change. Moreover, the availability of an untreated group in the post period (mobile users) affords us a control group in order to separate changes in behavior from a secular trend.

3 Dataset and Sample Design

We obtained proprietary data from eBay to evaluate the effect of messaging on bargaining breakdown. We study bargaining *interactions*, defined as a buyer-item pair in which we observe at least one offer. Our main dataset includes all interactions for which the first buyer offer was made during an eight-week window, constructed to

Figure 2: Messages and Best Offer



(a) eBay.com

(b) eBay.de (before May 23, 2016)

Notes: Panel (a) depicts the "Make Offer" panel for eBay.com, the US site, where we have highlighted the "add message to seller" button. Panel (b) depicts the pre-treatment panel on eBay.de, where there is no option to send a message.

be four weeks before and four weeks after the introduction of messaging on May 23, 2016.[10]

3.1 Summary Statistics

Table 1 presents summary statistics for the main sample across several dimensions: listings, buyers, sellers, and interactions. This sample includes 3.3 million interactions involving 2.2 million unique listings, 444 thousand sellers, and 1.6 million buyers. The listings span all categories of eBay excluding real estate, automobiles, and tickets. As

[10]This window was chosen at the outset of our work to isolate the effect of the policy change from other changes to the platform. We explore the robustness of our estimates to different windows in Appendix Section ??.

documented by Backus et al. (2020), Best Offer is more frequently used in categories with substantial heterogeneity, such as Collectibles. Listings observed in our sample have on average 1.5 interactions (note that our sample excludes listings with zero interactions) and the distribution is highly skew: 79 percent have only one; 91 percent have two or fewer, 96 percent have three or fewer, and there is a right tail with many more.[11] 55 percent ultimately sell through the Best Offer mechanism.[12]

Buyers and sellers engage in multiple interactions–on average, 2.1 and 7.4, respectively–with substantial positive skew, suggesting a large right-tail of highly-active participants. About 60% of buyers are only observed in a single interaction, while the top decile participates in four or more interactions. Instead, 34% of sellers are observed only once, while the top decile participates in twelve or more interactions. This difference motivates our interpretation of buyers as short-run players and sellers as long-run players, which will be important for the analysis of learning in Section 6. These patterns hold for sales and purchases as well.

At the interaction level, which is the unit of observation for the empirical analysis that follows, we see that 44% of interactions end in a sale. These facts are consistent with findings in prior work using data from the U.S. site eBay.com (Backus et al., 2019, 2020). The final two interaction-level variables of Table 1 are of unique interest to this paper: interactions may be initiated by buyers using the desktop version of eBay (54%) or the mobile version (46%), which is important because only desktop buyers have the opportunity to send messages after the policy change. Finally, 50% of our sample falls after the policy change on May 23, 2016.

3.2 Additional Controls

In addition to the basic characteristics summarized in Table 1, we have a number of controls available to predict bargaining breakdown. These include the asking price of the seller; dummy variables by product category and condition (new, used, refurbished, or unknown); dummy variables for the day of the week on which the first offer in an interaction is made (to allow for differential behavior on weekends and weekdays),

[11] Partly for this reason, and partly because Best Offer listings tend to take longer to sell than, say, auctions, we ignore the possibility of competing offers by buyers. We find this natural, since sellers also have the option to hold an auction on the eBay platform, and they should do so if they anticipate many buyers.

[12] Listings that sold at the Buy-it-Now price are not classified as an interaction because no negotiation is involved. These sales do not appear in our dataset.

Table 1: Summary Statistics

	Mean	Std. Dev.	Skewness	Min	Max
Listing-Level Data					
Asking Price (USD)	100.6	151.5	2.889	1.050	999.9
Number of Interactions	1.490	2.566	82.62	1	1004
$\mathbb{1}$(Sold)	0.563				
N	2210575				
Seller-Level Data					
Number of Interactions	7.426	42.30	47.81	1	7142
Number of Sales	3.279	20.76	44.17	0	3115
N	443644				
Buyer-Level Data					
Number of Interactions	2.101	3.102	20.24	1	396
Number of Purchases	0.928	1.526	25.26	0	389
N	1567995				
Interaction-Level Data					
Number of Offers	1.873	1.116	1.535	1	6
$\mathbb{1}$(Ended in Sale)	0.442				
$\mathbb{1}$(Buyer on Desktop)	0.537				
$\mathbb{1}$(First offer after May 23)	0.501				
N	3294362				

Notes: This table presents summary statistics for the main dataset of Best Offer interactions taking place within an eight-week window, four weeks before and four weeks after the policy change on May 23, 2016. Fixing that set of interactions, we have constructed the summary statistics four ways: where the unit of observation is the listing (i.e., the product), the buyer, the seller, or the interaction itself.

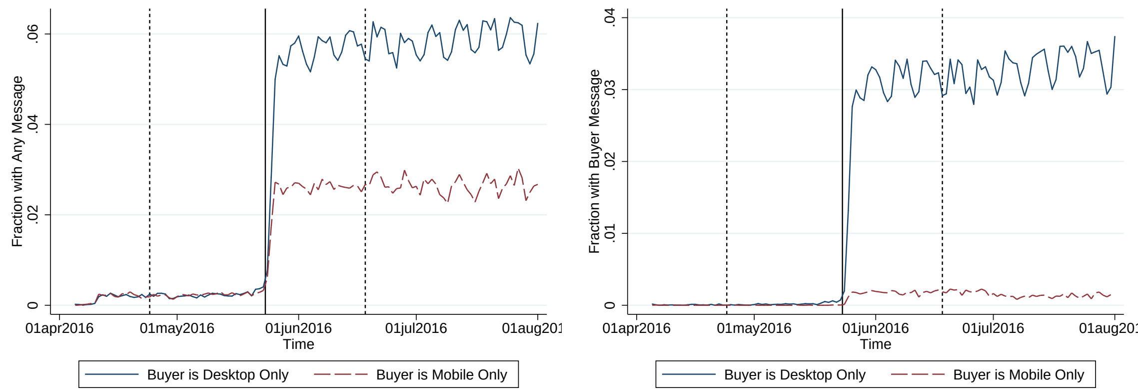
and a "holiday" dummy, which encodes all publicly observed holidays in Germany, as there is a particularly large number of them in May.[13] We also include controls for the weather in Frankfurt, Hesse, which we expect to be correlated with weather elsewhere in Germany, as weather conditions are both serially correlated and anecdotally cited as affecting online activity. These include a dummy for precipitation as well as the deviation of temperature from a linear trend over the sample.

3.3 Adoption on May 23, 2016

As depicted in Figure 3(a), users adopted messages rapidly: within days, the hazard rate with which bargaining interactions by desktop-only buyers involved messaging stabilized at approximately six percent. Both panels in Figure 3 distinguish between

[13]Just as differential behavior can occur on weekends and weekdays, the same may be true for holidays where people may either have more time to engage online, or may choose to disengage and travel. Publicly observed holidays in our sample include: May 1, Labor day; May 5, Ascension Day; May 8, Mother's Day; May 16, Whit Monday, and May 26, Corpus Christi. Just as

Figure 3: Launch of Messaging Feature



(a) Any Message

(b) Buyer Message

Notes: Panel (a) depicts the fraction of interactions, by first offer date, in which a message was included by the buyer or the seller. Panel (b) restricts attention to cases where the message was sent by a buyer. Both panels split the sample by whether the buyer was active on the desktop or mobile version of the platform. The solid vertical line represents the policy change. Dashed vertical lines depict the bounds of the eight-week window, centered at launch, of our main sample.

the case where the buyer is either exclusively making offers from a desktop computer or exclusively making offers from a mobile device. We do this because there was initially no messaging feature for the eBay.de mobile app. Therefore in Panel (a), which plots the hazard rate at which interactions involve any message (both seller and buyer messages), the hazard rate for the case where the buyer is on the mobile app is much lower. These are exclusively messages sent by sellers, to mobile-only buyers who cannot read them. In Panel (b), we plot the hazard rate of buyer messages, which confirms that mobile-only buyers are not sending messages after the change.[14]

4 Theory of Bargaining Breakdown

Theoretical models of bargaining with complete information, such as the Rubenstein-Stahl model which our setup mirrors in structure, have equilibria which are *efficient* and *immediate*. This is not what we find in our empirical setting, which is expected because the assumption of complete information seems implausible in our setting. Only 44.2% of bargaining transactions succeed, and we see substantial back-and-forth

[14]About 0.1% of offers included seller-side messages prior to the feature launch, which are rarely in German, indicating anomalies related to sellers registered on multiple sites. There are rare cases where mobile-only buyers appear to send messages, which seem to be database errors.

activity: 51.55% of transactions in our dataset go for at least two rounds, 21.72% for three or more, and 8.76% for four or more.[15]

Therefore our preferred theoretical lens is a model of incomplete information bargaining in the spirit of Myerson and Satterthwaite (1983), who offer a general treatment with independently distributed buyer and seller types that are private information, and that abstracts away from the bargaining protocol. An efficient bargaining mechanism would guarantee trade whenever the buyer's valuation v exceeds the seller's cost c, yet the seminal result of Myerson and Satterthwaite (1983) is that such a mechanism does not exist if we also require it to be incentive compatible, individually rational, and unsubsidized. Therefore, in general, we expect both efficient breakdown (when $v \leq c$) as well as inefficient breakdown (when $v > c$), regardless of the particular bargaining protocol.[16]

How would communication affect equilibrium play in bargaining games? Perhaps the most direct approach was taken by Farrell and Gibbons (1989), which considered the Myerson and Satterthwaite (1983) setup with a first stage cheap talk game. While they showed that agents send informative signals, these do not change the second best frontier. Valley et al. (2002) paints a different picture. In an experiment that implements the uniform case with $(v, c) \in [0, 1]^2$, they find that without text communication, subjects bargaining success and failure is remarkably well-predicted by the "linear solution" of Chatterjee and Samuelson (1983), in which trade happens when $v > c$ at a price that is the average of the bids in the double-auction game. However, when communication is available, agents outperform the Myerson-Satterthwaite second best, which is inconsistent with the standard assumptions of rational incomplete-information bargaining.

