Customer Discrimination in the Workplace: Evidence from Online Sales*

Erin Kelley,† Gregory Lane,‡ Matthew Pecenco,§ & Edward Rubin¶

May 25, 2022

Abstract

Discrimination by customers who prefer to interact with certain types of workers can affect worker productivity. In this paper, we measure the impact of gender-based customer discrimination on the productivity of online sales agents working across Sub-Saharan Africa. Using a daily randomization design that varies the gender of names presented to customers while holding other characteristics fixed, we find the assignment of a female-sounding name leads to significantly fewer purchases by customers. The results appear to be driven by relatively lower interest in engaging with female workers. We find no evidence of differential bargaining or harassment.

*Funding for this project was graciously provided by PEDL. Borui Sun, Victoria Yin, Edwin Kasila and Steven Wandera provided excellent research assistance. All views expressed are those of the authors and do not necessarily reflect the opinions of any of the funding or data providing organizations. AEA RCT identification number: 0006698.

†Development Impact Monitoring Evaluation Unit, World Bank erinmkelley@worldbank.org

‡Department of Economics, American University glane@american.edu

§Department of Economics, Brown University matthew_pecenco@brown.edu

¶Department of Economics, University of Oregon edwardr@uoregon.edu

1 Introduction

Women work less outside the home, earn lower wages, and run less profitable businesses, particularly in low-income countries (UN, 2022). Recent research focuses on understanding the sources of these disparities—e.g., discrimination, differences in productivity or skill, occupational choice, cultural norms (Jayachandran, 2015; Blau and Kahn, 2017a). Yet, customer-based discrimination—customers exerting preferences for/against certain types of workers—has received relatively little attention. If consumers prefer men to women, these preferences may lead them to purchase fewer products from women—or bargain more forcefully with women—leading to lower female productivity and ultimately lower female wages, promotions, and job prospects. This source of discrimination is potentially large and persistent (Becker et al., 1971; Bartlett and Gulati, 2016), as consumer-based discrimination may not disappear via market competition. Disentangling the precise role of customer-based discrimination is challenging because many factors affect worker productivity—e.g., the worker's own skills and behavior, customers' behaviors and preferences, and the workplace environment.

We overcome these challenges through a randomized experiment with an online sales company in Sub-Saharan Africa. We study workers—specifically sales agents—who chat with customers online to answer questions and increase sales. Two aspects of the context provide a novel framework for estimating the causal effect of customer-based discrimination. First, the names of workers—and implied genders—were randomized daily, providing plausible variation in customer beliefs about worker gender.[1] Customers could only infer agents' gender from their names, as they did not receive any other information about the agent. Second, while agents were aware of the experiment, they were unaware of their particular assigned name due to a web plugin that masked the assigned name from their view.

[1] Changing the names of workers appears to be relatively common in online sales settings (e.g., LiveAgent).

This step ensured that agents' behavior was not directly affected by their name assignment. Consequently, any change in consumer behavior towards sales agents could only occur if consumers responded to the randomly assigned names.

This research design is unique and marries the advantages of two common methods to study employer-based discrimination: audit and correspondence studies. Audit studies recruit individuals who share many salient characteristics except one to apply to or interview for jobs in person. This generates detailed outcomes through the resulting in-person interactions. However, the chosen actors in these studies may differ in many ways, making it difficult to isolate the impact of a specific characteristic. Correspondence studies send fictitious applications to employers. The fictitious identities enable clear, causal comparisons but are typically unable to measure the same set of detailed interactions. The name randomization in our context of online workplace interactions overcomes these challenges. Agents' names were randomized daily, generating clear counterfactuals: we compare the same agent to themselves when randomly assigned to male- and female-associated names. This setup also provides an opportunity to collect detailed outcomes, including overall outcomes—e.g., the likelihood of purchase—and specifics of the interaction, e.g., bargaining behavior.

The context of this study is important for three reasons. First, barriers to female labor force participation are higher in low-income countries (Jayachandran, 2015), and reducing gender-based pay differentials are central policy goals for governments and international institutions alike (Bank, 2011; O'Donnell et al., 2020).[2] Second, in Sub-Saharan Africa, many social norms favor men over women as economic agents and business. For example, data from The World Bank Development Indicators (WDI) show that households in Sub-Saharan Africa are more likely to agree that men make better business leaders and that women have no say in decisions on

[2] For example, The African Union Strategy for Gender Equality includes the promotion of laws to achieve pay equality—as the gender wage gap is persistent across most industries where women work (ILO, 2019)

large household purchases (Jayachandran, 2015). Such norms may also contribute to higher rates of customer discrimination. Finally, the service sector is growing across the continent and issues related to gender-based differences in customer interactions are increasingly relevant.

We find that randomly assigned female names reduce the likelihood that customers make any purchase, the number of purchases, and the value of the purchases. Specifically, the likelihood of any purchase decreases by 3.9 percentage points—a substantial effect given the low baseline purchase rate (6.2%). We observe similarly large reductions in the total number of purchases and the total value of goods purchased.

These results suggest that productivity differences between workers are not just a reflection of worker attributes in this context but also of differential customer responses. While the literature has identified productivity differences as a key determinant of the gender wage gap (Sin et al., 2020; Gallen et al., 2017; Blau and Kahn, 2017b; Caliendo et al., 2017), our results demonstrate that underlying customer-based discrimination can drive these productivity differences. This finding has important implications for policies such as "equal pay for equal work," because the concept of "equal work" may not account for the difficulties women face while doing the same job. Government policies may mitigate the impact of customer-based discrimination by eliminating employers' discriminatory hiring and wage practices. However, incentive-based pay schemes commonly found in many customer-facing roles may still lead to disparate outcomes due to customers' discrimination.

To confirm that our treatment effects result from the implied gender of agents' assigned names, we investigate whether customers are aware of the agents' names. In 11% of chats, customers mention the agent's assigned name—suggesting names are a salient feature of the interaction. Helpfully, we can also rule out a potential confound—customers responding to a "mismatch" between the gender implied by agents' assigned name and the agents' actual gender. This confound cannot explain our results because every agent in our sample is female, and any mismatch would come from being assigned a male-sounding name. Consequently, this "mismatch effect" would attenuate our estimated effect of discrimination against female agents.[3] In sum, our main results suggest that receiving a female name reduces productivity via consumers' gender preferences.

