Price and Prejudice: Gender Discrimination in Online Marketplaces*

Sher Afghan Asad† Husnain Fateh Ahmad‡ Hadia Majid§

July 10, 2024

PRELIMINARY DRAFT - PLEASE DO NOT SHARE OR CITE

CLICK HERE FOR THE LATEST VERSION

Abstract

We investigate gender discrimination in an online marketplace in Pakistan. Employing buyer profiles that signal gender, we experimentally engage in transactions with sellers on the platform. While there is no significant discrimination in pricing and other economic variables, non-price discrimination, potentially bordering on harassment, persists. Female buyers are significantly more likely to receive unsolicited messages and friend requests from sellers post-transaction. The linguistic analysis shows male sellers displaying greater verbosity and informality towards female buyers. The paper highlights that while online marketplaces may lower the cost of accessing markets, they may come at the cost of harassment and unwelcome advances.

Keywords: discrimination, online-marketplaces, taste-based discrimination, statistical discrimination, harassment, language-processing.

JEL Codes: J71, D91, C93

*We thank participants at the brown bag seminar at Lahore University of Management Sciences, Pathways to Development Conference, 2022 in Lahore, Pakistan, and Advances in Field Experiment Conference 2023 in Chicago, United States for their helpful comments. We are also grateful to Muhammad Saad Siddique, Sidra Mazhar, Maham Rasheed, Eesha Nadeem, Syed Haris Ali, Taimoor Naeem, Humza Malik, and Amara Ashfaq for their excellent research assistance. We acknowledge partial funding from the Lahore University of Management Sciences and the Sewanee: University of the South. This research was approved by the Lahore University of Management Sciences IRB and registered on the AEA RCT registry, ref. AEARCTR-0009049. In preparing this manuscript, we utilized ChatGPT, a language generation model developed by OpenAI, for the initial draft of a few sections of the paper and to improve the clarity of the write-up.

†Lahore University of Management Sciences (sherafghan@lums.edu.pk)

‡Sewanee: the University of the South (hfahmad@sewanee.edu)

§Lahore University of Management Sciences (hadia.majid@lums.edu.pk)

1 Introduction

Gender discrimination is an enduring and pressing issue that significantly impedes individual advancement and overall socioeconomic development (Duflo, 2012). In recent years, online marketplaces have emerged as potential catalysts for change in this context. Through their digital infrastructure, these platforms hold the potential to grant members of marginalized groups access to a broader market, effectively circumventing many of the offline limitations they may routinely encounter. Despite their potential, there is limited empirical evidence concerning online marketplaces' role in alleviating gender discrimination. We seek to address this knowledge gap by utilizing experimental methods to document and study gender discrimination in an online marketplace in Pakistan.

To investigate, we deploy a well-powered experiment to audit seller behavior on an online marketplace (Facebook marketplace, henceforth simply marketplace) in Pakistan. Based on a repeated weekly census of listings, we contact sellers who regularly sell on the platform through buyer profiles that unambiguously signal gender without revealing caste, ethnicity, or other economic markers. Each seller is contacted twice, once by each gender, following carefully crafted and predetermined bargaining scripts. We record and analyze both traditional economic variables, such as prices and product quality, and also track non-economic variables that may capture discrimination and non-economic costs. Furthermore, we track post-transaction unsolicited attempts from sellers to communicate, such as messages, phone calls, and and friend requests. We also perform linguistic analysis to determine if there are any differences in sellers' interactions with buyers.

We find no evidence of discrimination on economic variables but find that non-economic costs, namely unsolicited messages and friend requests, are significantly higher for female buyers. Our results highlight that while online marketplaces may have reduced some barriers to female participation, non-economic hurdles persist. Our research highlights that while online marketplaces may lower the cost of accessing previously hard-to-access markets, they come at the cost of the same old threat of sexual harassment and unwelcome advances.

The paper makes several contributions to the literature on gender discrimination. It contributes most directly to the extensive experimental literature on gender discrimination (see Croson and Gneezy (2009) and Neumark (2018) for surveys of the literature). While there is abundant evidence of discrimination against women, there is debate on the nature and mechanism behind observed discrimination. Our research contributes to this literature by highlighting the non-pecuniary nature of discrimination in an online marketplace setting of a developing country where sexual harassment is widespread (Duflo, 2012). To our knowledge, we are the first to systematically investigate discrimination against women in a product market setting.

Our work contributes most directly to the emerging literature on discrimination on online platforms. There is growing evidence of racial discrimination in auction prices (Ayres et al., 2015), wages (Hannák et al., 2017) and rents (Edelman et al., 2017). Asad et al. (2023) also find evidence of discrimination by workers towards black managers on Amazon's online platform, Mechanical Turk. In the developing country context, Chen (2024) finds discrimination against female physicians on an online health platform in China. We extend the literature by studying both gendered price discrimination in an online product marketplace and by extending discrimination beyond traditional economic variables.

The paper is more narrowly related to field studies on discrimination in bargaining. Examples include gender and racial differences in bargaining over car prices (Ayres and Siegelman, 1995), race, age, and gender differences in bargaining over sports cards (List, 2004), taxi fares (Castillo et al., 2013; Michelitch, 2015), medicine prices(Fitzpatrick, 2017), and more recently, perceived-income-based price discrimination in retail electricity markets in Australia (Byrne et al., 2022).

Another contribution of our research lies in the experimental control of the bargaining process. Extensive literature in economics documents women's inability to negotiate better deals (see Exley et al. (2020) for discussion of related issues) as the reason for poorer outcomes for women. In our design, we control the bargaining strategy, allowing us to eliminate such concerns and explain any observed biases as being solely driven by sellers.

Our methodological contribution stems from the unique bargaining design in which we send repeated signals of buyers' valuation to sellers, which would also help us determine whether outcome differences are driven by differences in seller perceptions of buyer values or due to consistent gendered taste biases that stay stable across the various stages of bargaining. We can disentangle belief-based discrimination (Phelps, 1972) from taste-based discrimination (Becker, 1957). However, we note that taste-based discrimination in our setting can go in either direction; for example, it may be disadvantageous for women if driven by in-group bias (see, for example, Chen and Li (2009) and Hedegaard and Tyran (2018)); or may work in their favor if it induces a preference for negotiating with female buyers.[1]

2 Experimental Design

We conduct our experiment on an online marketplace platform in Pakistan (Facebook Marketplace), using fictitious buyer profiles and following very strict bargaining scripts. In this section, we provide details on our subjects (sellers), our treatments (buyers and bargaining scripts), and the flow of the experimental procedure.

2.1 Sellers Selection

We selected sellers who regularly posted products on the Facebook marketplace as they were likely to use the platform as a business. Specifically, we restrict our sample to sellers who had posted at least 50 times on the marketplace using their profiles before the census day. We also ruled out sellers of used goods, as we wanted to ensure sellers had enough stock to fulfill both of our orders.

[1]The latter is similar to what is documented in (Castillo et al., 2013) where tax drivers exhibit a preference for women riders and quote lower fares to them in exchange for the company of women riders.

2.2 Product Selection

The marketplace had a variety of products under various categories. To get a sense of products listed on the marketplace, we conducted a census of sellers who deliver to Lahore, Pakistan, on January 5th, 2022. Appendix Table B2 presents the summary of posts in various categories. There were 31,120 posts on the census date, which Facebook automatically categorized into 177 generic categories. There was large variation in the kind and price of products within and across categories. Due to budgetary constraints and to allow for a wide range of products, we restricted attention to the top ten most frequently listed categories for which the 75th percentile of posted price was less than PKR 3,500 (≈20 USD). This yielded the (auto-generated) categories of arts, health, home-decor, bags, shoes, men's, women's, kids-clothing, bedding, and portable-audio-video. Table B3 briefly describes each of these categories.[2] Some categories (such as clothing and shoes) have products that come in various sizes, designs, or colors; for these products, we reached out to sellers for the product that was listed first.[3] We exclude products that require customization, such as engraving a name, etc. A wide range of categories allowed us to examine discrimination across a broader spectrum and is also more representative of overall discrimination on the platform. After restricting to these categories, for budgetary reasons, we imposed one final restriction: we only contacted a seller if the posted price was below PKR 2,000. If posts in selected categories were listed without price information, we contacted sellers but stopped bargaining if the first quoted price was above PKR 2,000.

2.3 Buyer Profiles

We created buyer profiles such that the profile name was an unambiguous signal of gender and did not contain any other ethnic, caste, or economic markers. To arrive at the representative list of names, we relied on publicly available tax data published by Pakistan's central tax authority, the Federal Board of Revenue (FBR). The latest published directory of 2018 contains information on more than 2.7 million taxpayers. We tabulate the most frequent first and last names for both genders and then randomly assigned first names to last names. We exclude caste, sect, or ethnic indicators (such as Khan, Chaudhry, Sheikh, Rao, etc.) from names to avoid any potential contamination. Table B4 lists the final names selected for our buyer profiles. To avoid suspicion, we do not approach the same seller with profiles with the same last or first name; instead, each seller is contacted using two entirely different names.

2.4 Bargaining and Ordering Scripts

Once a product and a seller are selected, bargaining starts with a first message from a randomly selected gendered profile, that asks for the price of the posted item (irrespective of the existence of a poster price). Once the seller quotes an offer, the buyer responds by asking for a discount

[2] There is obvious overlap between some categories used by Facebook; for example, women's shoes are likely to be categorized under both the categories of shoes and women.

[3] On some occasions, the designs/options are shared by a seller over messages; in these cases, we continued to select the first presented option.

without giving any counteroffer.[4] There are three possible seller replies to this: first, the seller may ask the buyer to quote a price; second, the seller may quote a discounted price; or third, the seller may refuse to give any discount. In all three cases, we exhaust the bargaining process and continue to nudge the seller to give discounts, if possible. Either the seller will concede to providing a discount or refuse. Upon agreement of the final price, which may be discounted or not, negotiations end, and the buyer moves the discussion toward placing the order.

We employed sequential bargaining, as our design was motivated by an extension of sequential interactions in Bohren et al. (2019) to our setting. The design allows us to disaggregate any observed price discrimination into its source.[5] As per our model, differences in the first quoted price would reflect both statistical and taste-based discrimination, while differences in the final price would only be driven by the latter.

Figures C1 and C2 detail both the scripts used and illustrate their flow. The order of scripts were assigned to each seller randomly.

The ordering process starts after the buyer and seller have agreed on a price. Each buyer for a seller is randomly assigned to one of the two ordering scripts (shown in Figures C3 and C4). The ordering scripts begin by confirming the mode of payment, and we only proceed with an order if cash on delivery is acceptable.[6] It is during the ordering stage that the buyer shares their contact details, including the address for delivery. Each seller is assigned to two addresses, one for each buyer, in random order.

2.5 Post-Delivery

After an order had been delivered, we downloaded all conversations with the seller on Facebook, WhatsApp, and text. We also recorded the entire call log history with each seller. Upon delivery, we inspected items thoroughly to see any differences in quality.

2.6 Sample Size

Our choice to use within-subject design is motivated by concerns to maximize power. Bellemare et al. (2014) show that a between-subject design requires between 4 to 8 times more subjects than a within-subject design to reach an acceptable 80 percent level of statistical power. Similarly, List et al. (2011) shows that within-subject design dramatically reduces the variance of unobservables, increasing the precision of the estimated average treatment effects. A disadvantage of within-subjects design is the possibility of order effects, i.e., subjects' behavior may depend on the order of the treatment. We address the latter concern by randomizing the order in which each gender contacted the seller.

We conducted a brief pilot of the design before the experiment's launch to determine the required sample size and get a sense of the minimum detectable effect. Based on the pilot

[4] As mentioned earlier, the negotiation stops if the quoted price is above PKR 2,000.

[5] For the full model, see Appendix Section A.

[6] Given the low penetration of financial products in Pakistan, cash on delivery is the most common payment method for online shopping.

results, we designed our experiment to detect a difference in prices of 0.25 standard deviations.[7] Given our within-subject design, we can detect this effect by contacting 128 sellers twice (256 contacts).[8] We also identified multiple primary variables of interest (5) and so planned to adjust for multiple hypothesis testing, using Anderson sharpened q-values (Anderson, 2008).[9]

Multiple hypothesis testing requires us to further adjust our sample size (List et al., 2019). Even though we use a false discovery rate adjustment (q-values), for power calculations, we use the more conservative Bonferroni correction , where the researcher takes the threshold value for the probability of a type one error and divides it by the number of hypotheses tested. We have 5 primary hypotheses (detailed later) and so to account for multiple hypothesis testing, we change the probability of type-I error from the conventional 0.05 to 0.01. This increases the required number of sellers to 191. We rounded that up and aimed to negotiate prices with 200 randomly selected sellers twice (400 matched purchases). As explained in Section 3, we were able to negotiate and agree on prices with 224 sellers each, which required contacting 619 sellers, further increasing power.

2.7 Experiment Flow

The experiment flows as follows;

1. At the beginning of each week, we conducted a census of posts that meet our criteria, recording basic information of each post, such as product name, category, and posted price.

2. We then randomly selected one of the ten categories.

3. From the selected category, we randomly drew a post and checked if the selected post was posted by a seller that meets the criteria specified in sub-section 2.1. If the post had already been drawn or the seller had been selected for another post, we redrew another post.

