From Margins to Mainstream? Gender Discrimination in Digital Marketplaces

Sher Afghan Asad* Maham Rasheed†

Abstract

Digital marketplaces may expand women's access to self-employment by lowering search, mobility, and transaction costs, but they can also reproduce gender-based discrimination through buyer behavior and platform algorithms. We study gender gaps on Facebook Marketplace in Pakistan—the country's largest social commerce platform—using a nationwide dataset of listings and seller profiles. We collect data on over 52,000 sellers and a projected one million listings across 37 divisional headquarters, and apply large language models to classify products into a harmonized taxonomy and to construct indices of listing completeness and marketability. Using these data, we document gender gaps in market participation, product selection, pricing, and listing quality, and examine how these gaps vary with market thickness, average price levels, and the extent to which transactions require in-person interaction. Our empirical analysis informs the design of a planned randomized audit experiment that will causally identify buyer discrimination by posting identical listings under male and female seller profiles, varying information disclosure and market conditions to distinguish taste-based from statistical discrimination.

Keywords: gender discrimination, digital marketplaces, market thickness, entrepreneurship, Pakistan

JEL Codes: J16, L81, O12, C93

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

†Northeastern University (rasheed.m@northeastern.edu).

1 Introduction

Persistent gender gaps in labor market participation and earnings remain a central development challenge, shaped not only by differences in skills and opportunities but also by mobility constraints, restrictive social norms, and discrimination (Jayachandran, 2021). Digital markets have the potential to relax some of these barriers by lowering search and transaction costs and enabling remote participation (Goldfarb and Tucker, 2019). Yet if the same social frictions that constrain women offline—through customer behavior or platform algorithms—are reproduced online, digitalization may offer less inclusion than often presumed.

We study gender discrimination on Facebook Marketplace in Pakistan, the country's largest social commerce platform. Pakistan is an empirically important context: women face strong mobility and norm-based constraints (Field and Vyborny, 2022; Cheema et al., 2019), and social media platforms have become a major channel for informal commerce. The project asks whether women who attempt to sell online face demand-side barriers—lower buyer interest, worse offers, or higher non-pecuniary costs—and whether these barriers vary with market conditions.

The study makes three contributions. First, it provides novel descriptive evidence on gender gaps in a large, informal digital marketplace in a developing country. While a growing literature documents discrimination on platforms such as Airbnb, Uber/Lyft, and online freelancing sites (Edelman and Luca, 2014; Ge et al., 2016; Hannák et al., 2017), evidence from social commerce platforms in low-income settings remains scarce. Second, it develops a methodological pipeline that uses large language models (LLMs) to structure the unstandardized, multilingual text and images that characterize listings on informal platforms into a consistent product taxonomy with quality indices. This approach addresses a practical challenge that has limited quantitative research on informal digital markets. Third, the empirical analysis directly informs the design of a randomized audit experiment that will causally identify discrimination and disentangle its mechanisms—a connection we discuss in detail below.

The project is organized around three core research questions: (1) Do buyers discriminate against female sellers in online product markets? (2) Is observed discrimination consistent with taste-based preferences, statistical discrimination, or algorithmic discrimination? (3) How do market thickness and competition shape the magnitude of discrimination?

This paper focuses on Phase 1 of the project: the construction and analysis of a nationwide descriptive dataset. We present the data collection strategy, the LLM-based processing pipeline, and the empirical framework for documenting gender gaps in participation, pricing, and listing quality. We also describe how patterns in the descriptive data—particularly variation in female participation across product categories and regions of differing market thickness—will guide the stratification and design of the planned audit experiment (Phase 2).

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the institutional setting. Section 4 details data collection and the LLM-based processing pipeline. Section 5 presents the empirical framework for the descriptive analysis. Section 6 outlines the audit experiment design and its connection to the empirical findings. Section 7 concludes.

2 Related Literature

This study contributes to several strands of the economics literature.

2.1 Customer Discrimination in Product Markets

A substantial body of work documents customer discrimination in labor and product markets. Foundational theories distinguish taste-based discrimination, where individuals derive disutility from transacting with members of certain groups (Becker, 1957), from statistical discrimination, where differential treatment arises from imperfect information about individual quality (Phelps, 1972; Arrow, 1973). Empirical studies have documented customer discrimination against employees in service industries (Leonard et al., 2010; Bar and Zussman, 2017), and a more recent literature examines discrimination in online settings, including racial discrimination on Airbnb (Edelman and Luca, 2014), gender and racial disparities in ride-hailing services (Ge et al., 2016), bias on freelancing platforms (Hannák et al., 2017), and discrimination in online job networks (Evsyukova et al., 2025). Our study extends this literature to social commerce, shifting attention from workers to entrepreneurs who set their own terms of trade but remain dependent on buyers' willingness to engage.