These reductions in worker productivity may be explained by several mechanisms: general customer disinterest in working with female agents, differential bargaining, or overtly negative interactions.[4] Data from agent-customer chat interactions suggest that customer disinterest is the most likely channel. We find that consumers respond more slowly to female agents—only responding after receiving additional messages from the agent.[5] This result suggests some consumers are hesitant to engage with female agents unless the agents persist through additional messaging. We also find that consumers are less likely to express any tone, which we interpret as another measure of engagement with the agent.

The data do not support other possible mechanisms. We find no evidence that consumers differentially bargain when agents receive female-sounding names. This result is interesting as bargaining is common in these interactions (occurring in 15% of conversations), and differences in bargaining by gender feature prominently in studies of wage gaps and job-application behavior (e.g., Card et al. (2016); Rousille (2021); Castillo et al. (2013)). We also find no differences in hostile or harassing behavior—although any form of harassment is rare in this context.[6]

This experiment's results are substantially different from the correlation we observe between an agent's gender and their productivity, which shows no difference. This suggests that while the company hired males and fe-

[3] More broadly, roughly two-thirds of the employees working in these roles at the company are female.

[4] These behaviors are consistent with various theories of discrimination, including statistical and taste-based discrimination. Differentiating between these theories is not the intent of this study.

[5] Agents always send the first message. The assigned name of the agent is revealed with this first message.

[6] We cannot assess possible mechanisms such as homophily in gender-gender interactions because we do not observe consumers' identities.

males with similar levels of productivity, the productivity metrics do not account the customer discrimination women face.

This paper makes three contributions. First, we add to the existing literature on the effect of discrimination in the labor market. Most of these studies focus on employer discrimination using correspondence studies in high-income countries (Bertrand and Duflo, 2017; Baert, 2018; Bertrand and Mullainathan, 2004). The relatively smaller set of studies on customer-based discrimination often test for racial or ethnic discrimination in labor market settings using cross-sectional variation in consumer attributes across establishments (Leonard et al., 2010; Holzer and Ihlanfeldt, 1998; Bar and Zussman, 2017; Kahn and Sherer, 1988; Combes et al., 2016).[7] While these studies benefit from studying diverse labor market settings, they have been in high-income contexts and are not able to fully control for differences in unobservable characteristics between consumers—introducing the potential for bias. As in Doleac and Stein (2013), we circumvent unobservables issues by using an online sales context in which we randomly assign seller attributes to consumers. We show that gender-based discrimination represents a meaningful impediment to women's labor market outcomes in a low-income country context.

Second, this paper contributes to a growing literature identifying women's labor-market barriers in low-income countries. Recent work documents a variety of constraints: norms and bargaining dynamics within the household (Lowe and McKelway, 2021; Bursztyn et al., 2020; Dean and Jayachandran, 2019; Heath and Tan, 2020; Field et al., 2021; McKelway, 2021a,b); workplace attributes (Subramanian, 2021), safety during commutes (Borker, 2021), market demand Hardy and Kagy (2020), and employer discrimination (Jayachandran, 2015; Duflo, 2012; Sin et al., 2017). We add to this literature by showing how consumers' preferences can create potentially important—and understudied—barriers to women's success in labor mar-

[7] Related work also finds that *employer*-based discrimination may be driven by *customers'* preferences: (Kline et al., 2021) find that employer-based discrimination is higher in consumer-facing roles; (Neumark et al., 1996) illustrate how restaurants discriminatory hiring practices could be driven by their patrons' preferences.

kets. We discuss three ways these preferences may manifest (bargaining, harassment, and disinterest) and their potential impact on female worker productivity. A related paper by Delecourt and Ng (2021), uses an audit-study approach to show that customers do not discriminate against female-led small businesses. We use our novel framework to test for customer discrimination among employees and find large impacts.

Finally, we contribute to a growing policy discussion about workplace gender equity in low-income countries (World Bank, 2013). While some companies have attempted to equalize opportunities for women and men (e.g., better wages and flexible hours), customer discrimination can still significantly affect women's productivity in the workplace. Our results provide an opportunity to design new solutions that mitigate the effect of customer preferences.

2 Context

2.1 Service sector in Sub-Saharan Africa

We study consumers' discriminatory behavior when engaging with online sales agents in Sub-Saharan Africa. As service-sector jobs increase, customer-facing roles are increasingly common across the continent. For example, the share of working-age individuals employed in services throughout Sub-Saharan Africa rose 12% from 2011 to 2019 (WDI). Women largely drove these trends: the share of working-age women employed in services increased by 16% over the same period—currently at 39.7%.

This trend will likely persist as internet connectivity spreads across the continent and service-sector jobs increasingly interface with clients online. In 2010, only 8.3% of Africans had internet access. By 2017, access had increased to 22.3% (WDI). Online shopping, in particular, has increased 18% annually between 2014 and 2017 (UNCTAD, 2018)—and estimates suggest that almost 50% of digital buyers in Africa are female (Statista, 2019). The COVID-19 pandemic has likely accelerated these trends as consumers increasingly head online. Reflecting this growth and importance: in 2020, the value of African e-commerce was estimated at 20 billion USD—a 42% increase over 2019 (IFC, 2021).

2.2 Company and study details

We evaluate an experiment at an online sales company based in Sub-Saharan Africa. The company employs sales agents who answer customer questions, field complaints, and increase purchases.[8] In the company, approximately two-thirds of the sales agents are women, and these women account for 83% of chats. The experiment included six agents, all of whom were female.[9] The company primarily sells tourism-related products. Consequently, products are relatively standardized, non-gendered, and highly valued relative to other online products.