4. Once a post was selected, we randomly selected a gender with which to contact the seller and then selected a random profile from the list of four profiles of that gender.

5. Once the profile was selected, we negotiated a bargaining script that was also randomly selected.

6. After the conclusion of bargaining, we ordered the item using a randomly selected ordering script.

7. After at least 24 hours, we contact the seller again with a profile from the other gender. We re-randomize without replacement from steps 4 to 6.

[7]Our pilot had an average difference in prices of ≈20 PKR with a standard deviation of the difference at ≈85. This was admittedly based on very few observations and only suggests an effect size.

[8]We use Stata's *power pairedmeans* command to calculate the sample size.

[9]For other experimental papers that deploy sharpened q-values, seeBanerjee et al. (2015), Bryan et al. (2021) and Ahmad et al. (2024).

We repeated steps 2 to 7 to get more observations weekly and continued the process until the target number of observations was reached.

3 Data and analysis

3.1 Sample

We ran the experiment for about seven and a half months, from March 25th, 2022, to November 8th, 2022. We initiated bargaining with 670 sellers for 1,236 bargaining attempts during this period. A total of 172 instances were identified in which bargaining was initiated from one gender but not the other. These discrepancies occurred when the seller removed their post after our first contact, or due to an early coding error that assigned buyers to sellers incorrectly, or finally due to instances where one of the research assistants failed to initiate/continue the conversation with a corresponding seller. As these observations violated the experimental protocol, lacked data on both genders and impeded the feasibility of within-subject comparisons, we excluded them from the analysis. This left us with 1,064 bargaining attempts with 532 sellers.[10]

Table 1 provides an overview of the collected data disaggregated by treatment. Of the total 1,064 attempts, approximately 39 percent culminated in successful order completions. The remaining instances where orders could not be finalized were attributed to various factors, including seller unresponsiveness (24 percent), product unavailability (13 percent), quoted prices exceeding the established threshold of PKR 2,000 (13 percent), and sellers requiring advance payments (10 percent). Furthermore, not all orders placed resulted in successful deliveries, with only 77 percent of orders ultimately being fulfilled, with a higher proportion among male buyers. On price, we find that approximately 55 percent of negotiations resulted in an agreement on the final price. However, for within-subject comparisons, we would need prices for both genders; this was reached for only 42 percent of sellers. Finally we find that products are almost equally represented across the selected (Facebook ascribed) categories.

Table 2 presents information on sellers. Overall, 63 percent of our sellers were male. On average, male sellers have more friends and followers than female sellers. Most sellers (61 percent) used their personal Facebook accounts to sell items on the marketplace, though female sellers were relatively less likely to use their personal accounts for marketplace activities. Most seller profiles are public (more so for male sellers than females).

3.2 Primary variables and analysis

We identified and pre-registered five primary variables for the main analysis. They include quoted prices (first and final agreed prices), product quality, unsolicited communications, and language/sentiment analysis.

[10]As per our pre-registration, the target was to collect data on 400 transactions from 200 sellers. Since every contact with a seller did not lead to the transaction's completion (defined as an agreement on the final price or delivery of the item), we continued contacting sellers until we reached the target number of transactions. Our final sample has a slightly higher number of contacts where bargaining led to agreement on the final price for both genders (446) and a much lower number of attempts where bargaining led to the item's delivery (321).

Regarding quoted prices, our theoretical model, an application of Bohren et al. (2019) to our setting, allows us to disentangle any observed discrimination to either statistical or taste-based discrimination.[11] Of interest is the result that any observed differences in the first price the seller quotes are driven by both statistical and taste-based discrimination, while differences in the final price should capture only the latter, as the identical signals (bargaining script) should update the seller's beliefs about the buyer's willingness-to-pay in a similar direction. As such, we use the first and final prices (including delivery fees) as two variables of interest. As mentioned, we use a balanced within-subject design for our analysis, so we utilize a linear mixed-effects model with clustered standard errors at the seller level (following de Chaisemartin and Ramirez-Cuellar (2024)).

We also identify product quality as a variable of interest. Sellers may discriminate between buyers by varying the product quality while keeping the price constant. To capture this, we visually inspect all products and identify whether, for a given seller, the products received are identical or if the quality is higher for one gender over another.

For non-economic variables, we identify two variables of interest: unsolicited communication attempts and seller tone/sentiment. For unsolicited communication attempts, we record the number of calls, messages, and friend requests each buyer receives daily across all platforms (including phone calls, SMS, WhatsApp, and Facebook). We test for differences using a Poisson regression of count variables on gender (Cameron and Trivedi, 2022).

Finally, we investigate if there are any differences in the way sellers communicate with buyers. To do this, we hired a team of local undergraduates to read and annotate chats with sellers and record their perceptions of buyer and seller language. We use their responses to create a simple additive index for seller interaction and test for differences across genders using a within-subject design with additional annotator fixed effects.

Although our primary variables of interest are pre-registered, we still adjust for multiple hypothesis testing (MHT) using false discovery rate (FDR) adjustments. Particularly, for all primary variables, we report Anderson Sharpened q-values (Anderson, 2008).[12]

Finally, at the design stage, we specified a within-subject design, so we restricted attention to a balanced sample for all our primary analyses. We report results from our full, unbalanced samples in the appendix as robustness checks (Appendix Table B1).

3.3 Exploratory analysis

Beyond our pre-registered primary variables, we collected a rich set of data as part of the project. We conduct exploratory analysis[13] on this as they provide insights into the mechanisms driving

[11] The full model, with all assumptions and derivations, is presented in Appendix A.

[12] FDR methods require the researcher to specify the expectation of a false rejection (or false discovery), referred to as the "q-value". Anderson's sharpened q-values sidestep this pre-specification, and the methodology conducts the step-down FDR method (Benjamini et al., 2006) for all possible q-values and reports the lowest q-value for which a particular hypothesis would be rejected. A sharpened q-value then is similar to the more common p-value but for the expectation of false rejection. So, for example, a sharpened q-value of 0.1 suggests a hypothesis would be rejected if we set the threshold/tolerance for a false discovery at 10% in expectation.

[13] Some of these were also identified as secondary variables in our pre-registration.

our main results. In particular, for prices, we expand our analysis to all stages of the bargaining process. Similarly, for non-price variables, we disaggregate them into their components to identify the main drivers, and we extend our exploratory analysis to analyze heterogeneous effects by observable seller characteristics. Beyond simple disaggregation, for language analysis, we also deploy an alternative method of coding, namely artificial intelligence-based natural language processing. All exploratory analysis expands on the primary analyses and relies on the same statistical tests. As this analysis is exploratory, we do not account for it in our multiple hypothesis testing adjustments.

4 Results and discussion

4.1 Primary results

Results from our primary analysis reveal no evidence of discrimination on traditional economic variables. Table 3 columns 1 to 3 show no evidence of price or quality discrimination. On average, our buyers received similar prices from sellers, suggesting that online markets may successfully reduce discrimination. However, we note that this may not indicate a lack of biases, as various sources of bias (statistical or taste-based) may cancel each other out. In particular, our model suggests that discrimination in the final price may be attributed solely to taste-based discrimination, and the in-group bias and preference for communicating with the opposite gender may indeed cancel each other out. Similarly, on average, we find that buyers receive the same quality of goods, with our data showing most matched pairs receiving identical goods (90%), and when there are visual discrepancies, on average, both groups are equally likely to receive the higher quality goods.

We do find evidence of discrimination in non-economic realms. In particular, we find evidence that female buyers receive significantly more unsolicited messages daily than their male counterparts (Table 3 column 4), suggesting that intangible costs for women to access markets continue to persist online. Beyond the statistical results, the scale of difference is even more stark visually, as illustrated in Figure 1. Our results show significant discrimination in how female buyers are approached (potentially harassed) after participation in online marketplaces, highlighting the existence of non-price costs that women must (continue to) pay for participating in economic activity. Together with our results on price and quality, it appears that while online markets help lower price barriers, non-price barriers remain a major hurdle.

Finally, we find weak evidence of differences in seller interaction with our buyers. When analyzed, our annotators, on average, marked seller interactions with male buyers as more professional than those with female buyers (Table 3 column 5). While any information on gender (gendered language, etc.) was stripped from chats shared with annotators, as a robustness check, we also asked annotators to rank buyer interactions at the same time, and our results highlight how there were no differences in buyer interactions reported by our annotators (Appendix Table B1 column 4).

4.2 Exploratory findings

For each of our primary variables, we have access to a rich set of accompanying data to explore the dynamics and drivers of our main results. We note that as these are exploratory (and were not fully per-specified), we do not account for them in our multiple hypothesis testing, nor are we powered for null results.

Sequential prices We first investigate potential gender differences in the sequence of prices we observed during bargaining. For this analysis, we restrict our sample to observations from sellers where bargaining led to an agreement on a final price for both genders, thereby allowing within-subject comparisons. Figure 2 presents the raw quoted prices from sellers at various stages of negotiations. While the minor fluctuations around the 45-degree line indicate some variability, overall, we do not see significant price differentials at any bargaining stage. To test this, we employed the Wilcoxon signed-rank test (Wilcoxon, 1945) to evaluate the equality of price distributions for each matched pair. These results consistently indicate no overall statistically significant difference in prices.

Furthermore, we extend our primary analysis of prices using a linear mixed-effects model with clustered standard errors at the seller level to estimate the price differences at each negotiation stage. We also conduct sub-sample analysis by seller gender.[14] As presented in Panel A of Table 4, initially, female buyers receive slightly more favorable first-price offers, though this difference is statistically indistinguishable from zero. However, as the bargaining process unfolds, this distinction diminishes and eventually becomes negligible for the final agreed-upon prices. It is worth noting that sellers treat female buyers favorably regarding delivery charges, sometimes waiving them or handling deliveries themselves. This practice appears to tip the scale slightly in favor of female buyers, though these price differences remain statistically similar and economically small. In light of these results, we conclude that, on average, there are no systematic price differences between male and female buyers at any stage of the bargaining process. Panels B and C report results for female and male sellers, and once again, we confirm that there is no significant difference in prices for male and female buyers across seller gender.

Finally, in connection with the conceptual framework in Appendix A, there seems to be no firm evidence of taste-based or statistical discrimination in prices, though the latter may be because opposing taste preferences may, on average, cancel each other out. However, according to the model, the observance of favorable final prices for female buyers (though statistically insignificant) indicates a preference among sellers for interacting with female sellers.

Non-price bargaining outcomes We also explore differences in bargaining outcomes beyond those related to price. Table 5 reports differences in outcomes related to bargaining, both before and after the completion of the process. Panel A in Table 5 shows no statistical difference

[14]We note that while we do not correct for MHT here, in examining the evolution of quoted prices, our dependent variables, ranging from first to final prices, are inherently correlated. The literature suggests that MHT adjustments are less critical when dealing with highly correlated tests that explore different dimensions of the same underlying process (Anderson, 2008; List et al., 2019)

in outcomes such as the probability of withdrawing from bargaining (column 1), the number of stages it takes to agree on a price (column 2), and the probability of requiring advance payment (proxy for trust in buyer to pay at the time of order delivery) before the delivery of order (column 3). However, interestingly, we find that the sellers are significantly more likely to complete the order for female buyers than male buyers (column 3), which is primarily driven by female sellers as shown in column 4 of Panel C.[15]

Similarly, we do not observe any statistically significant differences in the probability of the order being delivered conditional on order placement (column 5), the time it takes, in days, to deliver the product conditional on the delivery of the order (column 6) and the probability the order is same as ordered by the buyer (column 7). Panels B and C confirm no differences in these outcomes by the seller's gender.

Unsolicited communication attempts We break down our main results for unsolicited communication attempts by mode of communication and present our results in Table 6. We extend our aggregate analysis using exponentiated coefficients from Poisson regressions of count variables on gender (Cameron and Trivedi, 2022) and report results by mode of communication. We find a significantly higher incidence of post-transaction messages to females; specifically, on average, sellers send about 1.24 messages to a female buyer for every message to a male buyer. These messages are typically marketing messages, confirming order delivery, requesting to review the order, etc. In addition, female buyers received 1.4 phone calls and 1.6 messages for every call or message received by the male buyer. Similarly, the incidence ratio of receiving unsolicited messages on Facebook and WhatsApp is about nine times higher than that of male buyers. Female buyers also receive a disproportionately higher share of friend requests on Facebook than their male counterparts. The coefficient for friend requests is substantial because no male buyer ever received a friend request, and all requests were sent to female buyers.

It is important to note that we revealed little to no information about the buyer to the sellers. Therefore, these differences in communication attempts are driven entirely by the buyer's gendered name. Had we included pictures or other information about gender, such as marital status, these attempts may have been even more severe. Therefore, we interpret these effects as a lower bound of harassment and expect that the incidence of harassment would be far higher in an uncontrolled environment.

Language and sentiment analysis We are able to analyze various aspects of sellers' conversation style such as verbosity, responsiveness, use of emoticons, use of honorifics such as sir or madam, use of casual lexicons given by colloquial use of words such as 'bro' and 'sis,' and request to talk on WhatsApp. We find that sellers, on average, are significantly more verbose when bargaining with female buyers, driven primarily by male sellers (Table 7). We also observe that male sellers are more likely to use informal lexicons with male buyers than with their female counterparts.