2.2 Gender, Entrepreneurship, and Digital Markets

Women entrepreneurs in developing countries face well-documented barriers including restricted mobility (Field and Vyborny, 2022; Barbanchon et al., 2021), restrictive social norms (Jayachandran, 2021; Bursztyn et al., 2020), and low returns to capital in female-run enterprises (De Mel et al., 2008). Digital platforms may relax some of these constraints by reducing the need for face-to-face interaction and expanding geographic reach (Goldfarb and Tucker, 2019). Recent evidence from Jordan suggests that providing female entrepreneurs with access to virtual storefronts and digital marketing training can increase business survival and revenue, particularly among mobility-constrained women (Alhorr, 2024). However, online platforms also expose women to new risks: Asad et al. (2025) document that female buyers on online marketplaces in Pakistan are more likely to receive unsolicited messages and that male sellers exhibit greater flirtatiousness toward female

buyers, suggesting that gender norms around women's visibility are partially mirrored online. Our study extends this evidence by examining whether women who sell on digital platforms face systematic demand-side penalties.

2.3 Market Thickness and Discrimination

Economic theory generates competing predictions about the relationship between market thickness and discrimination. In Becker's (1957) framework, competitive pressure in thicker markets should reduce the scope for taste-based discrimination, since biased buyers risk losing profitable transactions to competitors who prioritize price and quality. However, if the availability of substitutes allows biased buyers to indulge their preferences at low cost, discrimination may persist even in thick markets (Arrow, 1973; Coate and Loury, 1993). Moreover, if some buyers derive utility from engaging with female sellers in non-serious or harassing ways, market thickness may not reduce—and could even increase—the effective costs women face (Levine and Modica, 2016). We contribute to this debate by documenting how female participation and gender gaps in pricing correlate with market thickness across product categories and regions.

2.4 Algorithmic Discrimination

A growing literature examines whether platform algorithms generate disparate outcomes by gender or race, even absent intentional targeting. Lambrecht and Tucker (2019) document that an explicitly gender-neutral STEM career ad was shown to fewer women due to cost optimization in ad delivery. Imana et al. (2021) find systematic gender differences in job ad delivery on Facebook. Our study extends these concerns to product marketplaces, where ranking and recommendation algorithms may differentially expose male and female sellers' listings to potential buyers.

3 Institutional Setting

Facebook Marketplace is the dominant social commerce platform in Pakistan. Integrated directly into the Facebook app, it allows users to list products for sale, browse nearby listings, and communicate with sellers via Messenger or WhatsApp. Unlike structured e-commerce platforms such as Amazon or Daraz (Pakistan's main dedicated e-commerce site), Marketplace listings are unstandardized: sellers write free-form titles and descriptions, upload their own photos, and set prices without reference to a common catalogue.

Several features of the platform make it a valuable setting for studying gender discrimination. First, seller identity is visible through profile names and photos, making gender a salient attribute

in buyer–seller interactions. Second, the platform's default search radius of approximately 40–50 miles creates geographically defined local markets, allowing us to exploit variation in market thickness across regions. Third, the informal nature of listings—with wide variation in the quality of descriptions, images, and pricing—creates scope for both statistical discrimination (where buyers use seller gender to infer product quality) and taste-based discrimination (where buyers prefer transacting with male sellers independent of product attributes). Fourth, the platform spans a wide range of product categories, from mobile phones and electronics to property and vehicles, enabling analysis of how discrimination varies across market segments with different characteristics.

Pakistan's social and economic context amplifies the relevance of these dynamics. Women's labor force participation is among the lowest globally, shaped by strong norms around gender segregation, restricted physical mobility, and limited access to public spaces (Cheema et al., 2019; Field and Vyborny, 2022). Digital platforms represent a potentially important margin for women's economic participation, particularly in self-employment. Understanding whether demand-side discrimination limits this margin is therefore critical for both platform design and broader development policy.