The company provides sales agents with a chat interface to interact with customers. Customers can initiate interactions with sales agents by clicking on a chat button at the bottom of the webpage. Clicking the chat button reveals a chat window displaying the agent's first name and a short greeting message. Thus, agents always send the first message; either the agent or the customer can send subsequent messages as the conversation evolves.

The company was keen to partner with the research team to investigate whether they could optimize this chat/sales interface. This particular test aimed to identify how customer behavior changed with respect to agents' identities—specifically when agents were assigned male- versus female-sounding names. To this end, the company needed to (1) randomize whether the name appearing in the chat implied a male or female identity and (2) ensure agents were unaware of their assigned names. To ensure the randomization was correctly implemented, a software program pulled one name per sales agent per day from an existing list. The agent then received the drawn name within the chat/sales interface. A local field team com-

[8] 40-50% of chats with any customer response relate to purchases.

[9] We do not have access to other information about agents' demographics or wages within the company.

piled the list of names by drawing 1,198 names from local school yearbooks and assigned each name an implied gender. To limit the customers' inference of other dimensions of agents' identities, the interface only included only agents' (randomly assigned) first names.

Next, to ensure that agents could not see the names that were assigned to them, a web plugin was designed to omit the agent's name from the agent-facing interface. The company installed the plugin on each agent's internet browser with oversight from our field team. The plugin symbol was removed from the list of visible extensions—appearing as a light grey square when all browser extensions were listed. The plugin worked in the following way. Consider a day when agent *James* (real name) was assigned the name *Steve*. Whenever the customer typed "Steve" into the chat, James would only see "Agent" in his chat window. In contrast, the client would still see "Steve." This name masking included any references to the agent's assigned name in the chat transcript.[10]

2.3 Data

The analysis relies on two sources of administrative data. The first dataset records every purchase made by customers, including the sale amount. The second dataset contains the agent-customer chat interactions: the full chat transcript, a timestamp for each message, and the customer's country.[11] The sales data were matched to the chats using date and customers' IP address.[12]

To measure overall purchases, we include purchases directly made by customers and purchases made by agents on behalf of customers. Agents

[10] The vast majority of interactions occurred in English, removing concerns about gendered identifiers.

[11] We only know customers' approximate locations—we do not have access to any customer demographic data. However, estimates from other sources suggest that in 2019 nearly 50% of digital buyers in Africa were female (Statista, 2019).

[12] We restrict to observations with fewer than five previous purchases as some users may access the site using non-unique locations—e.g., public areas or businesses. This restriction retains 98% of observations.

may input customer details and purchase products on their behalf, which we capture by reading through the chat records and flagging instances where agents send final purchase confirmation details to customers. Customers then pay separately or at the time of receiving the order. When agents make purchases, we cannot measure purchase values. We do not include these observations in our analysis of the total value of purchases.

From the chat transcripts, we create objective and subjective outcome measures. Objective measures do not require human interpretation—for instance, whether a purchase occurred. Subjective measures represent outcomes that require human interpretation of the chat content—e.g., the overall tone, whether customers bargained with agents (e.g., asking for a discount), or whether customers harassed agents. Enumerators familiar with the cultural context hand-coded these subjective outcomes; 20% of the observations were double coded to ensure consistent measurement.

Six female sales agents worked during the study. Agents' jobs involved several sales-related activities, including assisting customers via online chat and phone. Each agent worked the chat interface six weekdays per month on average. On days when agents responded to chats, they spent 2.9 hours on the online sales interface with customers, engaging in approximately 8 unique chat conversations per day. The average chat lasted 22 minutes and contained 73 words. The sales agents did not all work during the full study period for institutional reasons, although they all worked a majority of the time.

3 Empirical strategy

The design of this study overcomes two major challenges to identifying the causal effect of customer-based gender discrimination. First, daily randomization of agents' names ensured customers were randomly exposed to female- or male-sounding names. This separates unobserved factors that correlate with gender from customers' perceptions of gender. Second, agents were not aware of the name consumers see—any revelation of the agent's name during the chat was masked automatically by a computer program and was not seen by the agent.[13] Therefore, agents' behavior cannot directly respond to the randomized name—only to customers' responses to the names. Together, these elements allow us to test for customer-based gender discrimination.

Treatment assignment occurred as follows. Six female agents were randomly assigned 'male' or 'female' each day (with replacement). Given the selected gender, the randomization then chose a specific name from the name database. This procedure occurred every day of the study period. The number of agents working varied daily. Some days, only one agent worked; other days, multiple agents operated the chat. We restrict our sample to weekdays when agents typically work regular schedules.

Using this randomization, we estimate the effect of customer discrimination on worker productivity. Our main specifications take the following form:

$$y_{iam} = \beta\, \mathbb{1}[\text{Assigned female}]_{ia} + \gamma_{am} + X_i + \varepsilon_{iam}$$

where y_{iam} is the outcome of interest for customer i, working with agent a, in month m. The indicator $\mathbb{1}[\text{Assigned female}]_{ia}$ is 1 if agent a (matched to customer i) is assigned female in that period. The term γ_{am} represents agent by month-of-sample fixed effects—restricting comparisons within this grouping. Accordingly, the research design compares (1) a consumer who chats with agent a in month m on a day when the agent was assigned female to (2) a consumer chatting with the same agent in the same month when the agent was assigned male. This specification flexibly controls for unobservable agent differences over time, which is relevant as not all agents work for the full study period and purchasing patterns vary over time. We further control for customer characteristics, X_i, including country, past purchase history, and past chat history for precision. We can augment this regression specification to estimate individual-agent treatment effects, but we

[13] See section 2.2 for details.

cannot estimate heterogeneity by customer gender or other demographics, as we do not observe them.[14]

Customers may have multiple interactions with agents in the same day if they are disconnected or return to ask additional questions. We account for this possibility in two ways. First, we two-way cluster our standard errors at the agent-day (the level of randomization) and customer-day levels. Second, we assign the customer the treatment status of their first chat of the day. This circumvents the possibility that customers can affect their treatment status by returning to chat with an agent of a different gender.