[15]An order is labeled "incomplete" if a seller becomes non-responsive, does not accept cash on delivery, or reports the item as out of stock.

In addition to the manual coding reported in our primary analysis, we perform linguistic analysis of sellers' responses using automated language processing. We employ OpenAI's GPT-4 (OpenAI et al., 2023) for nuanced sentiment and language analysis of conversations between buyers and sellers, where communication often blends Urdu (in Roman or traditional script) and English.[16] GPT-4, the latest iteration of the Generative Pre-trained Transformer models, stands out for its language understanding and generation capabilities. Its architecture is designed to handle diverse datasets, making it uniquely suited for analyzing the intricacies of mixed-language conversations in the local context (Baktash and Dawodi, 2023).[17] We leveraged GPT-4 to assess various aspects of seller communication, including politeness, clarity, formality, enthusiasm, friendliness, and assertiveness.For each trait, GPT-4 assigned values between 0 and 1, where a value closer to 1 indicated a more substantial presence of the trait.

Table 8 presents the results from a linear mixed-effects model with clustered standard errors at the seller level for various traits. We find that sellers were rated to be relatively more informal and more enthusiastic with female buyers. On other traits, such as politeness, clarity, friendliness, and assertiveness, we do not find differential treatment across genders, though we find some evidence of female sellers being slightly more friendly towards their in-group.

Heterogeneity across observable seller characteristics The average treatment effect can vary with the observable characteristics of sellers and products. We first examine how average treatment effects vary by product and seller-related aspects. Then, we explore how sellers' behavior, including bargaining-related outcomes and conversational features, correlated with their differential pricing behavior. These tests allow us to explore the characteristics likely driving the treatment effects and provide a deeper understanding of how the treatment effects vary by these factors.

In Table 9, we present heterogeneity results by product and seller characteristics for all sellers (columns 1 and 2) and separately for male (columns 3 and 4) and female (columns 5 and 6) sellers for the first quoted price and final agreed price (including delivery charges). We do not find evidence of differences in treatment effects by the gender orientation of the product, presence of religious content on sellers' profiles, account type (personal/business), account privacy (private/public), and personal photos on the profile. However, we do observe that sellers who publicly post their marital status as 'single' tend to quote significantly higher prices to females as compared to the prices they quote to male buyers. This, though, should be interpreted with caution since, for many profiles (75 percent), the marital status was not publicly posted.

[16] See Charness et al. (2023) for how generative AI is transforming scientific practices by aiding experiment design, implementation, and analysis.

[17] There is a burgeoning literature on text analysis of Roman Urdu using language processing programs, see, for example, Mehmood et al. (2019); Ghulam et al. (2019); Chandio et al. (2022); Mehmood et al. (2020) among other studies. This literature, however, is in the early stages of development, and the models are generally not applicable to our context.

5 Conclusion

Digital platforms offer a space for economic empowerment and have the potential to reduce biases that perpetuate inequality. We study the presence of gender discrimination in an online market in the developing country context and find that while there is limited evidence of price discrimination, challenges persist.

We find that male and female buyers face no differential treatment on economic variables such as prices and product quality. However, female buyers disproportionately receive significantly more unsolicited advances through messages, calls, and friend requests. These advances potentially border on harassment and are suggestive of a higher non-price cost for women's participation in the economy.

The paper has several limitations which may be explored further in future research. For example, this paper has focused on product prices at the bottom of the price distribution; discrimination likely plays out differently in high-ticket items such as cars, real estate, etc. Additionally, our only signal of gender was the name we used for buyer profiles, and we avoided using photographs or other information in profiles that could contaminate our results. Sellers would likely respond differently if profiles had pictures or additional information, and understanding the effects of such confounds would be interesting.

This paper opens various avenues for future research. To adequately realize the potential of online marketplaces in addressing gender discrimination, there needs to be an investigation into the supply side and the experiences women face when participating as entrepreneurs in these marketplaces (see, for example, Alhorr (2024) for how online marketplaces can help women entrepreneurs promote their businesses). This is a crucial margin to explore and may have far-reaching implications for the inclusivity of online marketplaces. In addition, this paper invites the exploration of policies or interventions that can address the challenges identified and ensure the full participation of women in the economy.

Figures

Figure 1: Average number of unsolicited communications daily by gender.

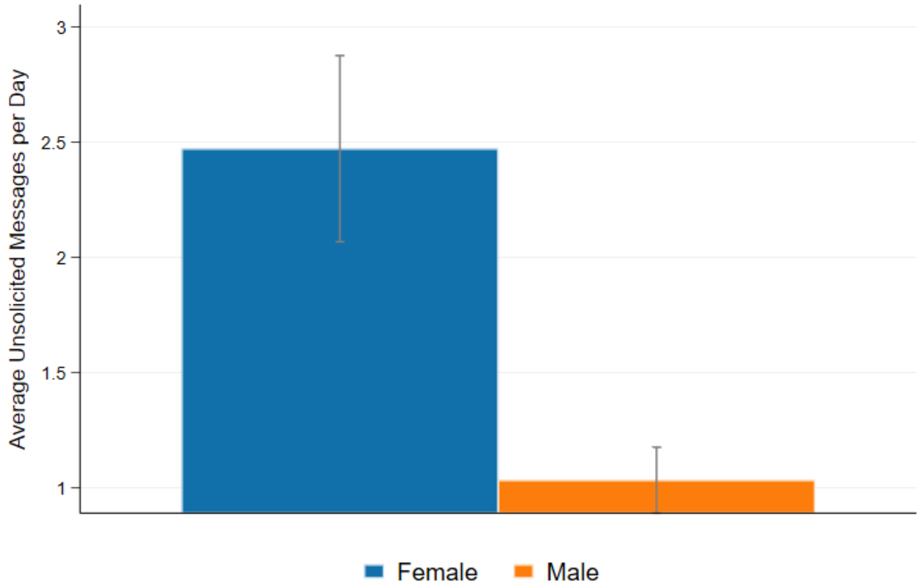
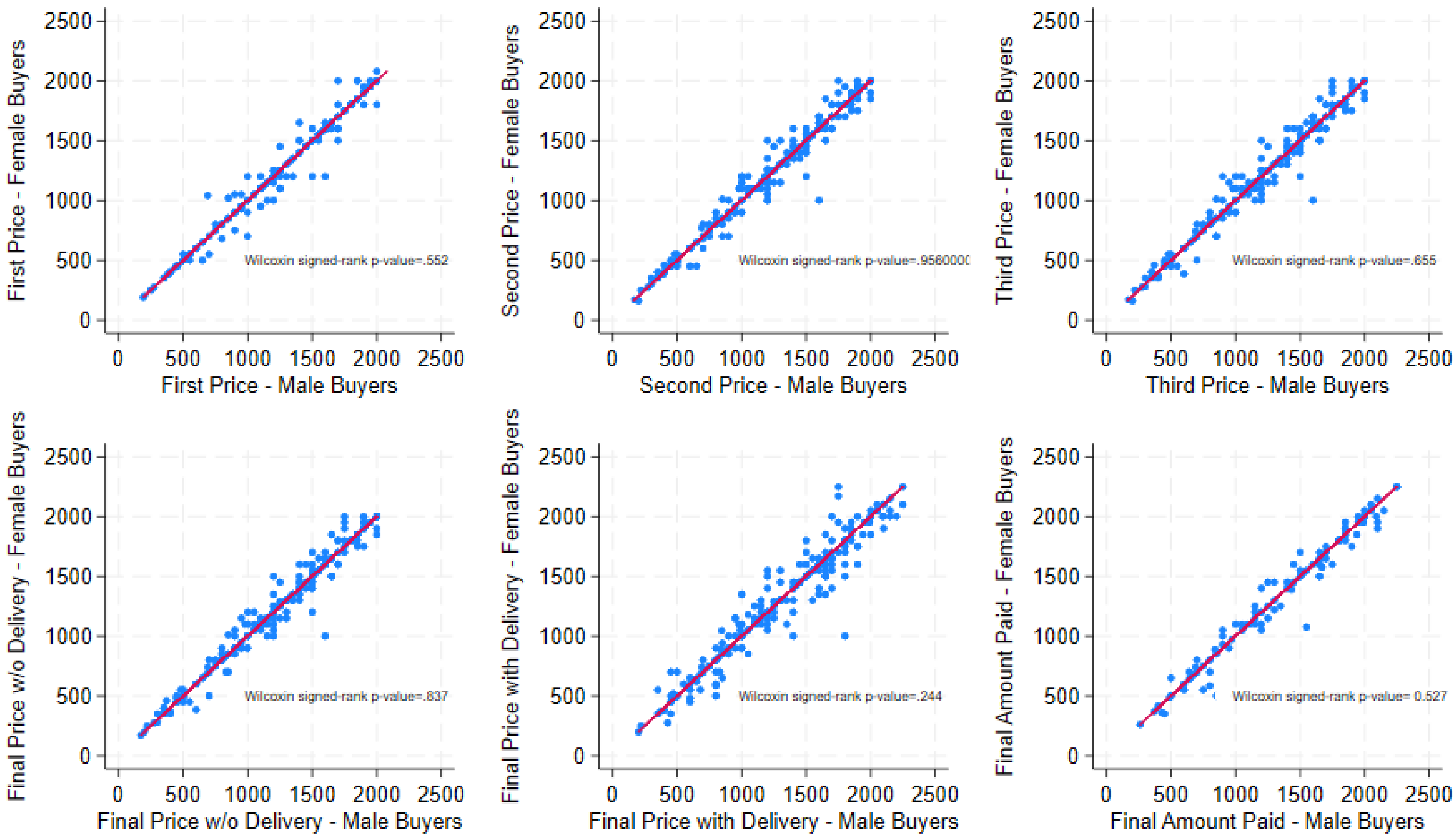


Figure 2: Comparison of Prices to Males and Females Buyers



Note: The figure presents the raw prices quoted to male (x-axis) versus female (y-axis) buyers. The 'First Price' corresponds to the first privately quoted price by the seller to a buyer at the start of the bargaining process. Similarly, the 'Second Price' and 'Third Price' reflect the prices quoted during the second and third stages of bargaining, while the 'Final Price' is the finally agreed price between the buyer and seller in response to the outlined bargaining process. 'Final Price w/ Delivery' add delivery charges, if any, to the 'Final Price.' 'Amount Paid' is the amount paid by the buyer at the time of the product's delivery. All prices are in PKR. Data in all panels is restricted to sellers where the sellers could agree on a final price for both genders.

Tables

Table 1: Summary of Bargaining Process.

	Buyer Gender		
	Female	Male	Total
	532 (50%)	532 (50%)	1064 (100%)
Bargaining outcome			
Order completed	223 (42%)	194 (36%)	417 (39%)
Seller stopped responding	120 (23%)	139 (26%)	259 (24%)
Item unavailable	66 (12%)	71 (13%)	137 (13%)
Quoted price > threshold price	67 (13%)	73 (14%)	140 (13%)
Seller required advance payment	56 (11%)	55 (10%)	111 (10%)
Delivery status			
Received	167 (75%)	154 (79%)	321 (77%)
Not Received	55 (25%)	40 (21%)	95 (23%)
Bargaining completed			
No	231 (43%)	253 (48%)	484 (45%)
Yes	301 (57%)	279 (52%)	580 (55%)
Bargaining completed for both genders			
No	309 (58%)	309 (58%)	618 (58%)
Yes	223 (42%)	223 (42%)	446 (42%)
Product Category			
arts	45 (8%)	45 (8%)	90 (8%)
bags	66 (12%)	66 (12%)	132 (12%)
bedding	50 (9%)	50 (9%)	100 (9%)
health	56 (11%)	56 (11%)	112 (11%)
home-decor	37 (7%)	37 (7%)	74 (7%)
kids-clothing	57 (11%)	57 (11%)	114 (11%)
mens	53 (10%)	53 (10%)	106 (10%)
portable-audio-video	47 (9%)	47 (9%)	94 (9%)
shoes	62 (12%)	62 (12%)	124 (12%)
womens	59 (11%)	59 (11%)	118 (11%)

Note: The table presents the frequency and percentage of various categorical outcomes related to bargaining, ordering, delivery, and products covered during the experiment. Product Categories refer to categories defined by Facebook marketplace.

Table 2: Summary of Sellers Data

	Seller Gender		
	Female	Male	Total
	198 (37%)	334 (63%)	532 (100%)
Number of Friends (mean)	255	473	392
Number of Followers (mean)	512	5336	3657
Business/Personal Account			
Personal	107 (54%)	218 (65%)	325 (61%)
Business	52 (26%)	74 (22%)	126 (24%)
Not Known	39 (20%)	42 (13%)	81 (15%)
Public/Private Profile			
Public	123 (62%)	239 (72%)	362 (68%)
Private	49 (25%)	62 (19%)	111 (21%)
Not Known	26 (13%)	33 (10%)	59 (11%)

Note: The table presents the summary statistics (mean, proportions, and percentages) about the sellers included in the study sample. "Business/Personal Account' refers to the type of seller profile with a 'Business' profile corresponding to profiles created solely for marketplace activities, while 'Personal' profiles represent profiles created for personal use. 'Public/Private Profile' represents the account's privacy setting. Religious Content refers to the public presence of any religious content on the seller's profile, and 'Marital Status' reflects the publicly posted marital status of the seller. 'Selfies/Personal Photos on Profile' refers to whether the seller publicly posts personal photos on the profile. The 'Not Known' under each category represents the instances when the relevant information could not be extracted from the profile's public information.