4 Data

4.1 Data Collection

We construct a nationwide dataset of active Marketplace listings by querying search results around 37 divisional headquarters and the federal capital. For each listing, we collect the title, description, images, posted price (and any discounted price), item condition, location, and seller name. We then scrape each seller's Marketplace profile, recording ratings, responsiveness badges, and the history of inactive and sold listings where available.

Our preliminary data cover over 52,000 unique sellers and a projected total exceeding one million listings. The geographic scope spans all major urban centers and divisional headquarters, providing variation in population density, economic development, and social norms.

4.2 LLM-Based Data Processing

A key methodological challenge is that Marketplace listings are unstandardized, multilingual (a mix of Urdu, English, and Roman Urdu), and vary widely in the quantity and quality of information provided. We address this using a two-step LLM-based processing pipeline.

In the first step, we prompt a multilingual large language model (Llama 3.3 70B) with listing titles and descriptions to construct a structured product taxonomy. The model assigns each listing

to a product category (e.g., electronics, vehicles, home and kitchen, property) and type (e.g., mobile phone, car, sofa), and extracts product-specific attributes (e.g., make, model, year, color, storage capacity for electronics).

In the second step, we feed listing text, images, and price information into a multimodal LLM (Llama 3.2 11B / LLaVA) to extract additional structured variables and construct quality indices. Specifically, we generate: (i) transactional variables such as unit quantity, stock availability, payment method, discount, and delivery charges; (ii) a *Completeness* index measuring how much relevant product information the listing provides relative to what a typical buyer would need; and (iii) a *Marketability* index capturing how well the listing is presented in terms of description quality, image quality, and overall professional appearance.

This pipeline transforms raw, unstructured listing data into a structured analytical dataset suitable for regression analysis. It also provides the variables needed to compare listing quality across male and female sellers—a comparison that is central to understanding whether observed gender gaps in outcomes reflect differences in what sellers offer or differences in how buyers respond.

4.3 Seller Gender Classification

We classify seller gender using a multi-step pipeline applied to Marketplace profile pictures. First, we use CLIP (Contrastive Language–Image Pre-training) to classify a sample of 1,000–2,000 profile images into categories (female, male, object/logo, etc.) using text prompts. Second, we manually review and correct these initial classifications. Third, we train a logistic regression classifier on the corrected labels. Fourth, we use the trained classifier to predict gender for the remaining profile images. We iterate this process until classification accuracy is satisfactory. All processing is performed locally to ensure data privacy and ethical compliance. For sellers whose profile pictures are objects, logos, or otherwise uninformative, we supplement with name-based gender inference using established Urdu and English name dictionaries.

5 Empirical Framework

The descriptive analysis addresses three sets of questions, each building toward the design of the audit experiment.

5.1 Female Participation Across Products and Regions

We first document where women are more and less represented as sellers, estimating:

$$\text{Female}_s = \alpha + \lambda_c + \mu_r + \varepsilon_{iscr} \tag{1}$$

where Female_s is an indicator for whether seller s is female, λ_c denotes category fixed effects, and μ_r denotes region fixed effects, for listing i in category c and region r. Observations are weighted by $w_{iscr} = 1/n_{scr}$ (the inverse of the number of listings by seller s in category c and region r), and standard errors are clustered at the seller level.

We then examine what characterizes markets with higher female participation:

$$\text{Female}_s = \alpha + \beta_1 \text{AvgPrice}_{cr} + \beta_2 \text{Thickness}_{cr} + \beta_3 \text{EntryBarrier}_c + \varepsilon_{iscr} \tag{2}$$

where AvgPrice_{cr} is the median price in category c and region r (a proxy for market profitability), Thickness_{cr} is a measure of competition intensity defined as $1 - HHI_{cr}$ (where $HHI_{cr} = \sum_s \text{share}_{scr}^2$ is the Herfindahl–Hirschman Index of seller concentration), and EntryBarrier_c is a mobility index equal to one if transactions in category c typically require physical inspection or in-person handover.

These participation regressions serve a dual purpose. Descriptively, they reveal whether women sort into particular market segments—for instance, lower-priced categories, less competitive niches, or product types that can be transacted remotely. For the audit experiment, they identify the category $\times$ region cells that exhibit meaningful variation in both female participation and market thickness, guiding the stratification of experimental listings.

5.2 Gender Gaps in Pricing and Listing Quality

We next estimate gender gaps in listing characteristics:

$$\ln(Y_{iscr}) = \beta_0 + \beta_1 \text{Female}_s + \beta_2 X_{iscr} + \beta_s + \beta_c + \beta_r + \varepsilon_{iscr} \tag{3}$$

where $Y \in \{\text{Price}, \text{Completeness}, \text{Marketability}\}$ captures listing-level outcomes, and X_{iscr} includes product attributes (e.g., make, model, condition) for the price regression. Standard errors are clustered at the seller level.