There are many possible "behavioral" models that can be explored in light of evidence of overcommunication in cheap-talk games.[17] We are unaware, however, of a theoretical model of bargaining that incorporates such features to study communication, let alone equilibrium selection in a Myerson-Satterthwaite setting. We therefore hope

[15]Here as elsewhere we are avoiding the temptation to count the initial posted price as an offer.

[16]In our empirical setting we cannot disentangle v and c, so we cannot distinguish between efficient and inefficient breakdown. This implies that the denominator of the effect we announce in the abstract, a 14% decline in bargaining breakdown, is too large relative to the real object of interest, which is the amount of inefficient breakdown, which suggests that our findings are understated.

[17]See Cai and Wang (2006), or Fréchette et al. (2021) for a more recent treatment.

that offering a first empirical perspective from the field will motivate further work on the question, as we discuss in further detail in our concluding discussion.

5 The Effect of Introducing Communication on the Likelihood of Bargaining Breakdown

5.1 Empirical Design and Identification

5.1.1 OLS and Endogeneity

Regressing a dummy for success on a dummy for the presence of a message is the wrong approach because the choice to include a message is endogenous. In particular, we expect a negative bias: bargainers send messages for items where bargaining is least likely to succeed, or in order to motivate a particularly aggressive offer.

We illustrate this by estimating a linear probability model in Appendix Section ??. The unconditional correlation between the presence of a message and bargaining success is negative, but once we condition on the log of the asking price it is small and positive. This reflects the fact that bargaining interactions with messages tend to involve more expensive goods, and that bargaining success is less likely for more expensive products, as documented by Backus et al. (2020) for eBay.com.

We expect this problem to be at least as important for unobservable characteristics than the asking price, and therefore rely on the policy change for a natural experiment and a more credible estimate of the effect of communication on bargaining breakdown.

5.1.2 Identification

The May 23, 2016 policy change, which introduced communication for desktop users, allows us to identify the effect of communication in two ways. The first is a simple pre-post design studying the change in the mean success rate of bargaining interactions before and after the policy change among desktop users. Our dependent variable is $\mathbb{1}$(success), a dummy for whether the bargaining interaction ends in a sale. Let X denote a set of controls; P (alternatively, "post") a dummy equal to one if the first offer in an interaction is in the post period, and D (alternatively, "desktop") a dummy equal to one if the buyer makes offers from the desktop version of eBay.de. Then, the pre-post estimate of the effect of communication on bargaining success is given by:

$$\hat{\beta}^{C}_{\text{pp}} = \mathbb{E}[\mathbb{1}(success)|X, P = 1, D = 1] - \mathbb{E}[\mathbb{1}(success)|X, P = 0, D = 1]. \tag{1}$$

An obvious objection here is that the pre-post design might conflate causal effects with secular trends. To deal with this, we can then use the untreated set of buyers who made offers on the mobile platform—where messaging was unavailable both before and after May 23, 2016—to control for common trends and alleviate concerns of endogeneity. This suggests a difference-in-differences design:

$$\begin{aligned}\hat{\beta}^{C}_{\text{dd}} =& (\mathbb{E}[\mathbb{1}(success)|X, P = 1, D = 1] - \mathbb{E}[\mathbb{1}(success)|X, P = 0, D = 1]) \\ &- (\mathbb{E}[\mathbb{1}(success)|X, P = 1, D = 0] - \mathbb{E}[\mathbb{1}(success)|X, P = 0, D = 0]).\end{aligned} \tag{2}$$

Both of the above strategies identify an effect that should be interpreted as an intent to treat (ITT) estimate, i.e., an estimate of the effect of the availability of communication, rather than the effect of actually choosing to communicate. In order to identify the treatment effect on the treated (TOT), we need to clarify who the compliers, i.e., those we actually consider to be treated, are.

5.1.3 ITT and TOT: Who are Compliers?

We say that a bargaining interaction is in the ITT group if the first offer occurs after May 23, 2016 and if the buyer uses the desktop version of eBay.de. Not every such interaction involves a message. By definition, compliers are interactions in our ITT group in which a message is sent by either party.[18] Note that the difference between the ITT group and the compliers group comes only from "never-takers," there are no "always-takers" since the feature did not exist prior to the policy change.

[18] Some messages were sent outside of this group, e.g., when a seller sends a message to a buyer who is using the mobile app. We exclude these for two reasons: first, because the buyer was mechanically unable to read the message, and second, because it turns out that using this less-restrictive definition will cause us to overstate the magnitude of the TOT estimate. Results are available on demand from the authors, however the intuition is simple: including messages in the pre-period or mobile set but not in the intent-to-treat group will inflate the coefficients on $\mathbb{1}$(Post) and $\mathbb{1}$(Desktop), respectively, and thereby depress the coefficient on $\mathbb{1}$(Post)$\cdot\mathbb{1}$(Desktop). Since the first-stage coefficient (which is in the denominator) is smaller, the IV effect will be inflated. Including all non-ITT messages will approximately double the estimate.

Table 2: Complier Characteristics

	$\mathbb{P}(x{=}1)$	$\mathbb{P}(x = 1 \mid \text{complier})$	$\frac{\mathbb{P}(x=1\mid\text{complier})}{\mathbb{P}(x=1)}$
Ask Price in (\$0,\$50)	0.52	0.41	0.78
Ask Price in [\$50,\$150)	0.27	0.29	1.08
Ask Price in [\$150,\$250)	0.09	0.11	1.30
Ask Price $\geq$ \$250	0.12	0.19	1.54
Friday, Saturday, or Sunday	0.42	0.42	1.00
Precipitation	0.46	0.67	1.46
Holiday	0.10	0.03	0.32
Post	0.50	1.00	2.00
Desktop	0.54	1.00	1.86

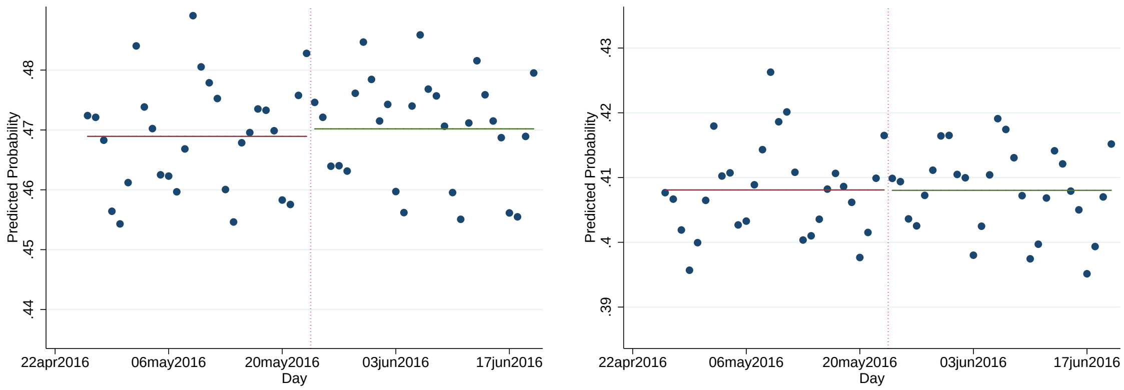
Notes: This table summarizes complier characteristics, i.e. the characteristics of interactions in the treatment group that "comply" and involve a message between bargainers. Each row represents a dummy variable which is taken to be x in the column formulas above. The unit of observation is an interaction between a buyer and an item.

Table 2 summarizes the characteristics of the compliers. Of particular note is the fact that bargainers are more likely to send messages when the asking price is higher, consistent with the intuition that communication is costly. The higher likelihood of precipitation and lower likelihood of a holiday is due mostly to early-May holidays and late-May rains, and we note that weekend users are no more likely to send messages.[19] In Appendix Table **??**, we extend Table 2 to include item categories and conditions. We find that bargainers are more likely to send messages for categories such as "Antiques & Art", "Cameras & Camcorders", and "Musical Instruments"; whereas categories like "Clothing & Accessories" and "Movies & DvDs" are underepresented by compliers. The prevalence of some of these categories may be explained by price; for instance, "Cameras & Camcorders", and "Musical Instruments" are on average the two highest priced categories in our data. Further, interactions are more likely to include a message when the item is refurbished, used, or unknown. This may be consistent with a story of incomplete information, whereby buyers use messages to learn about the properties of an item.