Table 3: Treatment effects for main variables of interest.

	(1) First Price	(2) Final Price w/Delivery	(3) High Quality	(4) Unsolicited Messages per Day	(5) Language Index
Female	-3.363	-8.610	0.00935	0.872***	-0.122*
	(5.031)	(9.027)	(0.0312)	(0.109)	(0.0628)
p-value	0.504	0.34	0.765	0	0.053
q-value	0.608	0.516	0.849	0.001	0.119
Constant	1,301***	1,321***	0.0467**	0.0330	2.264***
	(32.52)	(32.95)	(0.0205)	(0.0705)	(0.139)
Observations	446	446	214	2,504	2,219
Sellers	223	223	107	-	222

Note: The table presents our main results. Treatment effects/discrimination are captured by the coefficients of the variable Female. We use a balanced within-subject design as our main specifications. For prices and quality (1-3), we use a linear mixed-effects model. Unsolicited messages (4) are calculated for each buyer by summing up all messages and calls received over Facebook, WhatsApp, and mobile phone numbers, with results reported from a Poisson regression of count variables. The language index (5) measures the level of professionalism exhibited by the seller, as reported by a team of 30 evaluators, and we report the results of within-subject differences with annotator fixed effects. A higher value on the index indicates a higher level of perceived professionalism. Standard errors in specifications 1-3 and 5 are clustered at the seller level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of Buyer's Gender on Prices
Panel A: All Sellers

	(1) First Price	(2) Second Price	(3) Third Price	(4) Final Price	(5) Final Price w/ Delivery	(6) Amount Paid
Female	-3.36	-0.77	-1.37	-1.13	-8.61	-12.47
	(5.03)	(5.87)	(6.01)	(5.98)	(9.02)	(9.38)
Constant	1300.96***	1246.18***	1236.85***	1234.79***	1320.68***	1365.69***
	(32.48)	(31.78)	(31.70)	(31.70)	(32.91)	(37.28)
Observations	446	446	446	446	446	276
Clusters/Sellers	223	223	223	223	223	169

Panel B: Male Sellers

	First Price	Second Price	Third Price	Final Price	Final Price w/ Delivery	Amount Paid
Female	-5.62	-0.60	-0.17	-0.13	-4.30	-14.99
	(6.14)	(7.60)	(7.78)	(7.75)	(11.99)	(10.11)
Constant	1270.75***	1215.99***	1205.33***	1202.28***	1277.31***	1340.43***
	(40.39)	(39.48)	(39.37)	(39.36)	(40.08)	(44.97)
Observations	302	302	302	302	302	191
Clusters/Sellers	151	151	151	151	151	117

Panel C: Female Sellers

	First Price	Second Price	Third Price	Final Price	Final Price w/ Delivery	Amount Paid
Female	1.38	-1.13	-3.90	-3.21	-17.65	-6.81
	(8.79)	(8.85)	(9.00)	(8.98)	(12.20)	(20.43)
Constant	1364.32***	1309.49***	1302.96***	1302.96***	1411.64***	1422.32***
	(53.90)	(52.85)	(52.66)	(52.66)	(56.59)	(66.87)
Observations	144	144	144	144	144	85
Clusters/Sellers	72	72	72	72	72	52

Note: The table presents the results of a linear mixed-effects model with clustered standard errors at the seller level. 'Female' is the binary variable taking value one if the buyer's gender was assigned as female, zero otherwise. The 'First Price' corresponds to the first privately quoted price by the seller to a buyer at the start of the bargaining process. Similarly, the 'Second Price' and 'Third Price' reflects the prices quoted during the second and third stages of bargaining, while the 'Final Price' is the finally agreed price between the buyer and seller in response to the outlined bargaining process. 'Final Price w/ Delivery' add delivery charges, if any, to the 'Final Price.' 'Amount Paid' is the amount paid by the buyer at the time of the product's delivery. All prices are in PKR. Data in all panels is restricted to observations where the sellers could agree on a final price from both buyers. Panel A includes all the sellers, while Panels B and C restrict data to male and female sellers. P-values adjusted for multiple hypothesis testing using Westfall and Young (1993) method. Significance levels (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$) are based on these adjusted p-values.

Table 5: Effect of Buyer's Gender on non-Price Outcomes

Panel A: All Sellers

	(1) Incomplete Bargaining	(2) Bargaining Stages	(3) Required Advance	(4) Order Completed	(5) Order Delivered	(6) Delivery Time	(7) Same as Ordered
Female	-0.04*	0.03	0.00	0.05***	-0.04	4.06	0.02
	(0.02)	(0.03)	(0.01)	(0.02)	(0.04)	(3.93)	(0.04)
Constant	0.26***	1.42***	0.10***	0.36***	0.78***	45.76***	0.82***
	(0.02)	(0.03)	(0.01)	(0.02)	(0.03)	(4.13)	(0.03)
Observations	1064	1064	1064	1064	417	321	321
Sellers	532	532	532	532	265	214	214

Panel B: Male Sellers

	Incomplete Bargaining	Bargaining Stages	Required Advance	Order Completed	Order Delivered	Delivery Time	Same as Ordered
Female	-0.04	0.03	0.01	0.04	-0.08*	2.57	0.02
	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(4.01)	(0.04)
Constant	0.27***	1.46***	0.10***	0.41***	0.81***	49.07***	0.82***
	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(4.90)	(0.04)
Observations	668	668	668	668	286	222	222
Sellers	334	334	334	334	183	148	148

Panel C: Female Sellers

	Incomplete Bargaining	Bargaining Stages	Required Advance	Order Completed	Order Delivered	Delivery Time	Same as Ordered
Female	-0.03	0.03	-0.01	0.08***	0.06	7.93	0.03
	(0.03)	(0.05)	(0.02)	(0.03)	(0.06)	(8.70)	(0.07)
Constant	0.24***	1.35***	0.11***	0.29***	0.71***	37.87***	0.82***
	(0.03)	(0.04)	(0.02)	(0.03)	(0.06)	(7.66)	(0.06)
Observations	396	396	396	396	131	99	99
Sellers	198	198	198	198	82	66	66

Note: The table presents the results of a linear mixed-effects model with clustered standard errors at the seller level on various outcomes before the delivery of the product. 'Female' is the binary variable taking value one if the buyer's gender was assigned as female, zero otherwise. 'Bargaining Withdrawal' is a binary variable that takes a value one if a seller withdraws from the bargaining by not responding to the buyer. 'Bargaining Stages' refers to the number of stages before the price is finalized. 'Require Advance' takes a value of one when a seller requires advance payment before the item's delivery and a value of zero otherwise. 'Order Completed" is a binary variable that takes value one if the bargaining led to the successful placement of the order for the item. Panel A includes all the sellers, while Panels B and C restrict data to male and female sellers. P-values adjusted for multiple hypothesis testing using Westfall and Young (1993). Significance levels (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$) are based on these adjusted p-values.

Table 6: Incidence of Unsolicited Communication Attempts

	(1)	(2)	(3)	(4)	(5)	(6)
	Phone		Facebook		WhatsApp	
	Calls	Messages	Messages	Friend-Requests	Messages	Calls
Female	1.38***	1.58***	9.50***	1.9e+07	8.17***	1.90**
	(0.10)	(0.07)	(7.06)	(2.9e+10)	(0.67)	(0.51)
Constant	0.24***	0.63***	0.00***	0.00	0.13***	0.02***
	(0.01)	(0.02)	(0.00)	(0.00)	(0.01)	(0.00)
Observations	2504	2504	2504	2504	2504	2504

Note: The table presents results from the Poisson regression of count variables on gender. Post-Transaction Message (column 1) captures the number of times a seller messages the buyer after the completion of the transaction. Phone Calls (column 2) and Phone Messages (column 3) measure the number of calls and messages received per day per buyer during the experiment. Similarly, Facebook Message (column 4) and WhatsApp Message (column 5) measure the number of messages received on Facebook and WhatsApp per day for each buyer during the study duration. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Seller's Conversation Style by Buyer's Gender

Panel A: All Sellers

	(1) Verbosity	(2) Honorific Use	(3) Casual Lexicon	(4) Post Messages	(5) Transaction Messages - IRR
Female	46.81**	-0.14*	-0.04	0.20	1.24***
	(20.87)	(0.08)	(0.05)	(0.56)	(0.06)
Constant	364.48***	0.93***	0.30***	5.97***	0.68***
	(20.69)	(0.09)	(0.03)	(0.43)	(0.02)
Observations	1064	1064	1064	417	2568
Clusters/Sellers	532	532	532	265	

Panel B: Male Sellers

	Verbosity	Honorific Use	Casual Lexicon	Post Messages	Transaction Messages - IRR
Female	56.38**	-0.19*	-0.13***	0.65	1.22***
	(28.23)	(0.11)	(0.05)	(0.61)	(0.07)
Constant	389.16***	1.04***	0.35***	5.28***	0.51***
	(26.29)	(0.12)	(0.04)	(0.42)	(0.02)
Observations	668	668	668	286	2208
Clusters/Sellers	334	334	334	183	

Panel C: Female Sellers

	Verbosity	Honorific Use	Casual Lexicon	Post Messages	Transaction Messages - IRR
Female	30.67	-0.05	0.11	-0.83	1.27***
	(29.64)	(0.14)	(0.09)	(1.17)	(0.10)
Constant	322.83***	0.74***	0.20***	7.55***	0.31***
	(33.41)	(0.13)	(0.05)	(1.04)	(0.02)
Observations	396	396	396	131	2040
Clusters/Sellers	198	198	198	82	

Note: Columns 1-4 presents the results of a linear mixed-effects model with clustered standard errors at the seller level on various outcomes related to the sellers' conversation style. 'Female' is the binary variable taking value one if the buyer's gender was assigned as female, zero otherwise. 'Verbosity' captures the Verbosity of the seller (in words per conversation) when interacting with a particular gender, "Honorific Use" indicates if more formal terms were used to refer to the buyer, and 'Casual Lexicon' indicates whether more informal language was used. For post-transaction messages, we conduct two analyses. Column 4 identifies sellers by phone number and profile and conducts seller-level analysis of messages received. Messages- IRR (column 5) analyzes the number of messages received post-transaction by buyer phone number and social media accounts. Panel A includes all the sellers, while Panels B and C restrict data to male and female sellers. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Linguistic Analysis of Seller's Responses

Panel A: All Sellers

	(1) Polite	(2) Clear	(3) Formal	(4) Enthusiastic	(5) Friendly	(6) Assertive	(7) Flirtatious
Female	0.01	0.01	-0.02*	0.02**	0.01	0.01	0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Constant	0.82***	0.77***	0.66***	0.52***	0.75***	0.72***	0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations	963	963	963	963	963	963	963
Clusters/Sellers	516	516	516	516	516	516	516

Panel B: Male Sellers

	(1) Polite	(2) Clear	(3) Formal	(4) Enthusiastic	(5) Friendly	(6) Assertive	(7) Flirtatious
Female	-0.00	0.01	-0.01	0.02	-0.00	0.02*	0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Constant	0.83***	0.77***	0.65***	0.53***	0.77***	0.71***	0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations	609	609	609	609	609	609	609
Clusters/Sellers	325	325	325	325	325	325	325

Panel C: Female Sellers

	(1) Polite	(2) Clear	(3) Formal	(4) Enthusiastic	(5) Friendly	(6) Assertive	(7) Flirtatious
Female	0.02*	0.01	-0.03*	0.03	0.03**	-0.00	0.01**
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)
Constant	0.80***	0.78***	0.67***	0.50***	0.73***	0.73***	0.01***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
Observations	354	354	354	354	354	354	354
Clusters/Sellers	191	191	191	191	191	191	191

Note: The table presents the results of a linear mixed-effects model with clustered standard errors at the seller level for various traits from sellers' language analysis. 'Female' is the binary variable taking value one if the buyer's gender was assigned as female, zero otherwise. 'Verbose' captures the Verbosity of the seller (in words per message) when interacting with a particular gender, 'Polite,' 'Clear,' 'Formal,' 'Enthusiastic,' 'Friendly,' and 'Assertive' assume values between 0 and 1, where closer to 1 indicates a stronger presence of the trait. Panel A includes all the sellers, while Panels B and C restrict data to male and female sellers. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Treatment Effects Heterogeneity