This approach contains two potential concerns for external validity. First, the name-masking procedure could affect agents' productivity. For example, a male customer who believes the agent is female may banter with the agent. This behavior could confuse a male sales agent unaccustomed to being the target of this type of attention. This concern does not threaten our identification of the effect of agents' gender, but it may create interactions that are less reflective of reality. This situation is unlikely in our context because agents have little information about customers and thus may infer little about customer behavior. Second, agents have certain gender-specific language that could appear strange to consumers when assigned the opposite gendered name. For example, a male agent may use specific language that will confuse a customer who assumes they are speaking with a woman because of their female-sounding name—and this may reduce the chance of a sale. All of the agents in our sample are women, and could only potentially 'confuse' a customer with their language when they are assigned a male sounding name. However, we find that being assigned a female name reduces the likelihood of a sale, any 'confusing' behavior from a male-sounding would only attenuate our estimates.

We provide a validation of the randomization procedure in Table 1. In column (3) of this table, we regress observable customer characteristics (e.g.,

[14] Note that the regression can analogously be run at the (grouped) agent-day level after employing a two-step regression procedure that matches our customer (microdata) approach (Angrist and Pischke, 2008). The focus of the paper on customer behavior and the parsimony of the current approach motivates the customer-level analysis.

number of past purchases) and agent characteristics (e.g., number of daily chats) on an indicator for whether the agent received a female name, controlling for agent-by-month fixed effects. Female assignment does not correlate with any customer or agent characteristics at the 5% level. We fail to reject the joint null hypothesis that each of these effects is zero ($p = 0.65$). The table includes an additional row that identifies whether the customer mentioned the agent's actual name. This event occurs very rarely (mean is $<$0.01) and likely results from agents' names coincidentally matching a topic in the chat.

4 Results

4.1 Effect of name assignment

The experiment aims to identify the impact of gender on consumer behavior. This strategy requires that consumers pay attention to agents' assigned names. We confirm that customers notice agents' names by measuring how often consumers use agents' assigned names in chats. This test provides a lower bound for consumers' awareness of agents' names—and likely the names' implied genders. In our study sample, customers used agents' assigned names in 7% of all chats and 11% of chats in which consumers ever responded to agents' initial messages. We interpret this as a relatively high share of customer awareness as many chats are brief. Thus, agent names are indeed salient in chat interactions and could affect customers' behavior.

Table 2 presents the estimated effects of female-name assignment on outcomes related to customer purchases. We measure purchases within 24 or 48 hours of the chat to capture behavior plausibly related to the chat interactions rather than unrelated interactions that happen later.[15] We measure purchases in three ways: the probability of making any purchase, the number of distinct purchases, and the total price of purchases. As discussed in subsection 2.3, our measures of *any purchase* and *total number of purchases*

[15] Statistical power is also likely to be higher in the period directly after these focal events.

include purchases made by customers and by agents on customers' behalf. In contrast, the total-price measure only includes purchases by customers.

We find that consumers assigned to agents with female names are less likely to purchase products on the website. Column (1) shows that female-agent assignment decreases the probability that any purchase occurs (within 48 hours) by 3.9 percentage points ($p = 0.002$). The likelihood that a chat results in any purchase in the control group (male-sounding names) is only 6.2%. Thus, the point estimate implies a 63% reduction in the likelihood of making a sale. Column (2) shows that consumers also purchase 0.039 fewer total products ($p = 0.004$) when interacting with female-sounding names; column (3) shows that the total value of purchases falls by 3.5 euros. Columns (4-6) repeat the same outcomes but use a 24-hour window after the chat. The results are very similar (the 24-hour results may be slightly attenuated estimates by missing some consumer behavior).[16]

These results highlight the importance of customer-side discrimination in productivity differences between women and men in the workplace (for consumer-facing roles). Prior research on the gender wage gap suggests women receive lower pay partly because they are less productive (Sin et al., 2020; Gallen et al., 2017; Blau and Kahn, 2017b; Caliendo et al., 2017). We show that discriminatory behavior—on the part of consumers—can drive these productivity differences. In our context, for women and men to have similar productivity levels, women would need to overcome significant barriers created by consumers' behavior. These results also suggest that piece-rate wage structures—i.e., rewarding employees for their output levels—could further workplace inequality. While we cannot speak to optimal policy responses, regulation that prevents employer discrimination could limit some consequences of customer discrimination.[17]

[16] We also test for dynamic effects of female name assignment. We do not find evidence for this; the p-value of the joint test of the assignment to a female name in the previous two working days does not reject the null hypothesis of no effect either individually or jointly ($p = .377$).

[17] More broadly, wherever consumer-based discrimination can be monitored—e.g., online settings—-administrators may remove violators.

Because the study design assigns each agent to both treatment statuses (female and male) over time, we can test whether treatment effects vary across agents. If heterogeneity exists across agents, then policies that attempt to compensate employees for customer discrimination would likely need to take this heterogeneity into account. To estimate agent-specific treatment effects, we augment the baseline model by interacting treatment with agent-specific indicators. Agent-specific estimates may reflect (1) differences in agent characteristics and/or (2) the types of consumers that agents encounter—since the study's design does not randomize customers across agents. While the estimated effect of female-name assignment is negative for all agents (except one, whose positive coefficient is not statistically different from zero), we can reject that the treatment effects are the same across all agents ($p = 0.018$).[18] This result suggests that the impacts of consumer-based gender discrimination on productivity (sales) likely differ across agent and/or consumer types.

Our results are robust to various analysis choices. In particular, they are qualitatively and quantitatively similar when we remove customer controls (Table A2) and aggregate to the agent-day level (Table A3)—and remain statistically significant if we restrict to purchases in administrative records (Table A4).

4.2 Mechanisms

There are many reasons why purchases may fall when consumers chat with female agents. Our data allow us to explore three potential mechanisms. First, customers may be hesitant to engage with female sales agents because of taste-based or statistical discrimination. For instance, customers may dislike working with women or believe women are less efficient at helping with purchases. Second, recent work suggests women are more likely to face harassment and verbal abuse on the job—likely harming their productivity (Georgieva, 2018; Dupas et al., 2021; Folke and Rickne, 2020). Finally,

[18] Table A5 shows these agent-specific treatment effects.

an extensive literature documents women and men may face different bargaining processes (Ashraf, 2009; Rousille, 2021; Vesterlund, 2018; Castillo et al., 2013; Card et al., 2016)—a fact customers may attempt to exploit by bargaining more with female sales agents.