It is natural to wonder whether despite our controls, unobservable characteristics of the listings are generating compositional differences in the periods before and after the change. Therefore, in Figure 4 we document the *predicted success rate* conditional on those controls for all interactions for both samples—desktop and mobile—where the first-stage regression excludes the dummy for treatment as well as the time trend. We see no substantive change in the predicted probability of success conditional on

[19] We group Friday, Saturday, and Sunday based on the OLS coefficients reported in Table **??**, where weekends appear to be discretely different from weekdays.

Figure 4: Predicted Success Rates



(a) Desktop

(b) Mobile

Notes: Panel (a) depicts predicted success rates using a large set of controls (ln(ask price); category by condition fixed effects; day of week, precipitation, and holiday dummies and the temperature) for desktop users only. Panel (b) replicates this for the mobile users. The vertical axes on both plots are scaled identically subject to a location shift.

observable covariates for either group. While this does not rule out compositional changes coming from unobservable characteristics that are uncorrelated with observed characteristics, it does make it harder to tell such a story, since it would be surprising for such omitted variables to be uncorrelated with our controls (Oster, 2019).

5.2 Empirical Results

5.2.1 Regression Analysis

In what follows, we first generate precise estimates $\hat{\beta}^C_{pp}$ and $\hat{\beta}^C_{dd}$, and second, we identify a treatment effect on the treated because, recall from Figure 3, only a small fraction of interactions in the post period involve a message. To accomplish the former, we reformulate (1) and (2) in terms of a linear probability model and estimate them using OLS (in all of what follows we omit observation indices):

$$\mathbb{1}(success) = \beta^0_{pp} + X\beta^1_{pp} + P\beta^C_{pp} + \varepsilon, \tag{3}$$

and

$$\mathbb{1}(success) = \beta^0_{dd} + X\beta^1_{dd} + P\beta^2_{dd} + D\beta^3_{dd} + PD\beta^C_{dd} + \varepsilon. \tag{4}$$

Next we consider the treatment effect on the treated. We estimate this by reformulating (4) as an instrumental variables regression:

$$\mathbb{1}(success) = \beta_{iv}^0 + X\beta_{iv}^1 + P\beta_{iv}^2 + D\beta_{iv}^3 + \mathbb{1}(complier)\beta_{iv}^C + \varepsilon, \quad (5)$$

where $P \cdot D$ is an instrument for $\mathbb{1}$(complier), and $\mathbb{1}$(complier) is a dummy for interactions in the ITT group that include a message. Importantly, this does not employ any variation that is not already used in the difference-in-differences estimator. Instead, it rescales the estimate to the end of interpreting its economic meaning. As with all estimates of the treatment effect on the treated, it should be interpreted with caution, as bargainers' decision of whether to send a message may be correlated with unmodeled heterogeneity in the treatment effect.

Results for each of these three estimators, both with and without the controls discussed in Section 2, are presented in Table 3. We see a large effect of the policy change on success for desktop users, 0.46 percentage points in model (1), which is attenuated by the inclusion of controls to a statistically insignificant effect of 0.23 percentage points in model (2). As we show in the next section, the control that attenuates the result is the time trend. In the post period, there is a substantial positive drift, which stabilizes after a few weeks. Is this a delayed treatment effect, or a secular trend on the platform? To distinguish between these hypotheses, we need a suitable control group; here, interactions involving buyers who use the mobile app.

Critically, as we see in models (3) and (4), our estimates of β_{pp}^C for the placebo sample of mobile users are very close to zero. Therefore, difference-in-differences estimates in models (5) and (6), where the relevant estimate is the term on the interaction effect, are of the same order as those in (1): we estimate effects of 0.40 and 0.42, respectively. In particular for specification (6), with the rich set of controls, the mobile group is distinguishing common time trend from the ITT estimate.

Finally, models (7) and (8) report the IV estimates which are rescaled to obtain the TOT estimate. As only a small fraction of users, approximately six percent, actually send messages in the post period, this means that the roughly half-percent effect on conversion for the Best Offer environment at large translates to a substantially larger and economically important effect on the treated. To be precise, versus a baseline probability of success near 44% for the full sample (from Table 1), interactions that involve messages in the treated group are 7.44 (or 7.73, with controls) percentage points more likely to succeed. The inclusion of these controls (time trend; ln(ask price); category by condition fixed effects; day of week, precipitation, and holiday dummies and the temperature) allows us to rule out the most salient alternative hypotheses by which the effect is driven by compositional changes spurred by the introduction of communication, e.g. if messaging prompts buyers to bargain over cheaper goods which also have lower rates of breakdown. This leaves is with our main result: a 14% decrease in the rate of bargaining breakdown among the compliers.

Treatment effects on the treated should be interpreted with caution, but there are two senses in which our estimate is conservative. First, as we have noted, we cannot disentangle efficient and inefficient bargaining breakdown. If the true object of interest is the latter, then the denominator of that 14% calculation, which includes all breakdown, is too large. Second, however, is that this estimate is net of any increased rate of platform disintermediation. A common concern among online platforms is that users will use communication options to take their transaction off-platform, avoiding fees. This would appear as bargaining breakdown in our data, and so this theory would suggest that our results understate the true effect of communication on bargaining breakdown. Both of these points suggest that our already economically significant estimate of 14% could be interpreted as a lower bound.

Finally, a note on size and power: we are looking at extremely small effects, on the order of half a percentage point. They are small because take-up of the messaging feature among the ITT group, i.e., the compliance rate, is low, at approximately six percent. While we have a lot of data in our main sample—3.41 million bargaining interactions—this turns out to be close to what we need. Simple power calculations for the detection of an effect of 0.005 with a baseline success rate of 0.442 (borrowed from Table 1) with a confidence of $\alpha = 0.05$ and a power of fifty percent implies that we need a dataset of 2.72 million experimentally generated observations equally divided between treated and not. Therefore there are limitations on how far we can push the data to understand heterogeneity in the effects of communication.

While the low compliance rate is, in that sense, an empirical challenge, there is also a sense in which it is an advantage of our environment. If the greater fraction of buyers sent messages—and those messages were important for bargaining—then their communication might have an equilibrium effect on the quantity, composition, or listing style of goods on the platform. In the language of Angrist et al. (1996), these are concerns about the stable unit treatment value assumption (SUTVA)—i.e., that treatment of some observations has spillover effects on the effectiveness of the treatment for others. The fact that the compliance rate is very low reassures us that

Table 3: Effect of Messaging on Success Rate

	Desktop		Mobile		Differences		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Post)	0.0046*	0.0023	0.0006	0.0031	0.0006	0.0005	0.0006	0.0009
	(0.0013)	(0.0019)	(0.0013)	(0.0019)	(0.0013)	(0.0015)	(0.0013)	(0.0015)
$\mathbb{1}$(Desktop)					0.0599*	0.0482*	0.0599*	0.0482*
					(0.0013)	(0.0011)	(0.0013)	(0.0011)
$\mathbb{1}$(Post) $\cdot$ $\mathbb{1}$(Desktop)					0.0040*	0.0042*		
					(0.0013)	(0.0012)		
$\mathbb{1}$(Complier)							0.0744*	0.0773*
							(0.0240)	(0.0226)
Controls		✓		✓		✓		✓
N	1770261	1770261	1524101	1524101	3294362	3294362	3294362	3294362

Notes: This table depicts our main regression results. Even-numbered specifications include controls (time trend; ln(ask price); category by condition fixed effects; day of week, precipitation, and holiday dummies and the temperature), while odd ones include none. Specifications (1) and (2) report a linear regression on a dummy for being in the treatment period on a sample of desktop users (who are treated), while specifications (3) and (4) conduct the same exercise on a placebo sample of untreated mobile users. Specifications (5) and (6) combine the samples in a difference-in-differences approach, and finally specifications (7) and (8) take an instrumental variables approach, using the same variation, to identify a treatment effect on the treated for the combined sample. Heteroskedacticity-robust standard errors, clustered by seller, are reported in parentheses, and * denotes statistical significance at $\alpha = 0.05$.

they are not economically significant, and that we can identify a partial equilibrium effect, i.e. conditional on the broader state of eBay.de in the Summer of 2016.

Additional robustness checks and extensions are discussed in more detail in Appendices **??**, **??**, and **??**. We find that the results are robust to the inclusion of seller fixed effects, and we also explore alternative outcomes such as price and number of offers.

5.2.2 Graphical Intuition

In order to offer some intuition for the results in Table 3, Figure 5 presents the data aggregated to the daily level, both with and without residualization on a large set of controls.[20] In both cases we see an apparent jump in the success rate of approximately half a percent. Two other features are apparent—first, there is substantial variation between days in the success rate of bargaining interactions, although this is rather smaller when we condition on the set of controls. Second, and more importantly, it appears that there is a positive drift in the residuals during the post period. In Section 6 we will offer a simple economic explanation for this finding: buyers and sellers are learning to communicate in the weeks following the policy change, leading to new behavior and better outcomes.

Consistent with models (3) and (4) of Table 3, we see no evidence of change in the likelihood that interactions are successful for buyers using the mobile platform. We also see no positive drift in the residuals in the post period. This rationalizes our finding for the difference-in-differences estimator $\hat{\beta}^C_{dd}$ in models (5),(6),(7), and (8).