	All Sellers		Male Sellers		Female Sellers	
	(1)	(2)	(3)	(4)	(5)	(6)
	First Price	Final Price	First Price	Final Price	First Price	Final Price
Buyer Gender						
Female	-20.26	7.33	-8.70	15.12	-38.57	-1.28
	(30.32)	(17.44)	(28.86)	(24.74)	(62.00)	(23.25)
Product Orientation						
Male Oriented	-59.82	-50.13	5.80	-34.58	-189.52	-44.27
	(147.08)	(63.17)	(167.11)	(74.10)	(266.65)	(126.78)
Female × Male Oriented	-22.17	-6.26	-19.11	-4.76	-15.72	-13.84
	(21.69)	(20.33)	(28.39)	(24.85)	(31.73)	(31.31)
Religious Content						
Yes	-464.98***	-30.56	-446.56***	-54.87	-555.48***	6.88
	(102.02)	(72.63)	(123.65)	(91.24)	(187.91)	(115.15)
Female × Yes	-0.53	-18.07	8.72	-26.02	-4.46	0.17
	(19.40)	(21.48)	(27.84)	(28.83)	(28.96)	(24.68)
Marital Status						
Single	-237.59	-138.08*	-214.19	-162.33*	-271.61	4.70
	(165.31)	(77.53)	(194.97)	(88.16)	(257.49)	(156.39)
Female × Single	52.55*	80.74***	71.80**	90.45***	3.04	47.92
	(28.59)	(25.71)	(36.07)	(32.29)	(37.38)	(31.49)
Account Type						
Personal	209.79	-37.68	166.29	-1.56	267.07	-76.25
	(167.75)	(64.95)	(207.45)	(87.93)	(282.25)	(95.24)
Female × Personal	31.67	-9.38	19.45	-18.34	38.27	-3.80
	(33.97)	(18.81)	(38.46)	(25.46)	(60.16)	(21.32)
Account Privacy						
Public	299.50*	49.88	385.32**	58.76	166.22	32.49
	(162.90)	(65.14)	(180.68)	(83.83)	(300.86)	(105.58)
Female × Public	-25.83	-32.26*	-53.02	-31.59	12.44	-32.24
	(33.42)	(18.29)	(39.76)	(25.26)	(63.71)	(25.27)
Personal Photos						
Yes	-513.46***	-17.48	-464.98***	-13.93	-540.49***	32.86
	(134.99)	(63.93)	(179.66)	(84.47)	(194.72)	(101.84)
Female × Yes	19.61	11.49	28.57	13.72	23.61	8.12
	(24.50)	(18.61)	(29.42)	(23.23)	(36.36)	(26.19)
Constant	1696.16***	1376.20***	1573.73***	1321.55***	1850.65***	1428.92***
	(177.67)	(52.93)	(131.26)	(67.08)	(378.86)	(88.18)
Observations	860	580	545	390	315	190
Clusters/Sellers	474	357	302	239	172	118

Note: The table presents the results of the heterogeneity tests using a linear mixed-effects model with clustered standard errors at the seller level. Columns (1) and (2) include all sellers while (3) and (4) ((5) and (6)) subsample to male (female) sellers. 'First Price' and 'Final Price' correspond to the seller's first privately quoted price and the final quoted price (including delivery charges). All prices are in PKR. 'Product Orientation' captures the predominant target gender of the product, 'Religious Content' refers to the public presence of any religious content on the seller's profile with the reference category of no or unknown content, 'Marital Status' reflects the publicly posted marital status of the seller with the reference category of any status other than single, including unknown. 'Account Type refers to the type of seller profiles with the reference category of a 'business' profile created solely for marketplace activities. 'Account Privacy' represents the account's privacy with the reference category of private profile. 'Personal Photos' takes the value of one if the seller publicly posts personal photos on the profile with the reference category of 'no' or 'unknown' presence of photos. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Model: Disaggregating price discrimination

This section presents a conceptual framework highlighting the behavioral forces involved in price-setting behavior. Our setup is inspired by Bohren et al. (2019). The framework presented here is closely tied to our experiment design and helps inform our treatments.[18]

Consider a buyer who has observable group identity $g \in \{F, M\}$ and unobservable valuation for a good $v \sim N(\mu_g, 1/\tau_v)$ with mean $\mu_g \in \mathbb{R}$ and precision $\tau_v > 0$. The buyer sends a sequence of messages at times $t = 1, 2, ...$ to the seller. Each message reveals a signal, $s_t = v + \eta_t$, of the true valuation of the buyer, where $\eta_t \sim N(0, 1/\tau_{\eta_t})$ is an independent random shock with precision $\tau_{\eta_t} > 0$. Lower signal precision at time t reflects greater uncertainty in valuation. This precision can be interpreted as the amount of subjectivity in judgment involved in evaluating valuation, with lower precision implying greater subjectivity. We assume that the valuation for good is fixed across time,[19] and higher valuation generates a higher expected signal.

A seller quotes a price to the buyer, $p_t \in \mathbb{R}$. Before quoting the price at time t, the seller observes the buyer's gender g, history of the past signals by the buyer $h_t = (s_1, ..., s_{t-1})$, where $h_1 = \emptyset$, and signal s_t. A seller's type θ_i determines her preferences and inference model, including her subjective belief about the relationship between gender and valuation. We assume that the seller's cost of production of the good is zero, and the seller's payoff from quoting a price p to a buyer of gender g is given as

$$\pi_{ig} = -\left(p - (v + c_g^i - \delta_g^i)\right)^2 \tag{1}$$

where c_g^i is a type-specific taste parameter Ã la Becker (1957). Normalize $c_M^i = 0$. $c_F^i > 0$ corresponds to distaste from transacting with female buyers. δ_g^i captures the type-specific benefit or perverse gratification from harassing the buyer of gender g 'a la Basu (2003). Normalize $\delta_M^i = 0$. $\delta_F^i > 0$ corresponds to positive utility from harassing a female buyer. The seller has subjective prior beliefs $\hat{\mu}_g$ about the average valuation of a buyer of gender g.[20]

A seller of type θ_i gets a disutility from transacting with female buyers if $c_F^i > 0$. A seller of type θ_i has a preference for harassing female buyers if $\delta_F^i > 0$. A seller of type θ_i has a belief favoring male buyers if $\hat{\mu}_M^i < \hat{\mu}_F^i$.

The seller learns about the buyer's valuation from the history of counter offers. Her posterior belief about valuation is derived using the Bayes rule, given her model of inference. Each seller chooses the price that maximizes their expected payoff concerning their posterior belief about valuation. Suppose a seller has type θ_i and let

$$p_i(h, s, g) \equiv arg \underset{p \in \mathbb{R}}{max} \hat{E}_i \left[-\left(p - (v + c_g^i - \delta_g^i)\right)^2 | h, s, g \right] \tag{2}$$

denote the optimal price conditional on observing history h and signal s from a buyer of gender

[18] See DellaVigna (2018) for motivation on designing experiments using a model of behavior.

[19] This is equivalent to assuming that the discount factor in the bargaining model is equal to unity, i.e., the buyer is patient and values the same price trade equally at different periods.

[20] A seller can have a misspecified model of the relationship between gender and valuation, in that case, the seller's subjective belief may differ from the true population average valuation, $\hat{\mu}_g^i \neq \mu_g$.

g, where $\hat{E}_i$ denotes the expectation concerning her model of inference. Then, the optimal price in period t is

$$p_i(h_t, s_t, g) = \hat{E}_i[v|h_t, s_t, g] + c_g^i - \delta_g^i \tag{3}$$

Discrimination is the disparate quoting of prices based on the group to which the buyer belongs, i.e., gender, rather than on individual attributes, i.e., signal and history. Gender discrimination occurs when a male and female buyer with the same history and signal receives different prices. Let

$$D_i(h, s) \equiv p_i(h_t, s, F) - p_i(h_t, s, M) \tag{4}$$

denote the difference between type θ_i's quoting of prices to a male and female buyer conditional on observing history h and signal s.

A.1 Discrimination in First Price

We first examine how the preferences and beliefs impact the first quoted prices by the seller. Consider the quoting of a price to a buyer of gender g by a seller who has subjective prior beliefs $(\hat{\mu}_F, \hat{\mu}_M)$ about average valuation, taste parameter c_F, harassment parameter δ_F, and observes signal s_1. The initial signal has conditional distribution $s_1|v \sim N(v, 1/\tau_{\eta_1})$. Given the prior beliefs and signal distribution, the seller's posterior belief about valuation conditional on observing s_1 is normally distributed, $v|s_1 \sim N\left(\frac{\tau_v \hat{\mu}_g + \tau_{\eta_1} s_1}{\tau_v + \tau_{\eta_1}}, \frac{1}{\tau_v + \tau_{\eta_1}}\right)$. From 3, the optimal price is equal to

$$p_1(h_1, s_1, g) = \frac{\tau_v \hat{\mu}_g + \tau_{\eta_1} s_1}{\tau_v + \tau_{\eta_1}} + c_g^i - \delta_g^i \tag{5}$$

Higher signals and higher expected valuation result in higher first prices - the optimal first price is strictly increasing in s_1 and $\hat{\mu}_g$.

Discrimination in the first price depends on the seller's preferences and prior beliefs about valuation. From 5, first price discrimination is independent of the signal and equal to

$$D(h_1, s_1) = \frac{\tau_v}{\tau_v + \tau_{\eta_1}}(\hat{\mu}_F - \hat{\mu}_M) + c_F^i - \delta_g^i \tag{6}$$

There is discrimination against females in the first price, i.e., $D(h_1, s_1) > 0$, if the seller has unfavorable beliefs about valuation ($\hat{\mu}_F > \hat{\mu}_M$) and/or if the distaste towards women is greater than the utility from harassment ($c_F > \delta_F$). On the other hand, the discrimination in the first price could be in favor of women, i.e., $D(h_1, s_1) < 0$, if the sellers benefit from harassing females ($\delta_F > 0$) more than the distaste from interacting with them ($c_F < \delta_F$). The intuition is that the seller benefits from harassing the female buyer and is willing to accept a lower price for the perverse gratification of harassment. Of course, the effects of distaste and harassment may cancel out each other, in which case, discrimination in the first price arises solely due to differences in beliefs about valuations.

Equation 6 shows that varying the level of subjectivity in judgment differentially impacts initial discrimination depending on whether it is due to preferences (distaste or harassment) or beliefs. This comparative static can be used to identify the source of discrimination.

A.2 Discrimination in Sequential Prices

We now study how discrimination evolves across a sequence of messages from the buyer. Beginning in the second period, signals from the buyer provide information about the buyer's valuation. In our experimental setting, the buyer is constantly requesting a discount, which, in the terminology of this model, is equivalent to sending signals such that $s_1 > s_2 > s_3 > ... > s_n$, and since such signals are expected to reveal the buyer's low valuation, it can reasonably be assumed that the precision of the signal is increasing which each request for a discount i.e., $\tau_{\eta_1} < \tau_{\eta_2} < ... < \tau_{\eta_n}$. In the second period, the seller observes the signal s_2 and once again uses the Bayes rule to form a posterior about the buyer's valuation, i.e., the seller posterior belief on observing s_2 is normally distributed, $v|s_1, s_2 \sim N\left(\frac{\tau_v\hat{\mu}_g+\tau_{\eta_1}s_1+\tau_{\eta_2}s_2}{\tau_v+\tau_{\eta_1}+\tau_{\eta_2}}, \frac{1}{\tau_v+\tau_{\eta_1}+\tau_{\eta_2}}\right)$. From 3, the optimal price is now equal to

$$p_2(h_2, s_2, g) = \frac{\tau_v\hat{\mu}_g + \tau_{\eta_1}s_1 + \tau_{\eta_2}s_2}{\tau_v + \tau_{\eta_1} + \tau_{\eta_2}} + c_g^i - \delta_g^i \tag{7}$$

Comparing quoted prices to gender g in time period 1 (equation 5) and time period 2 (equation 7) reveals that any difference in prices between the two periods is driven by the seller's beliefs about the buyer's valuation since the preference parameters are assumed to be fixed over time.[21]

The price discrimination in period 2 is analogous to discrimination in period 1, indicating that the price discrimination against females is driven positively by beliefs and distaste against female buyers and negatively by perverse gratification from harassment.

$$D(h_2, s_2) = \frac{\tau_v}{\tau_v + \tau_{\eta_1} + \tau_{\eta_2}}(\hat{\mu}_F - \hat{\mu}_M) + c_F^i - \delta_g^i \tag{8}$$

Comparing discrimination across periods helps us identify the source of the discrimination, i.e.,

$$D(h_1, s_1) - D(h_2, s_2) = \frac{\tau_v\tau_{\eta_2}}{(\tau_v + \tau_{\eta_1} + \tau_{\eta_2})(\tau_v + \tau_{\eta_1})}(\hat{\mu}_F - \hat{\mu}_M)$$

indicating that the difference in discrimination between the two periods is purely driven by differences in the beliefs about the valuations of each gender.

As buyers send more signals (request discounts), the n-period discrimination is given by:

$$D(h_n, s_n) = \frac{\tau_v}{\tau_v + \tau_{\eta_1} + \tau_{\eta_2} + ... + \tau_{\eta_n}}(\hat{\mu}_F - \hat{\mu}_M) + c_F^i - \delta_g^i \tag{9}$$

Equation 9 reveals that discrimination due to differences in beliefs (first term) decreases with an increase in the precision of the signals (τ_{η_i} for $i = 1, 2, ...$). This implies that as $\tau_{\eta_i} \to \infty$, the discrimination against female buyers arises only due to the preference of sellers, i.e.,

[21] However, it is possible that the taste parameters (δ_g^i and c_g^i) get activated only after some communication has taken place between the buyer and the seller. So, the initial offer may not include the effect of distaste or harassment, and only when the buyer starts negotiating does the seller feel the urge to harass the buyer or get disutility from the interaction. Our model does not allow for this dynamic endogeneity of preferences.

$$D(h_n, s_n) \quad \rightarrow \quad c_F^i - \delta_F^i \quad \text{as} \quad n \rightarrow \infty \tag{10}$$

Consistent with Fitzpatrick (2017), we are postulating that discrimination in the first price can arise due to the beliefs or preferences of the sellers; however, any discrimination in the final prices must only be due to differences in preferences towards a gender. However, we can only identify the net effect of distaste and perverse gratification in our experiment and cannot isolate the discrimination from each preference source. This implies that any discrimination in final prices could be against females if the distaste outweighs the perverse gratification. Conversely, discrimination favoring females would imply that perverse gratification is the dominant driving force of favorable discrimination in final prices. Of course, the two forces may cancel each other out, and we may not observe discrimination in any direction.