The coefficient β_1 captures the conditional gender gap in each outcome. If female sellers list at lower prices conditional on product attributes, this could reflect strategic pricing (anticipating lower demand), different reservation prices, or different product quality within categories. If female sellers score lower on completeness or marketability indices, this could indicate differences in digital literacy, effort, or strategic choices about information disclosure. Importantly, these descriptive patterns are not themselves evidence of discrimination—they capture supply-side differences that may or may not reflect demand-side constraints. However, they inform the audit experiment by identifying whether and where male and female sellers differ in listing quality, which is essential for constructing experimental listings that hold quality constant.

6 Audit Experiment Design

The descriptive analysis documented above directly informs the design of a randomized audit experiment that will provide causal evidence on buyer discrimination. We describe the planned design here; the experiment will be implemented after the descriptive phase is complete.

6.1 Overview

The audit experiment posts identical listings under seller profiles that signal different genders. The design randomizes three dimensions: (i) seller gender signal (male vs. female), (ii) information disclosure (minimal vs. complete product information), and (iii) market thickness (listings are posted across category $\times$ region markets that vary in competition intensity). This factorial structure allows us to test for discrimination, distinguish its mechanisms, and examine heterogeneity by market conditions.

6.2 Gender Signal

Seller profiles signal male or female gender through two channels: gendered names and AI-generated profile photographs. We create "twin" profile pairs in which male and female profile images are generated from the same base image, varying only gender-specific attributes while holding constant age, facial expression, camera angle, skin tone, hair color, and clothing colors. Following Evsyukova et al. (2025), these images are validated through a crowdsourced survey in which respondents assess each image for perceived gender, authenticity (distinguishing real from computer-generated photos), and attributes including trustworthiness, intelligence, friendliness, and attractiveness. Only image pairs that are correctly identified by gender, perceived as realistic, and rated comparably on non-gender attributes are used in the experiment. Each twin pair is assigned a unique Facebook account, phone number, and WhatsApp account.

6.3 Information Treatment

To distinguish statistical from taste-based discrimination, we experimentally vary the amount of product information in each listing. "Minimal" listings include a brief title, a small number of photos, and a short description. "Complete" listings include a detailed title, multiple high-quality photos, and a comprehensive description covering all product attributes (condition, specifications, provenance, etc.) with identical prices. If gender gaps in buyer response narrow when complete information is provided, this pattern is consistent with statistical discrimination: buyers use seller gender as a proxy for unobservable product quality, and detailed information reduces the scope for

such inference (Phelps, 1972). If gender gaps persist despite complete information, the pattern is consistent with taste-based discrimination (Becker, 1957).

6.4 Market Thickness

The descriptive analysis provides measures of market thickness—defined as competition intensity at the category $\times$ region level—that guide the stratification of experimental listings. Products are assigned to markets stratified by thickness, enabling tests of whether discrimination attenuates in more competitive settings (as predicted by Becker's theory) or persists because substitutability allows biased buyers to discriminate at low cost.

6.5 Outcomes

We will measure three classes of outcomes. *Pecuniary outcomes* include contact rates, offer rates, negotiated prices conditional on contact, and conversion to sale. *Non-pecuniary outcomes* include unsolicited friend requests, unwanted phone calls, message tone, and harassment indicators, measured through systematic text analysis of buyer conversations. *Exposure outcomes* include listing views and ranking position (where observable), which capture potential algorithmic discrimination in how the platform distributes visibility across male and female sellers.

6.6 Connection to the Descriptive Analysis

The empirical phase contributes to the audit experiment design in several specific ways. First, the participation regressions identify which product categories and regions exhibit meaningful variation in both female representation and market thickness, guiding the selection of experimental markets. Second, the pricing and quality regressions reveal the extent to which male and female sellers differ in listing characteristics within categories, informing the construction of experimental listings that credibly hold quality constant. Third, the product taxonomy and attribute extraction provide the structured data needed to select representative products and to write standardized descriptions at both information levels. Fourth, the completeness and marketability indices establish benchmarks for what "minimal" and "complete" listings look like in each category, ensuring the information treatment is ecologically valid. In this sense, the descriptive analysis is not merely a preliminary exercise but an integral component of the experimental design.