5.2.3 Week-Specific Treatment Effects: Parallel Trends and Dynamics

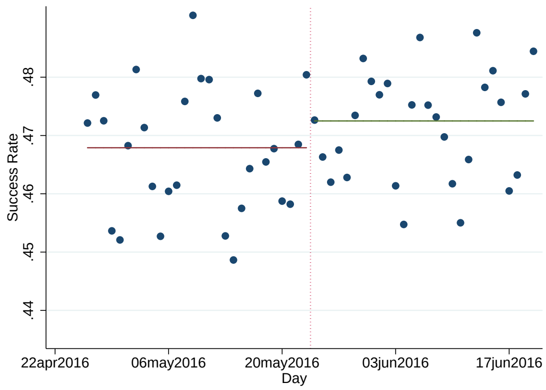
Next, we estimate a variant of (4) with week-specific effects over an extended version of our sample (six weeks before, and ten weeks after):

$$\mathbb{1}(success) = \beta^0_{dd} + X\beta^1_{dd} + P\beta^2_{dd} + D\beta^3_{dd} + \sum_t D \cdot \mathbb{1}(\text{week } t)\beta^{C,t}_{dd} + \varepsilon. \tag{6}$$

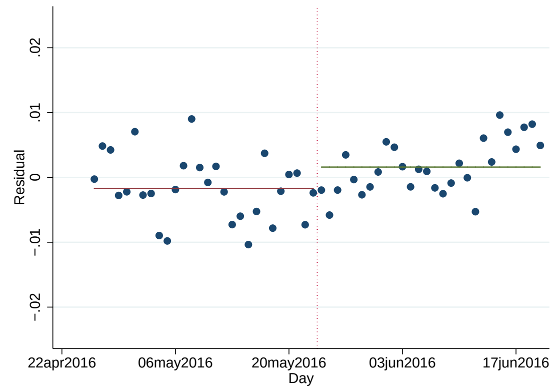
We are interested in this model for two reasons. First, the divergence in the residualized scatterplots for desktop, Figure 5(b), and mobile, Figure 5(d), in the post-treatment

[20]Note that while we will include it in the regression analysis that follows, we have excluded a time trend from the set of controls here. Coefficients from the OLS regression that generates these residuals are reported in model (6) of Table **??** in the appendix, where we present OLS results as a straw man alternative to our empirical design.

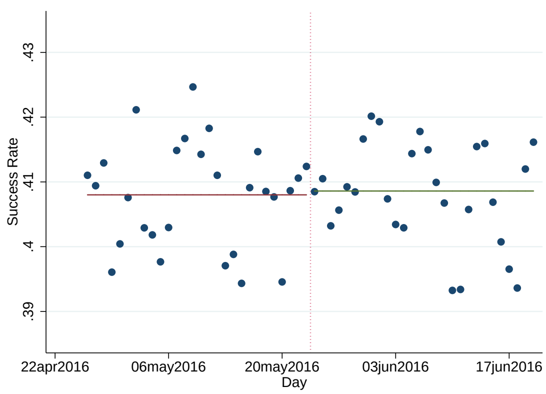
Figure 5: Bargaining Success Rates



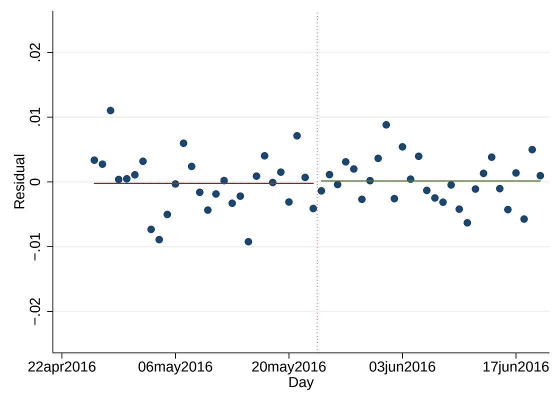
(a) Desktop: Raw



(b) Desktop: Residuals



(c) Mobile: Raw



(d) Mobile: Residuals

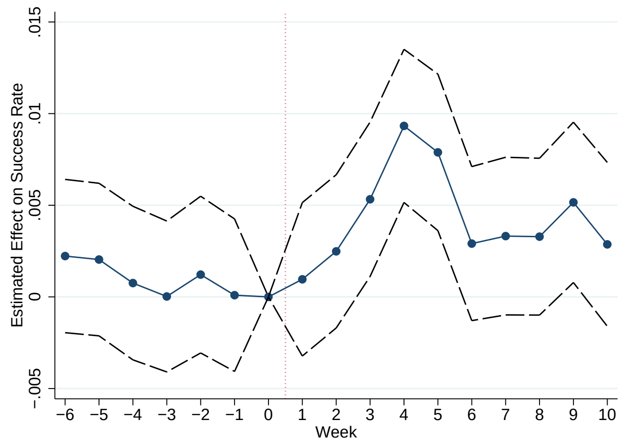
Notes: Panels (a) and (c) depict scatterplots of the raw daily success rates for bargaining interactions grouped by the date of the first offer for desktop and mobile, respectively. Panels (b) and (d) depict residuals from a linear probability model regressing a dummy for successful bargaining on a large series of covariates (ln(ask price); category by condition fixed effects; day of week, precipitation, and holiday dummies and the temperature) for desktop and mobile, respectively. The vertical axes on both plots are scaled identically subject to a location shift.

period suggest that the effect of communication on bargaining is not immediate but delayed. Estimating week-specific effects will allow us to characterize these dynamics.

Second, following Autor (2003), estimating week-specific effects in the pre-period allows a partial test of the parallel trends assumption. If estimated week-specific effects in the pre-period are significant, this would violate Granger causality—that is, the effect would precede the cause, marking a failure of our identifying assumptions.

Results are depicted in Figure 6. We normalize the effect for week zero (just before the policy change) to zero. In the pre-period, our test of the joint significance of the

Figure 6: Week-Specific Effects



Notes: This figure depicts week-specific effects using the diff-in-diff approach with the main sample, 6 weeks before and 10 weeks after the policy change. The omitted coefficient (normalized to zero) is the week just prior to the change. Dashed lines represent a 95% confidence interval with heteroskedacticity-robust standard errors clustered by seller.

coefficients fails to reject with a F-statistic of 0.41 (and an associated critical p value of 0.8737). Therefore, we find neither a violation of the parallel trends assumption nor of Granger causality for our sample. Furthermore, the model permits us to interpret the positive drift in the post period from Figure 5(b) as a time-variant effect of communication. Despite almost instantaneous adoption, it took several weeks for the effects to be fully seen in the probability of bargaining success—this is perhaps not surprising, as conventions for communication may have taken some time to stabilize. We investigate this further with the content of the messages in Section 6.

6 Evidence of Learning from Text Analysis

In Section 5, we found a positive and significant relationship between bargaining efficiency and messaging. Moreover, by analyzing the week-specific effects of messaging, we observed that it took time for the full effect of communication to materialize. In this section, we explore the messages themselves in order to offer a plausible economic story for the dynamics of the week-specific treatment effects depicted in Figure 6.

We compute the change in messaging content across weeks for both buyers and sellers, and are able to uncover some compelling patterns in messaging content. We find that messages sent by repeat sellers, or sellers who are sending multiple messages in our sample, are becoming more similar in content as the weeks pass. We also find convergent patterns in seller messages: specifically, the rate at which seller messages are changing is decreasing. These trends in seller messages are consistent with experience-based learning in which sellers adopt messaging strategies over time.

Similar patterns are not apparent for buyers in aggregate, which is consistent with our hypothesis. Recall from Section 3.1 that buyers are short-run players, and so we should not expect to see evidence of learning. We would expect similar findings for the subset of experienced buyers, but that sample turns out to be too small for our text analysis approach.

6.1 Messaging Data

We have 248,722 messages of buyer and seller interactions for the ten-week period succeeding the introduction of messaging beginning on May 25, 2016.[21] We process these messages through the following steps: First, we identify and keep only the messages sent in German in order to maintain a common corpus of words for our analysis; this makes up the vast majority (81.1 percent) of our dataset with the next most common language being English (which accounts for 6.4 percent of the messages). Second, the lower-cased messages are stripped of non-alphabetic characters, urls, extra spaces, and a list of common stop words—these are words such as "and" and "the" that provide little meaning in our messages. We then apply NLTK's German Snowball Stemmer (Bird et al. (2009)) to the tokens (ie. words) in each message so that they are transformed to their original stems. For instance, the word "angeboten" ("offered") is minimized to "angebot" ("offer"). This step is common practice in natural language processing and allows us to consider effectively synonymous words as the same word.[22]

The final reduced dataset amounts to 209,658 messages split into 93,577 buyer messages and 116,081 seller messages. 93.78% of messages are the first in an interaction. 99.64% are within the first two.[23] Table 4 provides descriptive statistics for buyer and seller messages split by experience level, which we define as the total count of messages sent by that seller or buyer over the ten-week period. Here we see that

[21] Our textual analysis starts two days after the change due to low take up on May 23-24, 2016.