B Additional Tables

Table B1: Robustness checks.

	(1) First Price	(2) Final Price w/Delivery	(3) High Quality	(4) Buyer Language Index
Female	-9.712	-7.802	0.00346	-0.00632
	(13.53)	(8.976)	(0.0208)	(0.0433)
p-values	0.473	0.385	0.868	0.884
Constant	1,721***	1,336***	0.0325**	2.463***
	(69.64)	(27.10)	(0.0143)	(0.0466)
Observations	860	580	321	2,219
Sellers	474	357	214	222

Notes: Columns 1-3 report the results from regression analysis of our main specifications with the full unbalanced sample, dropping the restriction of a balanced within-subject design. Column 4 reports the results for checking the Language index for buyer analysis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
Category									
arts	1,870	1,309	12	650,000	3,447	18,516	1650	2200	2900
health	1,859	1,284	11	123,466	1,594	4,842	500	1000	1600
furniture	1,810	1,513	11	3,150,000	27,861	91,422	1500	11500	30000
misc	1,784	1,344	11	9,000,000	287,721	1,046,039	600	1999	15000
cell-phones	1,708	1,672	11	1,111,111	29,653	42,051	14000	22500	32000
home-decor	1,562	1,110	11	150,000	3,261	9,659	500	1150	2500
kitchen-products	1,394	1,098	11	999,999	7,065	38,067	650	1699	5000
bags	1,373	1,004	11	35,000	2,172	1,654	1499	2000	2700
shoes	1,224	977	14	40,000	2,367	1,979	1350	1950	3000
mens	1,176	1,002	13	60,000	2,004	2,396	1000	1610	2400
womens	1,151	846	12	100,000	3,250	6,521	1699	2450	3200
kids-clothing	1,068	811	11	40,000	2,031	3,334	800	1499	2250
bedding	1,061	844	12	40,000	2,353	2,484	1150	1550	3000
computers	945	867	12	299,000	31,552	38,854	4000	19500	45000
appliances	840	709	12	999,999	18,146	59,726	1500	5000	16000
portable-audio-video	658	612	15	35,000	1,952	2,762	615	1400	2300
kids	619	555	12	160,000	4,883	9,657	750	2500	5500
home-lighting	614	439	11	125,000	3,263	8,586	649	1600	3200
autoparts	485	430	11	8,100,000	218,213	657,383	1200	7000	44000
home-audio-video	416	386	50	499,999	19,516	49,836	1295	3374.5	16000
media	414	344	30	1,234,567	8,608	67,990	400	1149.5	2200
cables-adaptors	359	333	40	123,456	1,415	7,093	220	350	850

[]

Continued on the next page

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
security-cameras	339	306	25	65,000	9,082	9,576	3499	5000	15000
tools	336	302	11	1,350,000	39,645	141,324	1234	4200	16500
bath-products	335	289	14	40,000	3,151	4,313	600	2200	4200
cell-phone-accessories	335	321	16	240,000	8,740	22,281	400	800	5800
home-heating-cooling	334	288	11	3,024,884	23,808	180,980	1449.5	2350	7500
scrap-metal	247	176	11	9,000,000	476,584	1,413,558	287.5	4350	111728
powersports	215	184	14	7,000,000	567,503	1,006,186	45000	122500	700000
guitars-basses	210	197	111	600,000	19,059	49,309	7000	10000	17500
video-games-consoles	210	200	50	240,000	13,733	26,036	1075	4050	15000
printers-scanners-fax	207	178	31	1,300,000	44,753	149,826	3500	14000	27999
sports-gear	205	163	18	69,999	4,826	9,011	700	2000	5000
costumes	200	139	13	25,000	1,743	2,600	400	1150	2200
power-adapters-chargers	189	178	15	111,111	2,940	10,265	400	778	1560
outdoor-recreation-gear	188	172	14	365,000	21,984	44,676	2700	8750	21000
cleaning-supplies	181	174	11	80,000	4,328	6,787	2000	3200	4000
baby-clothing	178	128	12	3,899	1,234	761	500	1242	1625
motorcycles	158	141	70	850,000	78,734	86,762	35000	68000	95000
antiques	128	89	14	3,645,000	51,545	385,830	950	1700	11500
exercise-fitness	125	105	15	176,699	24,520	38,072	975	6999	27000
planners	103	79	20	2,375,000	79,691	290,007	850	1600	8500
outdoor-cooking-equipments	102	88	99	1,000,000	51,029	133,602	3500	14500	35000
chalkboards	100	74	25	8,100,000	126,176	942,710	965	2000	12500
shipping-containers	73	55	45	7,070,000	354,560	1,286,983	550	1650	5300
adidas-hoodies	65	53	150	5,500	1,609	721	1300	1550	1800

[]

Continued on the next page

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
bathroom-faucets	65	56	12	19,500	4,460	4,015	692.5	5000	6650
dolls	61	45	100	12,345	1,809	2,047	750	1200	2000
flash-drives	60	56	42	23,500	3,793	4,907	925	1775	4500
audio-equipment	55	41	123	250,000	31,396	42,559	6500	17000	45000
label-makers	53	38	45	7,100,000	350,369	1,486,417	720	2649.5	7500
toy-vehicles	53	51	75	975,000	34,879	158,960	550	1000	2500
car-electronics	51	48	35	350,000	16,468	51,054	2800	4750	11000
stuffed-animals	51	40	100	16,500	2,904	4,200	350	1025	2650
walnut-lumber	50	44	16	88,000	5,045	18,006	500	1000	2050
cash-registers	49	38	85	1,700,000	170,627	407,561	1200	8500	35000
asphalt-paving	47	38	15	4,700,000	725,165	1,120,706	10000	216000	1600000
heated-blankets	45	33	150	123,456	11,007	29,165	1500	3150	6500
shoe-shine-kits	43	27	17	20,000	2,628	3,949	350	1799	3100
tire-machines	41	33	12	1,325,000	417,831	491,094	5000	120000	800000
apple-pencils	40	36	123	515,000	71,480	95,116	14999.5	57500	83999.5
tongue-and-groove-planks	40	34	65	1,234,645	82,356	269,014	350	1349.5	4500
electric-blankets	38	26	186	123,456	8,782	23,606	1900	3325	7500
caps	36	27	22	15,000	2,505	2,706	1199	2000	2800
cat-supplies	36	29	50	11,935	2,490	3,401	465	1050	2500
pretend-play-toys	36	30	480	20,000	2,853	4,349	950	1470	2150
microscopes	34	25	35	2,495,000	186,803	631,048	350	3000	7500
square-steel-tubes	34	27	20	123,456	16,275	33,713	220	1690	18000
fill-dirt	33	31	123	8,100,000	1,806,247	2,159,147	13000	1600000	2550000
pedestal-sinks	33	25	123	6,900,000	282,012	1,378,761	2000	4500	8500

[] *Continued on the next page*

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
stainless-steel-sinks	32	23	149	250,000	29,237	54,918	6000	11800	27000
lockers	31	23	365	10,000	3,105	2,449	1700	2500	3500
tailored-clothing	29	23	149	30,000	3,832	6,757	800	1450	4000
educational-toys	26	24	280	3,500	1,315	972	599.5	959.5	1899
action-figures	25	20	300	5,400	2,113	1,554	1020	1750	2787.5
bird-supplies	24	19	1,100	17,000	7,592	4,634	3500	7500	11500
cork-boards	23	16	25	5,850	1,641	1,381	675	1575	2150
dollhouses	23	19	15	45,000	6,248	11,230	1000	1600	4950
melodica-instruments	23	22	800	200,000	23,227	42,670	1950	13999.5	18000
pianos-keyboards	23	20	12	46,000	12,316	13,681	1675	4150	19500
fathers-day-gifts	19	16	150	3,850	1,259	1,156	335	1087.5	1775
nebulizers	19	11	123	21,500	6,066	6,945	2000	3500	5000
pallet-jacks	19	17	122	2,200,000	249,076	550,057	11700	42500	125000
quartz-counter-tops	19	15	123	1,234,567	87,699	317,370	470	3000	18000
wooden-toys	19	15	15	16,000	2,365	4,044	290	1450	2000
corrugated-sheets	17	12	95	850,000	75,342	244,063	124	2240	11500
bird-wildlife-accessories	16	13	80	1,799	891	687	300	500	1650
birthday-decorations	16	7	220	2,200	1,041	937	220	500	2000
laser-pointers	15	12	350	1,600	1,056	489	535	1075	1525
safety-jackets	14	8	450	5,500	2,955	1,822	1545	2800	4500
white-noise-machines	14	13	380	43,000	10,923	13,584	1499	2000	20000
corded-phones	13	11	45	4,000	1,004	1,156	200	585	1500
desk-organizers	13	12	110	7,200	1,481	1,996	300	750	1825
accordions	12	10	800	200,000	31,730	59,714	6000	16500	22000

[]

Continued on the next page

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
cameras	12	9	1,000	138,000	53,556	41,368	23000	52000	65000
magnifying-glasses	12	9	399	2,955	1,389	801	799	1600	1700
round-pens	12	9	123	21,500	6,736	7,654	1500	4500	5000
landline-phones	11	10	1,150	26,500	5,480	7,846	2000	2450	3000
alpinestars-motorcycle-riding-gear	10	8	300	16,000	5,831	5,551	1175	4500	9500
pet-collars-leashes	10	8	123	4,000	1,321	1,377	375	922.5	1924.5
teacher-supplies	10	9	248	9,500	2,866	2,927	600	2450	3500
centrifuges	9	4	135	40,000	14,034	17,779	3067.5	8000	25000
ice-melt	9	6	123	21,500	8,887	8,663	4200	4750	18000
nike-windbreaker-jackets	9	6	1,000	7,000	2,431	2,269	1234	1675	2000
pet-feeding-supplies	9	8	300	14,999	3,162	4,892	600	1600	2799.5
puzzles	9	7	170	780	370	196	250	350	400
roof-trusses	9	6	123	21,500	8,887	8,663	4200	4750	18000
walkie-talkies	9	9	150	240,000	38,850	77,588	3500	9500	15000
bicycles	8	8	3,500	245,000	81,500	99,335	10000	24250	167500
microphones	8	8	370	15,999	6,406	5,862	1190	4749	11500
paper-cutters	8	5	50	12,500	3,034	5,313	399	970	1250
toilets	8	6	450	1,100,000	184,583	448,462	600	1924.5	2599
cars	7	4	999	1,320,000	404,000	622,003	15499.5	147500	792500
laminators	7	6	140	1,250	412	424	150	267.5	399
the-grinch-shirts	7	7	170	1,250	527	404	170	300	750
couples-shirts	6	4	18	1,500	1,067	704	659	1375	1475
fog-machines	6	5	3,500	25,000	12,900	8,532	6000	15000	15000
packers-nfl-apparel	6	4	240	15,500	4,835	7,150	920	1800	8750

[] *Continued on the next page*

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
readymade-clothing	6	3	12	2,700	1,304	1,347	12	1200	2700
school-supplies	6	4	123	1,800	626	792	151.5	289.5	1099.5
studio-monitors	6	2	32,000	75,500	53,750	30,759	32000	53750	75500
credit-card-readers	5	4	35	67,000	18,046	32,718	92.5	2575	36000
envelopes	5	4	399	2,000	1,400	750	824.5	1600	1975
floor-tiles	5	4	138	2,200	662	1,025	144	155	1180
journal-notebooks	5	5	499	1,800	850	535	650	650	650
mail-organizers	5	5	250	1,080	566	378	250	400	850
mont-blanc-pens	5	5	399	4,500	1,650	1,656	600	1250	1500
mothers-day-gifts	5	5	350	2,700	950	984	499	599	600
paper-shredders	5	5	395	650,000	132,309	289,424	399	1250	9500
rf-modulators	5	4	1,490	70,000	18,618	34,255	1490	1490	35745
rolling-storage-carts	5	5	395	650,000	131,309	289,962	399	1250	4500
safes	5	5	18,000	3,900,000	1,084,876	1,670,697	30000	126378	1350000
whiteboards	5	3	799	2,000	1,266	643	799	1000	2000
a4-paper	4	4	50	20,000	6,375	9,357	250	2724.5	12499.5
award-ribbons	4	4	395	650,000	163,011	324,660	397	824.5	325625
drums	4	2	7,499	9,500	8,500	1,415	7499	8499.5	9500
gift-cards	4	3	550	9,000	3,400	4,850	550	650	9000
iphone-xr-black	4	3	33,000	75,000	52,667	21,127	33000	50000	75000
packing-moving-boxes	4	4	1,600	239,000	61,650	118,236	2050	3000	121250
place-card-holders	4	4	395	650,000	163,011	324,660	397	824.5	325625
plexiglass-shields	4	4	395	650,000	163,011	324,660	397	824.5	325625
poly-mailer-bundles	4	4	395	650,000	163,011	324,660	397	824.5	325625