We first explore whether customers are hesitant to engage with female agents. We investigate this along two dimensions. On the extensive margin—whether the customer engages with an agent at all—some consumers may be hesitant to chat with female agents or may entirely avoid female agents. On the intensive margin, consumers may engage differently by using different tones when they chat with female agents.

Columns (1-3) of Table 3 show the effect of female-name assignment on *extensive margin* consumer interactions. Mechanically, agents always send the first message; the conversations begin there. In column (1), female assignment leads to a negative but statistically insignificant effect on the likelihood the customer ever responds ($p = 0.184$).

However, agents can send multiple messages to customers to encourage their response, which means that measuring a binary variable of any response by the customer may not fully capture a lack of engagement. Column (2) shows that female-assigned agents send more messages before receiving a response ($p = 0.022$), suggesting lower customer engagement (higher hesitance). Finally, we test engagement using the number of messages the customer sends in their response to the agent. Column (3) shows that consumers send fewer messages when initiating a conversation with an agent with a female-sounding name ($p = 0.059$). Together, these three measures suggest that some consumers may hesitate to engage with female agents.

We investigate customer hesitancy along the *intensive margin* by analyzing the conversations' tones. While specific tones are likely imperfect proxies for genuine emotions, whether a tone exists may reflect a customer's level of engagement with the agent.[19] To this end, we construct a measure for any non-neutral tone detected in the conversation. Enumerators manually reviewed each chat and tagged whether the tone was neutral

[19] Table A6 also presents results for each tone separately.

and whether any bargaining or harassment occurred. Column (1) of Table 4 demonstrates a 2.7 percentage point reduction in the probability of any tone when customers engage with female-assigned agents, a 31% reduction relative to the control-group mean ($p = 0.053$). This result again suggests that customers exhibit weaker levels of engagement with female-assigned agents, echoing our extensive-margin findings.

The second possible mechanism—customers are more abusive toward women—is motivated by a growing literature that documents high rates of harassment for women in the workplace. However, the results in column 2 of Table 4 suggest this mechanism does not explain the differences in sales in our setting. The outcome measures any language classified as harassment within the chat. The data contain few instances of harassment interactions: 0.3% of conversations for the male-assigned (control) sample indicate harassment. The rate in the female-assigned sample is practically identical to the male-assigned sample and does not differ statistically.

Finally, column 3 of Table 4 tests whether customers bargain more often with female sales agents. While 15% of chats exhibit some bargaining behavior—for example, asking for deals—we find no significant effect of female-name assignment on the likelihood of bargaining. This null result is fairly precise relative to the baseline level and rules out changes to bargaining that exceed 20%. Therefore, in this context, differential bargaining does not appear to drive the observed productivity differences.

Together, our results suggest that customers interact differently with women and men in ways that can meaningfully reduce productivity. This result is especially consequential for the service industry, where customer-facing roles abound. As the service sector continues to expand, this issue may explain the persistence of the gender wage gap. Our investigation of the mechanisms behind this behavior suggests that consumers engage less with female agents—along extensive (any engagement) and intensive (tone used) margins. We find no evidence that customers differentially harass or bargain with women in our setting.[20]

[20] We are unable to test for other mechanisms—e.g., homophily (a customer's preference

4.3 Comparison to non-experimental estimates

We compare the results from our experimental research design to a simpler non-experimental comparison of male and female agents. The non-experimental results measure correlations between chat purchases and agent gender using chats *outside* the experimental sample.[21] We include similar controls in both specifications but cannot include agent fixed effects in the non-experimental comparison as they are collinear with gender.

Table 5 shows correlations between agents' actual gender and sales. We find no statistically significant differences between male and female agents across any of the three purchase outcomes. To compare the effect of female-name assignment in the experimental sample to the effect of being female in the correlational sample, we use seemingly unrelated regression. In this comparison, the outcome variable is the any-sale indicator measured using administrative sales records (within 48 hours).[22] A test of equality across the two 'female' coefficients rejects the null hypothesis at the 5% level ($p = 0.048$).

The difference between the experiment and correlational estimates sheds light on gender-based selection into this occupation. In particular, it suggests women in these jobs may be more productive than their male counterparts in the absence of customer discrimination. This is consistent with an equilibrium outcome in which males and females are paid similar wages, with female employees being taxed by customer bias. These results also illustrate the value of experimental analyzes in trying to identify how customer preferences affect worker productivity—an effect hidden in cross-sectional comparisons of workers.

to interact with an agent of the same gender)—because we lack sufficient customer information.

[21] The experimental sample is not representative of all sales at the company. Therefore this exercise is only suggestive.

[22] Only administrative sales records are available for the non-experimental sample.

5 Conclusion

This paper demonstrates that customer-based discrimination can cause negative effects on female workers' productivity. When sales agents randomly receive female-sounding names, the probability a customer makes a purchase falls by 3.9 percentage points. Consumers also purchase fewer total products, and the total value of their purchases falls. An exploration of potential mechanisms suggests these results are most consistent with customer disinterest in working with female agents—rather than differential bargaining or openly hostile behavior.

Our results have several implications. First, they suggest that female sales agents in our context may be more productive than their male counterparts when holding customer behavior constant. This result speaks to the "twice as hard" phenomenon whereby members of a discriminated group need to perform better than their counterparts in order to maintain their position in the workplace (Sofoluke and Sofoluke, 2021).