[22] Please refer to Appendix Section **????** for a more detailed discussion of the construction of our messaging dataset, including a description of the list of stop words we remove from our dataset.

[23] We have also run the analyses that follow using only the first message in an interaction, and found no qualitative differences.

Table 4: Buyer and Seller Messaging Characteristics Based on Experience

	Number of Messages	Unique Individuals	Average Message Length	Success Frequency
Buyer				
All Messages	93576	76415	9.293	0.309
One Message	64648	64648	9.250	0.339
Two - Four Messages	25595	11271	9.379	0.244
Five+ Messages	3333	496	9.475	0.240
Seller				
All Messages	116081	60076	8.513	0.228
One Message	40350	40350	8.218	0.243
Two - Four Messages	40541	16375	8.561	0.213
Five+ Messages	35190	3351	8.796	0.227

Notes: In the first column messages are split based on whether the message's corresponding buyer or seller sent one, two to four, or five plus messages. Average message length refers to the average number of tokens in each message for that group. Success is defined by whether that message ends in a sale.

Table 5: Buyer and Seller Most Common Tokens

Rank	Buyer Token	Translation	Frequency	Seller Token	Translation	Frequency
1	Hallo	Hello	0.031	Nicht	Not	0.032
2	Wurd	Would	0.030	Hallo	Hello	0.024
3	Gruss	Greeting	0.025	Gruss	Greeting	0.023
4	Versand	Shipping	0.018	Preis	Price	0.021
5	Nicht	Not	0.016	Euro	Euro	0.020
6	Euro	Euro	0.015	Versand	Shipping	0.016
7	Mfg	Kind regards	0.011	Leid	Unfortunately	0.015
8	Dank	Thanks	0.011	Dank	Thanks	0.013
9	Kauf	Purchase	0.010	Schon	Beautiful	0.012
10	Preis	Price	0.010	Mfg	Kind regards	0.011

Notes: This table reports the frequency of the ten most common tokens in our processed dataset for buyers and sellers. Tokens are translated to English for readability.

buyers tend to be "short-run" players, with only 496 buyers sending five or more messages. In contrast, individual sellers are more persistent in our dataset—there are 3,351 who are sending five or more messages. In Appendix Tables **??** and **??**, we include additional descriptive statistics for buyer and seller messages based on experience level. Specifically, we include statistics on asking price, categories for items listed, and the reported conditions of the items.

Next, Table 5 depicts the ten most common tokens for buyers and sellers. Here, sellers messages appear to be slightly more negative in messaging content as "not" is the second most common word; additionally, "unfortunately" is the eighth most frequent token to appear in seller messages. For a more thorough description of the messages in our dataset, see Appendix Section **??**.

6.2 Empirical Challenges and Methods

Representing textual data constitutes a unique and increasingly common challenge in empirical research. In order to analyze seller and buyer messages, we must first construct a measure for the content of each message.

To do this, we split each message into a series of bigrams, a two-word pairing formed from consecutive words. For example, message m as "this is my last offer" would be broken up into the parts: ["this is," "is my," "my last," and "last offer"]. Splitting the messages into bigrams, rather than single-word tokens, allows us to simplify each message while still incorporating some level of context in our textual analysis. An example of this are the words "not" and "fair:" put together, "not fair" has a much different meaning than the two words apart.

Next, we collapse our data into what is known as a "bag-of-words," or in our case, a "bag-of-bigrams," where each message is a row and each bigram a column. This exercise is frequently used in natural language processing (see Gentzkow et al. (2019a)); however, as analyzing textual data is still relatively new in economics, we offer a detailed exposition below.

Our bag-of-bigrams is in the form of the matrix $\mathbf{C}_i$, where element $c_{i,mj}$ corresponds to the number of counts for phrase j in message m for group i, ie. some group of buyers or sellers. $\mathbf{C}_i$ is a high dimensional matrix. For instance, $\mathbf{C}_s$ for all seller messages in our dataset makes up a 116,081-by-287,241 matrix accounting for 116,081 seller messages and the 287,241 distinct bigrams that appear in these messages. Similarly, $\mathbf{C}_b$ for buyer messages forms a matrix with dimensions 93,577-by-331,904.

We are interested in changes in messaging content at the weekly level: Thus, we collapse $\mathbf{C}_i$ to $\mathbf{C}'_i$. $\mathbf{C}'_i$ is a ten-by-X matrix with each row corresponding to a week in our ten-week sample and X representing the number of distinct bigrams sent by group i. Element $c'_{i,wj}$ then equates to the number of times phrase j appears in the messages sent during week w by group i.

Finally, we take vector $\mathbf{v}_w$ from row w of $\mathbf{C}'_i$ and compute the cosine distance between $\mathbf{v}_w$ and all rows in $\mathbf{C}'_i$. The cosine distance between vector $\mathbf{v}_{w1}$ and $\mathbf{v}_{w2}$ is given by

$$1 - \frac{\mathbf{v}_{w1} \cdot \mathbf{v}_{w2}}{|\mathbf{v}_{w1}||\mathbf{v}_{w2}|}, \tag{7}$$

where $\cdot$ represents the dot product and $|\mathbf{v}_w|$ is the ℓ^2 norm.

The cosine distance measures one minus the cosine of the angle between $\mathbf{v}_{w1}$ and $\mathbf{v}_{w2}$ and is a standard method for computing text dissimilarity. The normalizing term in the denominator of (7) is desirable in our case as it scales the distance between the two vectors by each vector's length. This is essential in our context as the sum of bigram counts varies week from week due to changes in the take-up of messaging, holidays, etc. Finally, the cosine distance is our metric of choice as it supplies an intuitive measure of the distance between bigram counts across weeks. Since our vectors are by construction composed of nonnegative values, the cosine distance in this case will always be in the range $[0, 1]$. Here, two weeks with orthogonal vectors will have a cosine distance of 1, whereas the cosine distance will approach 0 as the two vectors get more similar in counts.

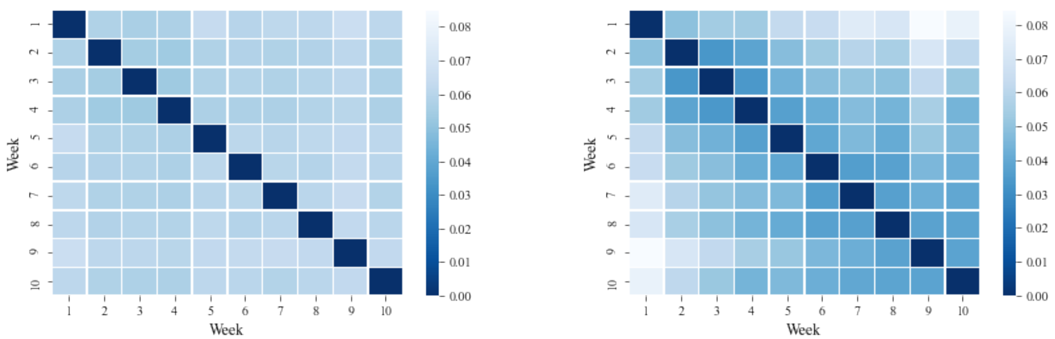
Finally, in order to clearly present the cosine distances across each pairing of weeks, we construct a ten-by-ten matrix of cosine distances with the $w_1 w_2$th element corresponding to the cosine distance between the vector of bigram counts in week w_1 and the vector of bigram counts in week w_2.

6.3 Dynamics of Communication by Experience

In Figure 7, we present our results through a heat map depicting the cosine distances between the bigram counts across weeks for buyers, Panel (a), and sellers, Panel (b). The colors indicate cosine distances in which the lighter boxes convey greater differences in messaging content, while the messages get more similar as the boxes get darker.

In Panel (a) of Figure 7, the cosine distances between buyer messages is stable across our ten-week period; we can see this as all the boxes in the heat map appear to be similar shades of blue. Table 4 offers a plausible explanation for this consistency in buyer messages across weeks: There are few buyers in our sample sending multiple messages. Specifically, only 3,333 messages are sent from buyers whose total message count is five or more. In contrast, 35,190 messages are sent by sellers with five or more messages. Due to the lack of repeat buyers, any changes in Panel (a) are presumably due to noise and week-effects. For instance, we observe differential usage of the word "urlaub" or "vacation" in our sample. Buyer messages include "vacation" 61 times during the week of July 20th, while they only mention "vacation" 21 times during the first week of our sample (May 25 – 31).

Figure 7: Cosine Distance of Buyer and Seller Messages by Week following the Introduction of Messaging



(a) Buyer (b) Seller

Notes: This figure presents the cosine distance of the bigram counts in the messages between each of the ten weeks following May 25, 2016, for buyers, in Panel (a), and sellers, in Panel (b).

Panel (b) of Figure 7 portrays starkly different results; here, there are clearly changes in content across sellers messages from week to week as made evident due to the patterns of color occurring in seller messages. Still, it is challenging to discern exactly how seller messages are transforming. In order to amend this, there are two ways in which we can more intuitively decipher the patterns presented in the heat maps. First, we can more closely observe the bottom gradient of the heat maps, where we are comparing the differences in messaging content between week w and week 10. Second, we can plot the off-diagonals of the heat maps in order to see the rate of change in message content. The former will tell us whether the content of the text is changing systematically, where the latter will tell us whether it is *convergent*. That is, if the rate of change is decreasing. We take this as evidence of the convex pattern that is signature of models of learning and information. To wit, our first experiences teach us more, on the margin, than those that follow.