7 Conclusion

This paper presents a research design for studying gender discrimination on Facebook Marketplace in Pakistan, the country's dominant social commerce platform. The project combines a large-scale descriptive analysis of existing marketplace data with a planned randomized audit experiment.

In the empirical phase—the focus of this paper—we construct a nationwide dataset covering over 52,000 sellers and a projected one million listings, and develop an LLM-based pipeline to structure the unstandardized, multilingual data into a consistent product taxonomy with quality indices. We document gender gaps in market participation, pricing, and listing quality across product categories and regions, and examine how these gaps correlate with market characteristics including thickness, price levels, and the extent to which transactions require physical interaction.

These descriptive findings will directly inform the design of the audit experiment, which will post identical listings under male and female profiles to causally identify discrimination and its mechanisms. By varying information disclosure across listings, the experiment will distinguish taste-based from statistical discrimination. By stratifying listings across markets of varying thickness, it will test whether competitive pressure attenuates or sustains discriminatory buyer behavior. By tracking listing visibility metrics, it will assess whether platform algorithms systematically disadvantage female sellers.

Understanding demand-side barriers to women's participation in digital markets is critical for both platform design and development policy. If discrimination is belief-based, interventions that reduce ambiguity—such as standardized listing templates, trusted-seller badges, or enhanced reputation systems—may help level the playing field. If discrimination is taste-based, broader competition and internal trade opportunities may matter more, but additional safeguards against harassment may also be necessary. If algorithmic bias contributes to gender disparities, platform design changes—such as adjusting ranking algorithms to avoid systematically penalizing female-led sellers—could improve outcomes. The evidence generated by this project will inform these choices.

References

Alhorr, L. (2024). Virtual windows through glass walls? Digitalization for low-mobility female entrepreneurs. *Policy Research Working Paper*, World Bank.

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets* (pp. 3–33). Princeton University Press.

Asad, S. A., Ahmad, H. F., and Majid, H. (2025). Price and prejudice: Gender discrimination in online marketplaces. *Journal of Development Economics*, 177, 103540.

Bar, R. and Zussman, A. (2017). Customer discrimination: Evidence from Israel. *Journal of Labor Economics*, 35(4), 1031–1059.

Barbanchon, T. L., Rathelot, R., and Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *Quarterly Journal of Economics*, 136(1), 381–426.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10), 2997–3029.

Cheema, A., Khwaja, A. I., Naseer, M. F., and Shapiro, J. N. (2019). Glass walls: Experimental evidence on access constraints faced by women. Technical report.

Coate, S. and Loury, G. C. (1993). Will affirmative action policies eliminate negative stereotypes? *American Economic Review*, 83(5), 1220–1240.

De Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, 123(4), 1329–1372.

Doleac, J. L. and Stein, L. C. D. (2013). The visible hand: Race and online market outcomes. *Economic Journal*, 123(572), 469–492.

Edelman, B. G. and Luca, M. (2014). Digital discrimination: The case of Airbnb.com. *SSRN Electronic Journal*.

Evsyukova, Y., Rusche, F., and Mill, W. (2025). LinkedOut? A field experiment on discrimination in job network formation. *Quarterly Journal of Economics*, 140(1), 283–334.

Field, E. and Vyborny, K. (2022). Women's mobility and labor supply: Experimental evidence from Pakistan. *ADB Economics Working Paper Series*, No. 655.

Ge, Y., Knittel, C. R., MacKenzie, D., and Zoepf, S. (2016). Racial and gender discrimination in transportation network companies. *NBER Working Paper* No. 22776.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1), 3–43.

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., and Wilson, C. (2017). Bias in online freelance marketplaces: Evidence from TaskRabbit and Fiverr. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1–19.

Imana, B., Korolova, A., and Heidemann, J. (2021). Auditing for discrimination in algorithms delivering job ads. *Proceedings of the Web Conference*, 3767–3778.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, 69, 576–595.

Lambrecht, A. and Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science*, 65(7), 2966–2981.

Leonard, J. S., Levine, D. I., and Giuliano, L. (2010). Customer discrimination. *Review of Economics and Statistics*, 92(3), 670–678.

Levine, D. K. and Modica, S. (2016). Peer discipline and incentives within groups. *Journal of Economic Behavior and Organization*, 123, 19–30.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62(4), 659–661.

Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. *Innovation Policy and the Economy*, 10, 57–81.