Second, these results indicate that *equal-pay-for-equal-work* policies may not fully resolve discrimination's effects when workers face discrimination from customers. This outcome is particularly relevant to the service industry, which often ties employee pay to productivity/output (e.g., number of sales) through piece-rate wages. In our setting, discrimination's effects might be eliminated by agents using gender-neutral names (or avoiding names altogether). More broadly, however, our results demonstrate that employers/institutions can play essential roles in mitigating discrimination—creating environments that do not compel individuals to obscure their identities. Solutions could involve further restricted unequal wage practices—mitigating the consequences of consumer-based discrimination. Alternatively, companies could transition away from individual-based incentivized pay schemes[23] or endeavor to sensitize customers to female workers. In our setting, this normalization could occur through the exclusive use of female-

[23] For example, employers could 'pool' performance-based bonuses—a common practice for sharing tips in the restaurant industry.

sounding names—acclimating customers to the presence of women in these roles. This solution is similar to Beaman et al. (2009), who show that prior exposure to female politicians improves individuals' positive perceptions of female leaders.

Finally, the results speak to specific barriers women face in the labor market in Sub-Saharan Africa. Like many regions around the world, significant gender disparities exist in formal-sector employment across Sub-Saharan Africa, where less than 15 percent of women work full-time for an employer (World Bank, 2013; Klugman and Twigg, 2016). Numerous economic models and empirical studies suggest that improvements in gender parity can drive substantial economic growth (World Economic, 2017). Recognizing this potential, governments throughout the continent are leading initiatives to address equity issues—including powerful provisions that support gender equality (e.g., USAID). This paper documents an understudied barrier to gender equity that institutions should address when designing policy.

References

ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly harmless econometrics*, Princeton university press.

ASHRAF, N. (2009): "Spousal control and intra-household decision making: An experimental study in the Philippines," *American Economic Review*, 99, 1245–77.

BAERT, S. (2018): "Hiring discrimination: An overview of (almost) all correspondence experiments since 2005," *Audit studies: Behind the scenes with theory, method, and nuance*, 63–77.

BANK, W. (2011): *World development report 2012: Gender equality and development*, The World Bank.

BAR, R. AND A. ZUSSMAN (2017): "Customer discrimination: evidence from Israel," *Journal of Labor Economics*, 35, 1031–1059.

BARTLETT, K. T. AND M. GULATI (2016): "Discrimination by Customers," https://ilr.law.uiowa.edu/print/volume-102-issue-1/discrimination-by-customers/.

BEAMAN, L., R. CHATTOPADHYAY, E. DUFLO, R. PANDE, AND P. TOPALOVA (2009): "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, 124, 1497–1540.

BECKER, G. S. ET AL. (1971): "The Economics of Discrimination," *University of Chicago Press Economics Books*.

BERTRAND, M. AND E. DUFLO (2017): "Field experiments on discrimination," *Handbook of economic field experiments*, 1, 309–393.

BERTRAND, M. AND S. MULLAINATHAN (2004): "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American economic review*, 94, 991–1013.

BLAU, F. D. AND L. M. KAHN (2017a): "The gender wage gap: Extent, trends, and explanations," *Journal of economic literature*, 55, 789–865.

——— (2017b): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 55, 789–865.

BORKER, G. (2021): "Safety First: Perceived Risk of Street Harassment and Educational Choices of Women," *Working Paper*.

BURSZTYN, L., A. L. GONZÁLEZ, AND D. YANAGIZAWA-DROTT (2020): "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, 110, 2997–3029.

CALIENDO, M., W.-S. LEE, AND R. MAHLSTEDT (2017): "The Gender Wage Gap and the Role of Reservation Wages: New Evidence for Unemployed Workers," *Journal of Economic Behavior & Organization*, 136, 161–173.

CARD, D., A. R. CARDOSO, AND P. KLINE (2016): "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly journal of economics*, 131, 633–686.

CASTILLO, M., R. PETRIE, M. TORERO, AND L. VESTERLUND (2013): "Gender Differences in Bargaining Outcomes: A Field Experiment on Discrimination," *Journal of Public Economics*, 99, 35–48.

COMBES, P.-P., B. DECREUSE, M. LAOUENAN, AND A. TRANNOY (2016): "Customer discrimination and employment outcomes: theory and evidence from the french labor market," *Journal of Labor Economics*, 34, 107–160.

DEAN, J. T. AND S. JAYACHANDRAN (2019): "Changing Family Attitudes to Promote Female Employment," *AEA Papers and Proceedings*, 109, 138–142.

DELECOURT, S. AND O. NG (2021): "Does gender matter for small business performance? Experimental evidence from India," .

DOLEAC, J. L. AND L. C. STEIN (2013): "The visible hand: Race and online market outcomes," *The Economic Journal*, 123, F469–F492.

DUFLO, E. (2012): "Women empowerment and economic development," *Journal of Economic literature*, 50, 1051–79.

DUPAS, P., A. SASSER MODESTINO, M. NIEDERLE, J. WOLFERS, AND T. S. D. COLLECTIVE (2021): "Gender and the Dynamics of Economics Seminars," Working Paper 28494, National Bureau of Economic Research.

FIELD, E., R. PANDE, N. RIGOL, S. SCHANER, AND C. TROYER MOORE (2021): "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms," *American Economic Review*, 111, 2342–2375.

FOLKE, O. AND J. K. RICKNE (2020): "Sexual harassment and gender inequality in the labor market," .

GALLEN, Y., R. V. LESNER, AND R. VEJLIN (2017): "The Labor Market Gender Gap in Denmark: Sorting out the Past 30 Years," 41.

GEORGIEVA, K. (2018): "Changing the Laws That Keep Women out of Work," https://www.worldbank.org/en/news/opinion/2018/03/29/changing-the-laws-that-keep-women-out-of-work.

HARDY, M. AND G. KAGY (2020): "It?s getting crowded in here: experimental evidence of demand constraints in the gender profit gap," *The Economic Journal*, 130, 2272–2290.

HEATH, R. AND X. TAN (2020): "Intrahousehold Bargaining, Female Autonomy, and Labor Supply: Theory and Evidence from India," *Journal of the European Economic Association*, 18, 1928–1968.

HOLZER, H. J. AND K. R. IHLANFELDT (1998): "Customer discrimination and employment outcomes for minority workers," *The Quarterly Journal of Economics*, 113, 835–867.