[] *Continued on the next page*

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
stationery-sets	4	4	99	1,350	537	556	199	349	874.5
usb-adapters	4	4	700	1,799	1,275	451	975	1300	1574.5
wedding-decorations	4	3	2,700	24,950	10,517	12,514	2700	3900	24950
apple-iphone-xr-unlocked	3	3	33,000	49,500	40,833	8,282	33000	40000	49500
bubble-wrap	3	3	65	700	421	325	65	499	700
electric-scooters	3	3	8,000	60,000	30,000	26,907	8000	22000	60000
guitar-pedals	3	2	400	70,000	35,200	49,215	400	35200	70000
insulation-boards	3	2	550	750	650	141	550	650	750
pendleton-apparel	3	2	500	1,900	1,200	990	500	1200	1900
software	3	3	500	13,500	5,167	7,234	500	1500	13500
string-instruments	3	3	4,000	13,500	10,167	5,346	4000	13000	13500
apartments-for-rent	2	2	25,000	38,000	31,500	9,192	25000	31500	38000
boats	2	2	50,000	94,000	72,000	31,113	50000	72000	94000
fire-extinguishers	2	2	500	15,000	7,750	10,253	500	7750	15000
houses-for-rent	2	1	2,200,000	2,200,000	2,200,000	.	2200000	2200000	2200000
jewelry	2	1	250	250	250	.	250	250	250
juneteenth	2	1	1,200	1,200	1,200	.	1200	1200	1200
micro-sd-cards	2	2	2,675	3,600	3,138	654	2675	3137.5	3600
peg-boards	2	2	291	399	345	76	291	345	399
pvc-pipes	2	2	6,999	6,999	6,999	0	6999	6999	6999
remote-car-starters	2	2	8,000	12,500	10,250	3,182	8000	10250	12500
ti-84-calculators	2	1	2,000	2,000	2,000	.	2000	2000	2000
trucks	2	2	1,290,000	2,558,500	1,924,250	896,965	1290000	1924250	2558500
vending-machines	2	1	7,800	7,800	7,800	.	7800	7800	7800

[]

Continued on the next page

Table B2: Summary of Posts on Facebook Marketplace

	N	Non-Missing	Min	Max	Mean	SD	P25	P50	P75
wind-instruments	2	2	123	45,000	22,562	31,733	123	22561.5	45000
batteries	1	1	1,234	1,234	1,234	.	1234	1234	1234
clipboards	1	1	500	500	500	.	500	500	500
garagesale	1	1	250,000	250,000	250,000	.	250000	250000	250000
playstation-5-controllers	1	1	16,000	16,000	16,000	.	16000	16000	16000
scissors	1	1	399	399	399	.	399	399	399
townhouses-for-rent	1	0	.	.	.	.			
trophies	1	1	399	399	399	.	399	399	399
vintage-school-desks	1	1	399	399	399	.	399	399	399
water-features	1	1	450	450	450	.	450	450	450
water-softeners	1	1	600	600	600	.	600	600	600
Total	31,120	25,151	11	9,000,000	44,657	352,746	1000	2200	8000

Note: This table summarizes the census of posts from Facebook Marketplace as of January 05th, 2022. Column N indicates the number of posts against a category, and Non-Missing indicates the number of posts with a positive posted price. Min, Max, Mean, and SD indicate the price's minimum, maximum, mean, and standard deviation by each category, respectively. P25, P50, and P75 are the 25th, 50th, and 75th percentiles of the posted price.

Table B4: Names of Buyers

First Name	Last Name	Gender
Shazia	Ali	Female
Samina	Rehman	Female
Saima	Iqbal	Female
Ayesha	Ahmed	Female
Muhammad	Iqbal	Male
Ahmed	Ali	Male
Abdul	Rehman	Male
Ali	Ahmed	Male

Notes: The table presents the selected names of buyers used for the experiment.

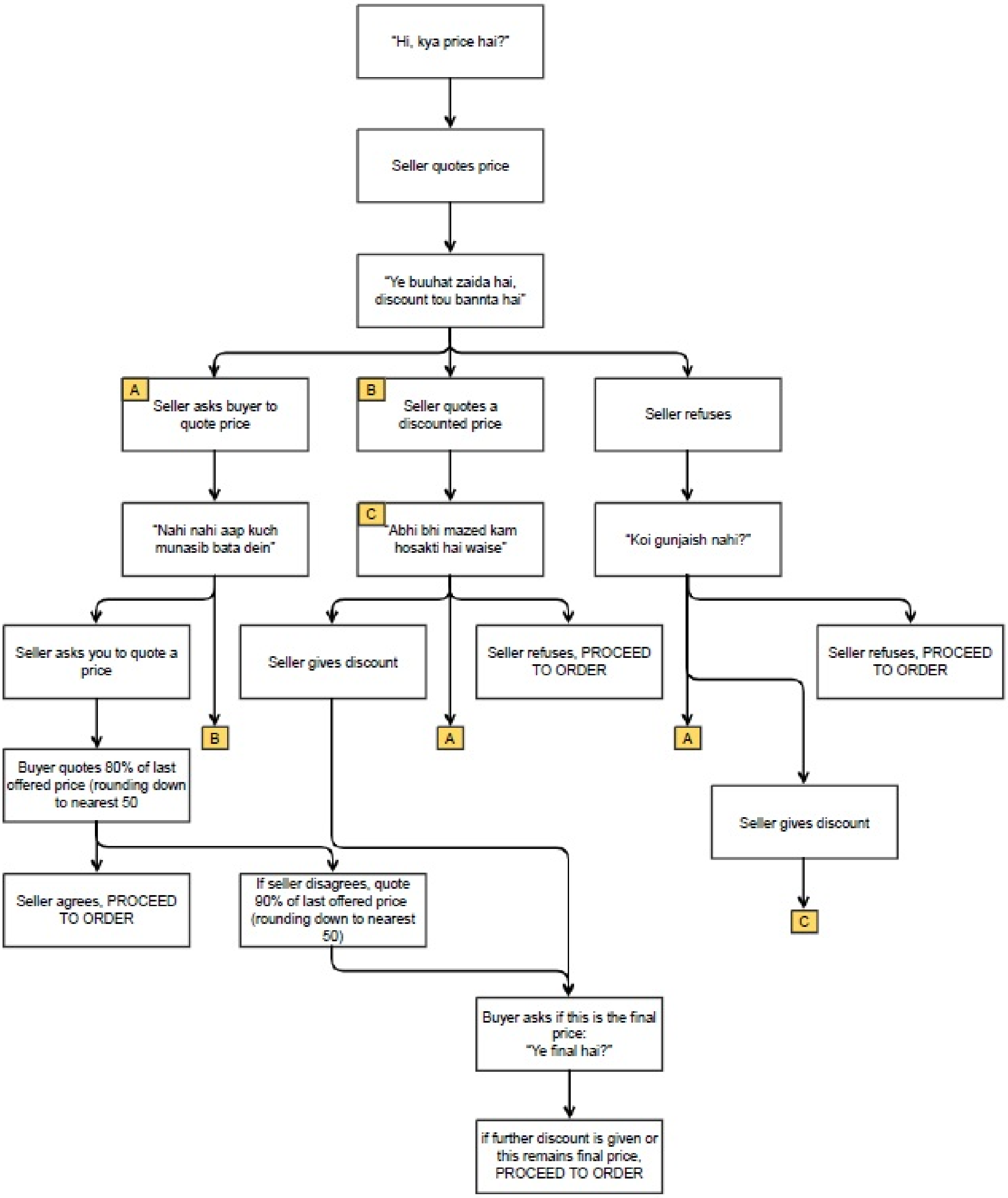
Table B3: Description of Selected Categories

Category	Description
arts	This mainly includes clothing articles for men and women with calligraphy, embroidery, and artwork.
health	This includes a variety of products ranging from skincare, hair care, beauty products, etc.
home-decor	This includes home decoration products such as frames, vases, clocks, lamps, etc.
bags	This includes bags such as handbags, wallets, clutches, pouches, etc. for men and women.
shoes	This includes shoes such as sandals, sneakers, boots, etc. for men and women.
mens	This includes products such as clothes, shoes, wallets, caps, etc. for men.
womens	This includes products such as clothes, shoes, wallets, caps, etc. for women.
kids-clothing	This includes clothing articles for kids.
bedding	This includes bed-sheets, comforter sets, pillows, blankets, etc.
portable-audio-video	This mainly includes earphones, headphones, portable speakers, etc.

Note: This table presents the description of categories that are selected for the study.

C Additional Figures

Figure C1: Bargaining Script 1



References

Ahmad, H., Banuri, S., and Bokhari, F. (2024). Discrimination in healthcare: A field experiment with Pakistan's transgender community. *Labour Economics*, 87:102490. 9

Alhorr, L. (2024). Virtual Windows Through Glass Walls? Digitization for Low-Mobility Female Entrepreneurs. 5

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, 103(484):1481–1495. 2.6, 3.2, 14

Asad, S. A., Banerjee, R., and Bhattacharya, J. (2023). Do workers discriminate against their out-group employers? Evidence from an online platform economy. *Journal of Economic Behavior & Organization*, 216:221–242. 1

Ayres, I., Banaji, M., and Jolls, C. (2015). Race effects on eBay. *The RAND Journal of Economics*, 46(4):891–917. 1

Ayres, I. and Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *American Economic Review*, pages 304–321. 1

Baktash, J. A. and Dawodi, M. (2023). Gpt-4: A Review on Advancements and Opportunities in Natural Language Processing. 4.2

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., ParientÃ©, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799. 9

Basu, K. (2003). The economics and law of sexual harassment in the workplace. *Journal of Economic Perspectives*, 17(3):141–157. A

Becker, G. S. (1957). *The economics of discrimination.* University of Chicago Press. 1, A

Bellemare, C., Bissonnette, L., and Kröger, S. (2014). Statistical power of within and between-subjects designs in economic experiments. 2.6

Benjamini, Y., Krieger, A., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507. 12

Bohren, J. A., Imas, A., and Rosenberg, M. (2019). The Dynamics of Discrimination: Theory and Evidence. *American Economic Review*, 109(10):3395–3436. 2.4, 3.2, A

Bryan, G., Choi, J., and Karlan, D. (2021). Randomizing religion: The impact of protestant evangelism on economic outcomes. *The Quarterly Journal of Economics*, 136:293–380. 9

Byrne, D. P., Martin, L. A., and Nah, J. S. (2022). Price Discrimination by Negotiation: a Field Experiment in Retail Electricity. *The Quarterly Journal of Economics*, 137(4):2499–2537. 1

Cameron, A. C. and Trivedi, P. K. (2022). Count-data models. In *Microeconometrics Using Stata*, volume II, chapter 20, pages 1021–1032. Stata Press, second edition. 3.2, 4.2

Castillo, M., Petrie, R., Torero, M., and Vesterlund, L. (2013). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics*, 99:35–48. 1, 1

Chandio, B. A., Imran, A. S., Bakhtyar, M., Daudpota, S. M., and Baber, J. (2022). Attention-Based RU-BiLSTM Sentiment Analysis Model for Roman Urdu. *Applied Sciences*, 12(7):3641. 17

Charness, G., Jabarian, B., and List, J. A. (2023). Generation Next: Experimentation with AI. 16

Chen, Y. (2024). Does the gig economy discriminate against women? Evidence from physicians in China. *Journal of Development Economics*, page 103275. 1

Chen, Y. and Li, S. X. (2009). Group Identity and Social Preferences. *The American Economic Review*, 99(1):431–457. 1

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–74. 1

de Chaisemartin, C. and Ramirez-Cuellar, J. (2024). At What Level Should One Cluster Standard Errors in Paired and Small-Strata Experiments? *American Economic Journal: Applied Economics*, 16(1):193–212. 3.2

DellaVigna, S. (2018). Structural Behavioral Economics. In *Handbook of Behavioral Economics: Applications and Foundations*, volume 1, pages 613–723. North-Holland. 18

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–79. 1

Edelman, B., Luca, M., and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, 9(2):1–22. 1

Exley, C. L., Niederle, M., and Vesterlund, L. (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy*, 128(3):816–854. 1

Fitzpatrick, A. (2017). Shopping While Female: Who Pays Higher Prices and Why? *American Economic Review*, 107(5):146–149. 1, A.2

Ghulam, H., Zeng, F., Li, W., and Xiao, Y. (2019). Deep Learning-Based Sentiment Analysis for Roman Urdu Text. *Procedia Computer Science*, 147:131–135. 17

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., and Wilson, C. (2017). Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1914–1933. 1

Hedegaard, M. S. and Tyran, J.-R. (2018). The Price of Prejudice. *American Economic Journal: Applied Economics*, 10(1):40–63. 1

List, J. a. (2004). The Nature and Extent of Discrimination in the Marketplace: Evidence From the Field. *Quarterly Journal of Economics*, 119(1):49–89. 1

List, J. A., Sadoff, S., and Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14:439–457. 2.6