Figure 8 depicts a plot of the the cosine distances between week w and week ten (the bottom gradient) for buyers and sellers. In this figure, we also depict the results for (nested) subsamples with different experience levels, where we define experience by the number of messages sent by that seller/buyer over our entire sample. Separating by experience level will allow us to isolate the sellers for whom we are more likely to observe patterns of learning. The scale of each plot is normalized to 1 in the last week;

we do this because sampling variation, which is more salient as we restrict the sample size, biases the cosine distance measure upwards and makes levels uninterpretable.[24] Note also that the scales of panels (a) and (b) are different.

Panel (a) depicts the differences in messages between week w and week 10 for buyers. The line for "All" includes all buyers, and corresponds directly to the values in the bottom row of the heat map in panel (a) of Figure 7. For buyers with one, two, or three or more messages, the cosine distance does not appear to be changing. For buyers with four, five, and six or more messages, the cosine distance is decreasing over time, but this is obscured by sampling variation due to the small number of buyers in these sets.

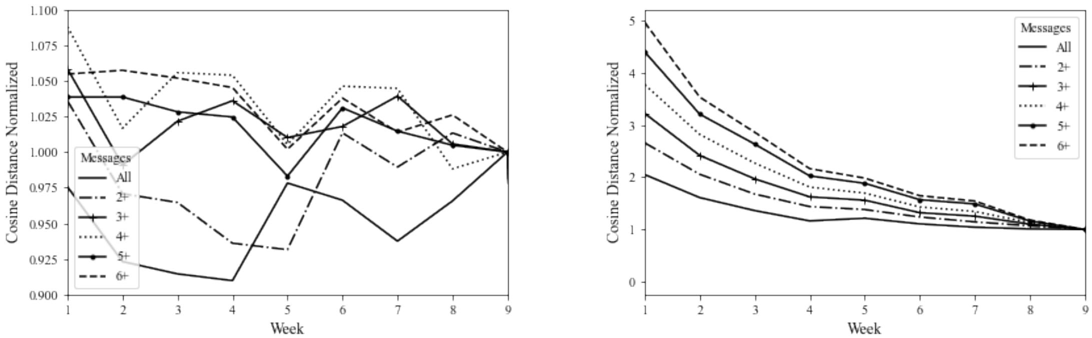
In Panel (b), we again observe sharply different results for sellers. Here again, the line for "All" includes all sellers, and corresponds directly to the values in the bottom row of the heat map in panel (b) of Figure 7. For all samples, we see that as the weeks get closer to week 10, seller messages are increasingly more similar in content to the messages from this last week. Note also that the change is more pronounced among more experienced sellers. For instance, we see the sharpest trend for sellers that sent 6 or more messages, where there is a substantial difference in the cosine distance between the set of week 1 and 10 messages and the set of week 9 and 10 messages. These patterns suggest that as the weeks pass, sellers may be adopting new messaging strategies; moreover, it appears that repeat sellers are driving the changes in Figure 7.

So far, we have established that repeat sellers' messaging strategies are changing as they accumulate experience. We would like to ask whether, in addition, the changes are convergent, i.e. whether they represent a pattern that reflects learning. To do this, we plot the cosine distance across different periods of the same length to see whether the differences in message content are decreasing at a slower or faster rate over time. The objects we are depicting correspond to the off-diagonal elements of the heat map in Figure 7. The different off-diagonals represent differences over varying lengths of time; we depict them all because we are concerned that high-frequency differences may exaggerate the ratio of noise to signal.

Figure 9, plots the off-diagonals for all buyers, Panel (a), and sellers, Panel (b). In this plot, Δx for week w corresponds to the cosine distance in messages between week w and week $w - x$. The astute reader will note that, e.g., the plot of $\Delta 3$ does not

[24]See Appendix Figure **??** for the non-normalized version.

Figure 8: The Bottom Gradient for Buyers and Sellers with Different Experience Levels



(a) Buyer (b) Seller

Notes: This figure presents the cosine distance between the messages sent in weeks x and 10 for buyers, Panel (a), and sellers, Panel (b). The cosine distance is scaled by the distance between week 9 and week 10 messages. Each panel is cut by groups, where All includes our entire sample of buyers/sellers, 2+ indicates our sample of messages sent by buyers/sellers who sent 2 or more messages, 3+ from our sample of messages sent by buyers/sellers who sent 3 or more messages, and so on.

begin until week 4 because we need four weeks of data to construct cosine distance on a three-week difference. As suspected, Panel (c) portrays a noisy plot with no obvious patterns in buyer messages; on the other hand, there appears to be a downward trend in the Δx's in Panel (d). Thus, while seller messages are becoming more similar over time, they do so at a declining rate.

Figure 9 depicts the off-diagonals corresponding to the "All" groups of Figure 8, but from the latter we saw that repeat sellers are changing their messaging content more than transient ones. So, in Figure 10 we reproduces the same exercise for sellers who sent 5 or more messages in our ten-week period. As a reference point, the heat map for this group of sellers is depicted in Panel (a). Already, we see starker patterns in the change in message content across weeks. Panel (b) presents the off-diagonal plot for this group of sellers. Again, the patterns that previously emerged for our entire dataset of sellers become much more explicit when we restrict our analysis to repeat sellers. For instance, from $\Delta 5$ in Panel (b) we can see that sellers' messages between weeks 8 and 3 are more similar than in comparison to the messages for weeks 7 and 2 and weeks 6 and 1. Or in other words, as repeat sellers learn to use similar messaging strategies over time, they are doing so at a declining rate as the weeks pass. This convergent path in which sellers are changing their messaging content faster in the beginning weeks following the messaging intervention, and then more slowly as time goes by, is consistent with experienced sellers learning how to use messages to facilitate transactions.

Most importantly, these results offer a simple explanation for the dynamics seen in Section 5.2.3, where we find a time-variant effect of communication on bargaining success that stabilizes following week five.

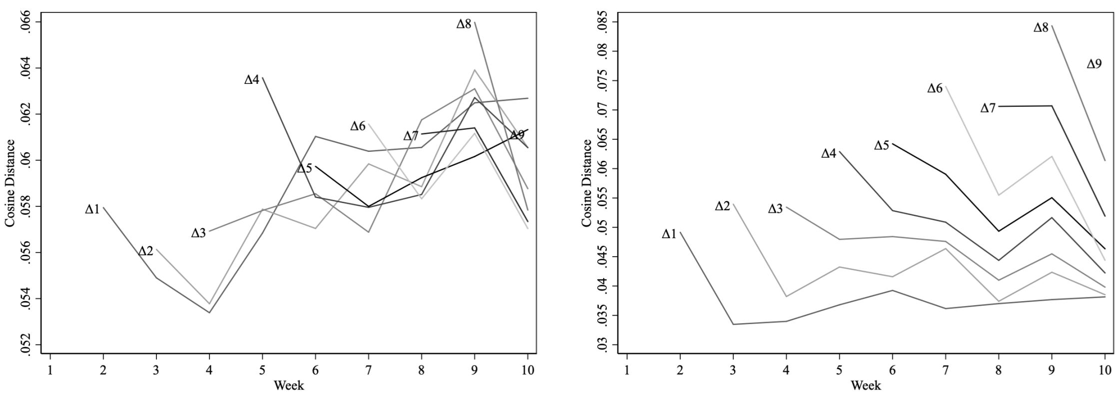
Further evidence is offered in Figure **??** of the Appendix where we have included the heat maps and off-diagonal plots for sellers who sent two, three, and four or more messages. In addition, Figure **??** in the Appendix presents the results for both sellers and buyers who appear only once in our dataset. Here, the effects presented above are completely attenuated when looking at this group of sellers, further providing evidence that these patterns of convergence are due to changes in messaging content by experienced sellers.

In Appendix Figure **??**, we also considered whether buyers or sellers are changing the length of messages they send over time. We find a noisy, but slight upward trend in the number of words used by sellers. For instance, after cleaning the messages, sellers send 8.44 words on average for the first five weeks of the sample compared to 8.58 words during the last five weeks. This pattern, again, is driven by more experienced sellers who send longer messages on average (as depicted in Table 4). Finally, in line with previous findings, the length of buyer messages is substantially noisier across time.[25]

To summarize, our foray into text analysis has taught us several things. First, the pattern is related to repeat play. We see this in both the comparison of buyers and sellers (Figure 8) as well as the comparison of less- to more-experienced sellers (Figure 9 panel (b) vs Figure 10 panel (b)). Second, the pattern is consistent with changes in the textual content over time, which visible for repeat players (sellers) in Figure 8 panel (b). Third and finally, that the rate of change is decreasing for repeat players, which we see in Figure 9 panel (b). Taken together, these stylized facts are consistent with a the hypothesis that repeat players are learning to use the communication feature. Next, we seek to bolster this with evidence that the endpoint of this process reflects less bargaining breakdown.