IFC (2021): "Women and E-Commerce in Africa," *International Finance Corporation*, 74.

ILO (2019): "Wages in Africa," .

JAYACHANDRAN, S. (2015): "The roots of gender inequality in developing countries," *economics*, 7, 63–88.

KAHN, L. M. AND P. D. SHERER (1988): "Racial differences in professional basketball players' compensation," *Journal of Labor Economics*, 6, 40–61.

KLINE, P. M., E. K. ROSE, AND C. R. WALTERS (2021): "Systemic discrimination among large US employers," Tech. rep., National Bureau of Economic Research.

KLUGMAN, J. AND S. TWIGG (2016): "Gender at Work in Africa: Legal Constraints and Opportunities for Reform," *African Journal of International and Comparative Law*, 24, 518–540.

LEONARD, J. S., D. I. LEVINE, AND L. GIULIANO (2010): "Customer discrimination," *The Review of Economics and Statistics*, 92, 670–678.

LIVEAGENT (2022): "Agent alias," https://www.liveagent.com/customer-support-glossary/agent-alias/.

LOWE, M. AND M. MCKELWAY (2021): "Coupling Labor Supply Decisions: An Experiment in India," *Working Paper*.

MCKELWAY, M. (2021a): "How Does Women's Employment Affect Household Decision-Making? Experimental Evidence from India," *Working Paper*.

——— (2021b): "Women's Employment in India: Intra-Household and Intra-Personal Constraints," *Working Paper*.

NEUMARK, D., R. J. BANK, AND K. D. VAN NORT (1996): "Sex discrimination in restaurant hiring: An audit study," *The Quarterly journal of economics*, 111, 915–941.

O'Donnell, M., U. Nwankwo, A. Calderon, C. Strickland, et al. (2020): "Closing Gender Pay Gaps," .

Rousille, N. (2021): "The Central Role of the Ask Gap in Gender Pay Inequality," *Working paper*.

Sin, I., S. Stillman, and R. Fabling (2017): "What Drives the Gender Wage Gap? Examining the Roles of Sorting, Productivity Differences, and Discrimination." .

——— (2020): "What Drives the Gender Wage Gap? Examining the Roles of Sorting, Productivity Differences, Bargaining and Discrimination," *The Review of Economics and Statistics*, 1–44.

Sofoluke, R. and O. Sofoluke (2021): *Twice As Hard: Navigating Black Stereotypes and Creating Space For Success*, New York, NY: DK.

Statista (2019): "Online Shoppers in Africa by Gender," https://www.statista.com/statistics/1190608/online-shoppers-in-africa-by-gender/.

Subramanian, N. (2021): "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment," *Working Paper*.

UN (2022): "Facts and Figures: Economic Empowerment," https://www.unwomen.org/en/what-we-do/economic-empowerment/facts-and-figures.

UNCTAD (2018): "UNCTAD B2C E-COMMERCE INDEX 2018: FOCUS ON AFRICA," Tech. rep.

USAID (2017): "Gender Equality and Women's Empowerment in Kenya | Kenya | Archive - U.S. Agency for International Development," https://2012-2017.usaid.gov/kenya/gender-equality-and-womens-empowerment-kenya.

VESTERLUND, L. (2018): "Knowing When to Ask: The Cost of Leaning-in," Working Paper 6382, Department of Economics, University of Pittsburgh.

WORLD BANK, G. (2013): "Gender at Work: A Companion to the World Development Report on Jobs," https://www.worldbank.org/content/dam/Worldbank/document/Gender/GenderAtWork

WORLD ECONOMIC, F. (2017): "The Global Gender Gap Report 2016 - World," https://reliefweb.int/report/world/global-gender-gap-report-2016.

6 Tables

Table 1: Placebo tests for female assignment

	N	Var. Mean	Female
Customer mention agent true name	2657	$<.01$	-.00128
			(.00234)
Customer amount of past chats	2657	.11	-.0463*
			(.0252)
Customer amount of past purchases	2657	.25	-.0175
			(.0496)
Agent first message length	2657	5.47	-.000659
			(.0041)
Agent chats (daily)	337	7.76	-.106
			(.596)
Agent hours worked (daily)	337	2.57	-.0272
			(.173)
Joint p-value			.65

This table shows customer and agent outcome means in column (2) and correlation between female name assignment and outcomes in column (3). The number of chats and hours worked by agents are at the day level, while the other variables are at the chat level. Controls include agent-month fixed effects. Female indicator determined in customer's first chat of the day. Standard errors in parentheses and clustered at agent-day level. Joint p-value tests equality of all coefficients with zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Effect of female assignment on purchase outcomes

	Purchases (48h)			Purchases (24h)		
	(1) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	-.039***	-.039***	-3.5***	-.037***	-.036***	-3.4***
	(.012)	(.013)	(1.2)	(.011)	(.012)	(1.2)
Control Mean	.062	.065	4.177	.057	.059	3.820
N	2653	2653	2653	2653	2653	2653

This table shows the effect of female name assignment on purchase outcomes. Any represents any purchase, Total represents number of purchases, and Total price is the cumulative price of all purchases in EUR. Any purchases and total purchases combine purchases by customer and by agent, while total price is based only on customer purchases. Purchases are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Effect of female assignment on conversation response

	Initial messages		
	(1)	(2)	(3)
	Ever respond (C)	Msgs to response (A)	Msgs to response (C)
Female	-.027	.0094**	-.11*
	(.02)	(.0041)	(.059)
Control Mean	.660	1.010	1.254
N	2653	2653	2653

This table shows the effect of female name assignment on customer and agent responses. Ever respond (C) is a 1 if the customer ever responded. Msgs to response (A) is the number of messages sent by agent before customer first response. Msgs to response (C) is the number of messages by customer in intial response. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect of female assignment on chat outcomes

	Tone	Negativity		Bargaining
	(1)	(2)	(3)	(4)
	Any	Harass	Any neg.	Any
Female	-.027*	.00018	.0064	.0081
	(.014)	(.0028)	(.0041)	(.018)
Control Mean	.086	.003	.009	.147
N	1742	1742	1742	1742