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22:773–793. 2.6, 14

Mehmood, F., Ghani, M. U., Ibrahim, M. A., Shahzadi, R., Mahmood, W., and Asim, M. N. (2020). A Precisely Xtreme-Multi Channel Hybrid Approach for Roman Urdu Sentiment Analysis. *IEEE Access*, 8:192740–192759. 17

Mehmood, K., Essam, D., and Shafi, K. (2019). Sentiment Analysis System for Roman Urdu. pages 29–42. 17

Michelitch, K. (2015). Does Electoral Competition Exacerbate Interethnic or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining. *American Political Science Review*, 109(1):43–61. 1

Neumark, D. (2018). Experimental Research on Labor. *Journal of Economic Literature*, 56(3):799–866. 1

OpenAI, :, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I., Berdine, J., Bernadett-Shapiro, G., Berner, C., Bogdonoff, L., Boiko, O., Boyd, M., Brakman, A.-L., Brockman, G., Brooks, T., Brundage, M., Button, K., Cai, T., Campbell, R., Cann, A., Carey, B., Carlson, C., Carmichael, R., Chan, B., Chang, C., Chantzis, F., Chen, D., Chen, S., Chen, R., Chen, J., Chen, M., Chess, B., Cho, C., Chu, C., Chung, H. W., Cummings, D., Currier, J., Dai, Y., Decareaux, C., Degry, T., Deutsch, N., Deville, D., Dhar, A., Dohan, D., Dowling, S., Dunning, S., Ecoffet, A., Eleti, A., Eloundou, T., Farhi, D., Fedus, L., Felix, N., Fishman, S. P., Forte, J., Fulford, I., Gao, L., Georges, E., Gibson, C., Goel, V., Gogineni, T., Goh, G., Gontijo-Lopes, R., Gordon, J., Grafstein, M., Gray, S., Greene, R., Gross, J., Gu, S. S., Guo, Y., Hallacy, C., Han, J., Harris, J., He, Y., Heaton, M., Heidecke, J., Hesse, C., Hickey, A., Hickey, W., Hoeschele, P., Houghton, B., Hsu, K., Hu, S., Hu, X., Huizinga, J., Jain, S., Jain, S., Jang, J., Jiang, A., Jiang, R., Jin, H., Jin, D., Jomoto, S., Jonn, B., Jun, H., Kaftan, T., Kaiser, Å., Kamali, A., Kanitscheider, I., Keskar, N. S., Khan, T., Kilpatrick, L., Kim, J. W., Kim, C., Kim, Y., Kirchner, H., Kiros, J., Knight, M., Kokotajlo, D., Kondraciuk, Å., Kondrich, A., Konstantinidis, A., Kosic, K., Krueger, G., Kuo, V., Lampe, M., Lan, I., Lee, T., Leike, J., Leung, J., Levy, D., Li, C. M., Lim, R., Lin, M., Lin, S., Litwin, M., Lopez, T., Lowe, R., Lue, P., Makanju, A., Malfacini, K., Manning, S., Markov, T., Markovski, Y., Martin, B., Mayer, K., Mayne, A., McGrew, B., McKinney, S. M., McLeavey, C., McMillan, P., McNeil, J., Medina, D., Mehta, A., Menick, J., Metz, L., Mishchenko, A., Mishkin, P., Monaco, V., Morikawa, E., Mossing, D., Mu, T., Murati, M., Murk, O., Mély, D., Nair, A., Nakano, R., Nayak, R., Neelakantan, A., Ngo, R., Noh, H., Ouyang, L., O'Keefe, C., Pachocki, J., Paino, A., Palermo, J., Pantuliano, A., Parascandolo, G., Parish, J., Parparita, E., Passos, A., Pavlov, M., Peng, A., Perelman, A., Peres, F. d. A. B., Petrov, M., Pinto, H. P. d. O., Michael, Pokorny, Pokrass, M., Pong, V., Powell, T., Power, A., Power, B., Proehl, E., Puri, R., Radford, A., Rae, J., Ramesh, A., Raymond, C., Real, F., Rimbach, K., Ross, C., Rotsted, B., Roussez, H., Ryder, N., Saltarelli, M., Sanders, T., Santurkar, S., Sastry, G., Schmidt, H., Schnurr, D., Schulman, J., Selsam, D., Sheppard, K., Sherbakov, T., Shieh, J., Shoker, S., Shyam, P., Sidor, S., Sigler, E., Simens, M., Sitkin, J., Slama, K., Sohl, I., Sokolowsky, B., Song, Y., Staudacher, N., Such, F. P., Summers, N., Sutskever, I., Tang, J., Tezak, N., Thompson, M., Tillet, P., Tootoonchian, A., Tseng, E., Tuggle, P., Turley, N., Tworek, J., Uribe, J. F. C., Vallone, A., Vijayvergiya, A., Voss, C., Wainwright, C., Wang, J. J., Wang, A., Wang, B., Ward, J., Wei, J., Weinmann, C., Welihinda, A., Welinder, P., Weng, J., Weng, L., Wiethoff, M., Willner, D., Winter, C., Wolrich, S., Wong, H., Workman, L., Wu, S., Wu, J., Wu, M., Xiao, K., Xu, T., Yoo, S., Yu, K., Yuan, Q., Zaremba, W., Zellers, R., Zhang, C., Zhang, M., Zhao, S., Zheng, T., Zhuang, J., Zhuk, W., and Zoph, B. (2023). GPT-4 Technical Report. 4.2

Phelps, E. S. (1972). The Statistical theory of Racism and Sexism. *American Economic Review*, 62(4):659–661. 1

Westfall, P. H. and Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons. 4, 5

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1:80–83. 4.2

Figure C2: Bargaining Script 2

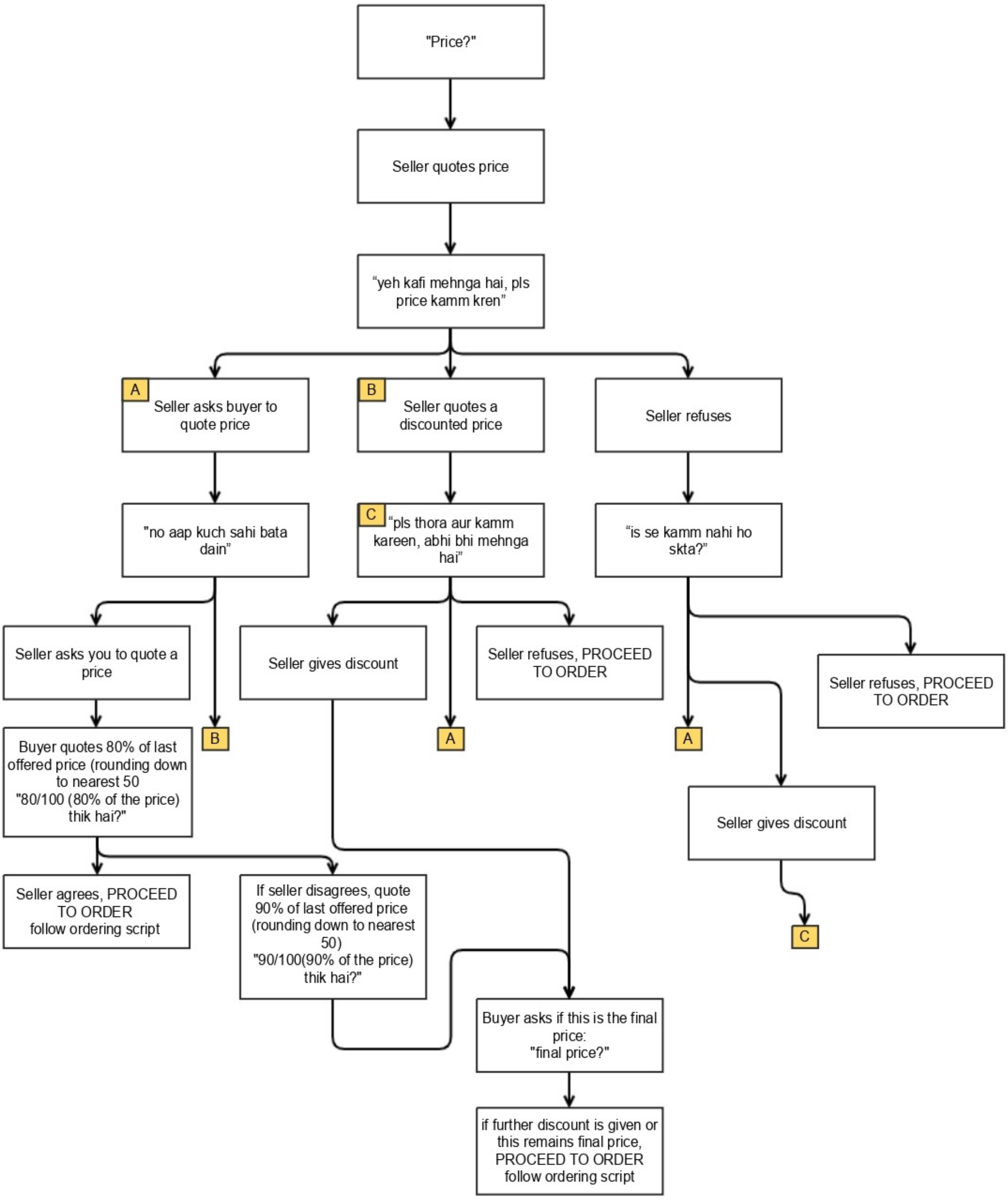


Figure C3: Ordering Script 1

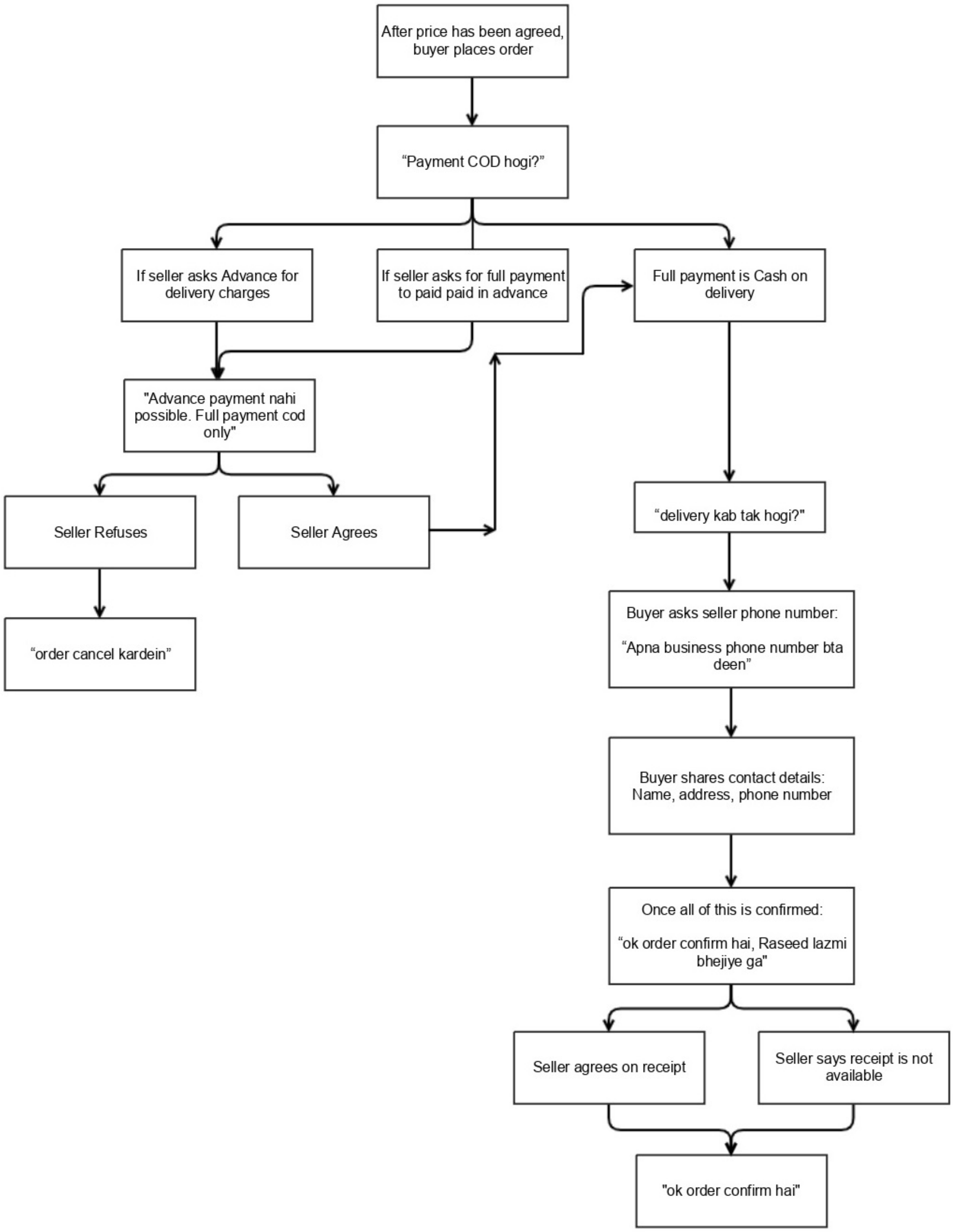


Figure C4: Ordering Script 2