[25]In these figures, there is an uptick in the length of messages for sellers in weeks 6 and 7 and buyers in week 7. We suspect that this is due to current events. For instance, the day with the longest messages ocurred the day after the Germany National Soccer/Football team lost to France in the European Semifinal Championships (July 7, 2016).

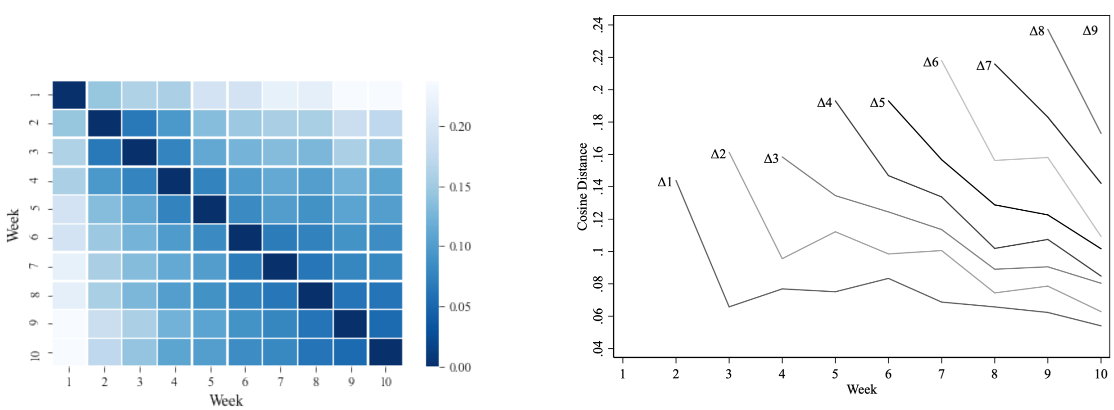
Figure 9: Off-Diagonals for Buyers and Sellers



(a) Buyer (b) Seller

Notes: This figure displays the change in cosine distance of bigram counts in the messages for different periods of time following May 25, 2016, separately for buyers, in Panel (a), and sellers, in Panel (b). Δx for week w indicates the cosine distance between week w and week $w - x$.

Figure 10: Convergence Among Sellers with Five or More Messages



(a) 5+ Messages (b) 5+ Messages

Notes: Panel (a) depicts the heat map representing the cosine distance of the bigram counts in the messages for each pairing of weeks following May 25, 2016 for sellers that sent five or more messages. Panel (b) plots the off-diagonals of the heat map from (a). Δx for week w indicates the cosine distance between week w and week $w - x$.

7 Message Experience Predicts Success

In Section 5.2.3, we found a time-variant effect of communication on the success rate of bargaining; namely, it took several weeks for the full effect on the success rate to manifest. Consistent with that, the results from Section 6.3 indicate that the content

Table 6: The Relationship Between Message Success and Cosine Similarity

	(1)	(2)	(3)
Sim(m, week 10)	0.0661*	0.0559*	0.0490
	(0.0226)	(0.0227)	(0.0428)
Message Length		0.0013*	0.0003
		(0.0002)	(0.0005)
N	101931	101931	62217
Controls	✓	✓	✓
Seller FE			✓

Notes: This table presents our results on message success and a measure of message experience. Sim(m, week 10) is the cosine similarity between a message and the set of week 10 messages excluding sellers who sent more than 21 messages. All models include our main set of controls: time trend; ln(ask price); category by condition fixed effects; day of week, precipitation, holiday dummies, and the temperature. Likewise, all models drop sellers sending more than 20 messages. Model (3) includes seller fixed effects. Robust standard errors are reported in parentheses and * denotes statistical significance at $\alpha = 0.05$.

of seller messages is changing over time, and in a convergent pattern. Here we close the loop by offering evidence that the endpoint of that convergence—what the sellers are learning—is actually bringing them greater success in bargaining.

We start by calculating the cosine similarity between each message and the aggregated set of bigram counts from the messages sent by sellers in week 10; in other words, how similar any given message is to the corpus of messages ten weeks after the change. We regress message success, a binary variable representing whether that seller's message ended in a sale, onto this cosine similarity measure.[26]

Table 6 presents our results. In all specifications we include the same set of controls as documented in Section 3.2. Finally, Models (2) and (3) also control for message length, as measured by the number of tokens in the processed message. Model (3) includes seller fixed effects.

In Model (1), we find a statistically significant effect that going from a completely orthogonal message to a message containing the set of week 10 messages is associated with an increase in success probability of 6.61 percentage points. In Model (2), we find that message length has a positive, but economically small, relationship with

[26] In this analysis we exclude a small set of sellers who send more than 20 messages. We do this because we believe that a) they are qualitatively different, professional sellers and b) they are overrepresented in message-level regressions. Overall, we are dropping 217 sellers that sent 21 or more messages; this is out of our original sample of 60, 076 sellers. In Appendix Section **??** we report variations, including using all sellers, which attenuates the results slightly, and the set of sellers that sent fewer than 11 messages, which does not qualitatively change the results.

the success rate. Finally, in Model (3) our result is no longer statistically significant when including seller fixed effects. We have only 3,134 sellers sending more than five messages, but fewer than 21; as a result, we lose a considerable amount of power in our estimates when restricting attention to within-seller variation. Still, the relationship between our cosine similarity measures and message success remains positive.

Altogether, Table 6 points to a positive correlation between the probability of message success and the similarity between the message and the set of week-10 seller messages. Notably, the point estimates here are similar in magnitude to the point-estimates from Table 3, where we found that among our treatment group, interactions that involved messages were 7.73 percentage points more likely to end in success.

A limitation of our approach is that we have not been able to address the direct mechanism in which messages contribute to higher rates of bargaining success. While we cannot provide causal evidence on this front, we borrow further from the literature on text analysis to offer suggestive evidence in the Appendix. In Appendix Section **??** we implement a distributed multinomial model from Taddy (2015); this helps us characterize what more experienced sellers are more likely to say. Table **??** presents the bigrams, both in English and the original German, that are the most and least predictive of seller experience, conditional on a set of controls. In addition, we use the bigram coefficients estimated from our model to compute message experience scores; here, a higher score indicates a message associated with higher levels of seller experience. Table **??** then shows the messages with the highest experience scores, while Table **??** includes the messages with the lowest scores.

Among the messages that are most correlated with experience, we see an emphasis on costs, especially those which may not be salient to buyers, such as the eBay commission.[27] This is consistent with prior work highlighting the effectiveness of cost rationales originally raised by Kahneman et al. (1986), and re-emphasized in bargaining settings by Lee and Ames (2017) and Bhattacharya and Dugar (2023). The latter work contrasts cost rationales with disparagement rationales, showing the latter to be ineffective. With respect to this, our focus on seller messages obviates disparagement, however we do observe that, among those messages least predictive of experience, arguments about the quality of the product are common. We also see

[27]The price that the buyer pays is not received by the seller. Instead, the platform takes a "final value fee" in the range of 10%. Buyers who have never been on the other side of the marketplace might easily be unaware of this fact.

that where those most correlated with experience are restrained, polite, and precise, those less so tend to use more effusive language, which is reminiscent of the finding, by Jeong et al. (2019), that warm language may be counterproductive in negotiations.

8 Discussion

We exploited a natural experiment in the availability of text messaging in bargaining to study the role of cheap-talk communication in avoiding bargaining breakdown. We found a statistically and economically significant effect: bargainers were on average eight percentage points more likely to transact, implying a fourteen percent reduction in the likelihood of bargaining breakdown among bargainers who used messages (against an average success rate of forty-four percent). Furthermore, using text analysis, we provide evidence of repeat-players learning how to communicate more effectively.

Our results encourage efforts to develop and test models that incorporate non-standard assumptions to understand the role of communication (Miettinen, 2013; Dufwenberg et al., 2017). There may be at least two mechanisms at play. One is that communication changes preferences in a way that encourages altruism. This is equivalent to increasing the buyer's willingness to pay and reduce the seller's willingness to accept. The empirical implication of this is to increase the likelihood of successful outcomes, which is consistent with our results.

A second mechanism may be the impact of communication on beliefs. For example, Feinberg and Skrzypacz (2005) show that heterogeneity in higher order beliefs disrupts one of the most celebrated, yet unrealistic results in dynamic bargaining models: that delay in bargaining cannot occur when offers can be made very frequently. It is possible, therefore, that communication may impact beliefs in ways that increase the likelihood of success. Indeed, in Table **??** we see some evidence, albeit not statistically significant, that both successful and unsuccessful bargaining interactions conclude in fewer rounds when communication is used.

We hope that our findings on communication in bargaining will encourage continued work. A promising direction that we have broached but not exhausted is to use developments in natural language processing and text analysis to study the content of what people say, above and beyond the fact that they say something. These questions have already been raised in experimental settings (see Lee and Ames (2017) for a recent contribution to, and summary of, this literature in social psychology). The availability of data from online bargaining marketplaces is an opportunity to see them in the field and directly link them to economic outcomes.

As Crawford (1982) writes, "the potential welfare gains from improving the efficiency of bargaining outcomes are enormous, perhaps even greater than those that would result from a better understanding of the effects of macroeconomic policy." Understanding how to improve bargaining outcomes is among the most central questions in economics and management, and also among the questions on which we have made the least progress, in part due to a lack of empirical work using data from real market interactions. Our paper tries to fill some of that gap, and our results highlight the importance of communication in bargaining. As Farrell and Gibbons (1989) write, "the economic importance of costless, nonverifiable, informal communication is much greater than its role in the literature suggests." Our results document empirically that this remains true today.

References

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

Aryal, Gaurab, Federico Ciliberto, and Benjamin T. Leyden, "Coordinated Capacity Reductions and Public Communication in the Airline Industry," 2021. forthcoming, *Review of Economic Studies*.

Autor, David, "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics*, 2003, *21* (1), 1–42.

Backus, Matthew, Tom Blake, and Steven Tadelis, "On the Empirical Content of Cheap-Talk Signaling: An Application to Bargaining," *The Journal of Political Economy*, 2019, *127* (4), 1599–1628.

_ , _ , Bradley Larsen, and Steven Tadelis, "Sequential Bargaining in the Field: Evidence from Millions of Online Bargaining Interactions," *Quarterly Journal of Economics*, 2020, *135* (3), 1319–1361.

Belkin, Liuba Y, Terri R Kurtzberg, and Charles E Naquin, "Signaling Dominance in Online Negotiations: The Role of Affective Tone," *Negotiation and Conflict Management Research*, 2013, *6* (4), 285–304.

Bhattacharya, Haimanti and Subhasish Dugar, "Undervaluation versus unaffordability as negotiation tactics: Evidence from a field experiment," *Journal of Economic Psychology*, 2023, *96*, 102609.

Bird, Steve, Edward Loper, and Ewan Klein, in "Natural Language Processing with Python," O'Reilly Media Inc., 2009.

Bochet, Olivier, Manshu Khanna, and Simon Siegenthaler, "Beyond Dividing the Pie: Multi-Issue Bargaining in the Laboratory," *The Review of Economic Studies*, 2023, *forthcoming.*

Bolton, Gary E. and Axel Ockenfels, "Does laboratory trading mirror behavior in real world markets? Fair bargaining and competitive bidding on eBay," *Journal of Economic Behavior and Organization*, 2014, *97*, 143–154.

_ , Kalyan Chatterjee, and Kathleen L. McGinn, "How Communication Links Influence Coalition Bargaining: A Laboratory Investigation," *Management Science*, 2003, *49* (5), 583–598.

Byrne, David P., Leslie A. Martin, and Jia Sheen Nah, "Price Discrimination, Search, and Negotiation in an Oligopoly: A Field Experiment in Retail Electricity," *Quarterly Journal of Economics*, 2019. forthcoming.

Cabral, Luis and József Sákovics, "Must Sell," *Journal of Econoomics and Management Strategy*, 1995, *4* (1), 55–68.

Cai, Hongbin and Joseph Tao-Yi Wang, "Overcommunication in Strategic Information Transmission Games," *Games and Economic Behavior*, 2006, *56*, 7–36.

Card, David and Gordon Dahl, "Bargaining and the Role of Expert Agents: An Empirical Study of Final-Offer Arbitration," *The Review of Economics and Statistics*, 2012, *94* (1), 116–132.

Charness, Gary and Martin Dufwenberg, "Promises and Partnership," *Econometrica*, 2006, *74* (6), 1579–1601.

Chatterjee, Kalyan and William Samuelson, "Bargaining Under Incomplete Information," *Operations Research*, 1983, *31* (5), 835–851.

Clark, Robert and Jean-François Houde, "The Effect of Explicit Communication on Pricing: Evidence from the Collapse of a Gasoline Cartel," *The Journal of Industrial Economics*, 2014, *62* (2), 191–228.

Coffman, Lucas and Paul Niehaus, "Pathways of Persuasion," *Games and Economic Behavior*, 2020, *124*, 239–253.

Cooper, David J. and Kai-Uwe Kühn, "Commuication, Renegotiation, and the Scope for Collusion," *American Economic Journal: Microeconomics*, 2014, *6* (2), 247–278.

Crawford, Vincent P., "A Theory of Disagreemnt in Bargaining," *Econometrica*, 1982, *50* (3), 607–637.

— , "Explicit Communication and Bargaining Outcomes," *American Economic Review*, 1990, *80* (2), 213–219.

— **and Joel Sobel**, "Strategic Information Transmission," *Econometrica*, 1982, *50* (6), 1431–1451.

— **, Miguel A. Costa-Gomes, and Nagore Iriberri**, "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 2013, *51* (1), 5–62.

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes, "Just Starting Out: Learning and Price Competition in a New Market," *The American Economic Review*, 2018, *108* (3), 565–615.

Dufwenberg, Martin, Maroš Servátka, and Radovan Vadovič, "Honesty and Informal Agreements," *Games and Economic Behavior*, 2017, *102*, 269–285.

Ert, Eyal, Stephanie Creary, and Max H. Bazerman, "Cynicism in Negotiation: When Communication Increases Buyers' Skepticism," *Judgement and Decision Making*, 2014, *9* (3), 191–198.

Farrell, Joseph and Robert Gibbons, "Cheap Talk Can Matter in Bargaining," *Journal of Economic Theory*, 1989, *48* (1), 221–237.

Feinberg, Yossi and Andrzej Skrzypacz, "Uncertainty about uncertainty and delay in bargaining," *Econometrica*, 2005, *73* (1), 69–91.

Fréchette, Guillaume R., Alessandro Lizzeri, and Jacopo Perego, "Rules and Commitment in Communication: An Experimental Analysis," 2021. Working Paper.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, "Text as Data," *Journal of Economic Literature*, 2019, *57* (3), 535–574.

— **, Jesse M Shapiro, and Matt Taddy**, "Measuring Group Differences in High-dimensional Choices: Method and Application to Congressional Speech," *Econometrica*, 2019, *87* (4), 1307–1340.

Gneezy, Uri, "Deception: The Role of Consequences," *The American Economic Review*, 2005, *95* (1), 384–394.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani, "Mediation, Arbitration, and Negotiation," *Journal of Economic Theory*, 2009, *144* (4), 1397–1420.

Green, Etan and Barry Plunkett, "The Science of the Deal: Optimal Bargaining on eBay Using Deep Reinforcement Learning," *Proceedings of the 23rd ACM Conference on Economics and Computation (EC '22)*, 2022, pp. 1–27. https://dl.acm.org/doi/10.1145/3490486.3538373.

Harrington, Joseph E. and Lixin Ye, "Collusion through Coordination of Announcements," *The Journal of Industrial Economics*, 2019, *67* (2), 209–241.

Hine, Michael J, Steven A Murphy, Michael Weber, and Gregory Kersten, "The Role of Emotion and Language in Dyadic E-Negotiations," *Group Decision and Negotiation*, 2009, *18* (3), 193–211.

Jeong, Martha, Julia Minson, Michael Yeomans, and Francesca Gino, "Communicating with Warmth in Distributive Negotiations is Surprisingly Counterproductive," *Management Science*, 2019, *65* (12), 5813–5837.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, "Fairness as a Constraint on Profit Seeking: Entitlemenets in the Market," *American Economic Review*, 1986, *76* (4), 728–741.

Karagözoğlu, Emin, "On "Going Unstructured" in Bargaining Experiments," in J-F. Laslier, H. Moulin, R. Sanver, and W. S. Swicker, eds., *Studies in Economic Design*, The Future of Economic Design, Springer, 2019, chapter 40, pp. 295–304.

Kopelman, Shirli, Ashleigh Shelby Rosette, and Leigh Thompson, "The Three Faces of Eve: Strategic Displays of Positive, Negative, and Neutral Emotions in Negotiations," *Organizational Behavior and Human Decision Processes*, 2006, *99* (1), 81–101.

Lee, Alice J. and Daniel R. Ames, ""I Can't Pay More" Versus "It's not Worth More": Divergent Effects of Constraint and Disparagement Rationales in Negotiations," *Organizational Behavior and Human Decision Processes*, 2017, *141*, 16–28.

Miettinen, Topi, "Promises and Conventions – An Approach to Pre-Play Agreements," *Games and Economic Behavior*, 2013, *80*, 68–84.

Moore, Don A., Terri R. Kurtzberg, Leigh L. Thompson, and Michael W. Morris, "Long and Short Routes to Success in Electronically Mediated Negotiations: Group Affiliations and Good Vibrations," *Organizational Behavior and Human Decision Processes*, 1999, *77* (1), 22–43.

Myerson, Roger B and Mark A Satterthwaite, "Efficient mechanisms for bilateral trading," *Journal of economic theory*, 1983, *29* (2), 265–281.

Oster, Emily, "Unobservable Selection and Coeffiicent Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Radner, Roy and Andrew Schotter, "The Sealed-Bid Mehanism: An Experimental Study," *Journal of Economic Theory*, 1989, *48* (1), 179–220.

Rubinstein, Ariel, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 1982, *50* (1), 97–110.

Taddy, Matt, "Distributed Multinomial Regression," *The Annals of Applied Statistics*, 2015, *9* (3), 1394–1414.

Valley, Kathleen, Leigh Thompson, Robert Gibbons, and Max Bazerman, "How Communication Improves Efficiency in Bargaining Games," *Games and Economic Behavior*, 2002, *38* (1), 127–155.