This table shows the effect of female name assignment on chat outcomes. Column (1) measures any non-neutral chat tone, column (2) measures any harassment of the agent, column (3) measures any negative words or phrases, and column (4) measures any bargaining. The sample include only chats with any consumer response. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, customer chat history, and hand coder fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Correlational relationship between female agent and sales outcomes

	Sales (48h)			Sales (24h)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Any	Total	Total price	Any	Total	Total price
Female	.0034	.0049	-1.6	.0018	.0013	-2.1
	(.0076)	(.0084)	(1.5)	(.0074)	(.0077)	(1.5)
Control Mean	.02	.03	4.70	.02	.02	4.48
N	8863	8863	8863	8863	8863	8863

This table shows the correlational effect of female agent on sales outcomes. Any represents any sale, Total represents number of sales, and Total price is the cumulative price of all sales in EUR. All outcomes based on customer purchases only. Sales are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix

A1 Tables

Table A1: Effect of female assignment on purchase outcomes (customer-day level)

	Purchases (48h)			Purchases (24h)		
	(1) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	-.04***	-.042***	-3.2***	-.039***	-.038***	-3.1***
	(.012)	(.014)	(1.2)	(.011)	(.013)	(1.2)
Control Mean	.064	.068	4.351	.059	.061	3.936
N	2169	2169	2169	2169	2169	2169

This table shows the effect of female name assignment on purchase outcomes. The data is aggregated to the customer-day level. Any represents any purchase, Total represents number of purchases, and Total price is the cumulative price of all purchases in EUR. Any purchases and total purchases combine purchases by customer and by agent, while total price is based only on customer purchases. Purchases are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors clustered at the agent-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Effect of female assignment on purchase outcomes (without controls)

	Purchases (48h)			Purchases (24h)		
	(1) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	-.038***	-.038***	-3.2***	-.036***	-.035***	-3.2***
	(.013)	(.014)	(1.2)	(.011)	(.012)	(1.1)
Control Mean	.062	.065	4.177	.057	.059	3.820
N	2655	2655	2655	2655	2655	2655

This table shows the effect of female name assignment on purchase outcomes. Any represents any purchase, Total represents number of purchases, and Total price is the cumulative price of all purchases in EUR. Any purchases and total purchases combine purchases by customer and by agent, while total price is based only on customer purchases. Purchases are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include agent-month fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Effect of female assignment on purchase outcomes (agent level)

	Purchases (48h)			Purchases (24h)		
	(1) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	-.04***	-.042***	-3.2**	-.038***	-.038***	-3.1***
	(.013)	(.014)	(1.3)	(.012)	(.013)	(1.2)
Control Mean	.065	.070	4.513	.060	.062	4.091
N	335	335	335	335	335	335

This table shows the effect of female name assignment on purchase outcomes. Data is at the agent-day level. Any represents any purchase, Total represents number of purchases, and Total price is the cumulative price of all purchases in EUR. Any purchases and total purchases combine purchases by customer and by agent, while total price is based only on customer purchases. Purchases are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include agent-month fixed effects while a first-step regression included these and additionally customer location, customer purchase history, and customer chat history fixed effects. The outcome is the coefficient on agent-day from the first step-regression. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Effect of female assignment on administrative purchase outcomes

	Purchases (48h)			Purchases (24h)		
	(1) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	-.02**	-.022**	-3.5***	-.018**	-.018**	-3.4***
	(.0097)	(.011)	(1.2)	(.0087)	(.0093)	(1.2)
Control Mean	.033	.037	4.177	.029	.030	3.820
N	2653	2653	2653	2653	2653	2653

This table shows the effect of female name assignment on purchase outcomes taken from administrative records. Any represents any purchase, Total represents number of purchases, and Total price is the cumulative price of all purchases in EUR. Any purchases and total purchases combine purchases by customer and by agent, while total price is based only on customer purchases. Purchases are measured within 24 or 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Effect of female assignment on any sales by agent (48 hours)

	Purchases (48h)
	(1) Any
Female * Agent 1	-.014
	(.014)
Female * Agent 2	-.023
	(.039)
Female * Agent 3	-.017
	(.033)
Female * Agent 4	-.04***
	(.011)
Female * Agent 5	.0058
	(.0099)
Female * Agent 6	-.05**
	(.019)
Control Mean	.062
Joint p-value	.02
N	2653

This table shows the effect of female name assignment on purchase outcomes by agent. Any represents any purchase. Purchases are measured within 48 hours of the start of the chat. Female indicator determined in customer's first chat of the day. Joint p-value tests equality of all coefficients. Controls include agent-month, customer location, customer purchase history, and customer chat history fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Effect of female assignment on chat tones

	Tone					
	(1) Any	(2) Angry	(3) Happy	(4) Ecstatic	(5) Impatient	(6) Sad
Female	-.027*	.002	-.02***	-.0017	-.017**	.01**
	(.014)	(.0053)	(.0076)	(.0011)	(.0074)	(.0051)
Control Mean	.09	.01	.04	.00	.03	.01
N	1742	1742	1742	1742	1742	1742

This table shows the effect of female name assignment on chat tone outcomes. Outcomes measure either any tone, or any of the specific types of chat tones. Female indicator determined in customer's first chat of the day. Controls include agent-month, customer location, customer purchase history, customer chat history, and hand coder fixed effects. Standard errors two-way clustered at the agent-day and customer-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Outcome Variable Descriptions

Purchases	
Any Purchases	Whether customer made any purchase 24 or 48 hours after the chat
Total Purchases	The total number of purchases that were by the customer made 24 or 48 hours after the chat
Total Price	The cumulative price of all purchases in EUR that were made by the customer 24 or 48 hours after the chat
Chats	
Ever Respond	= 1 if the customer ever responded
Messages to Response (A)	Number of messages sent by agent before customer first response.
Messages to Response (C)	Number of messages sent by customers in their initial response. Customers that never respond are coded as 0.
Tone	
	We employed two research assistants (RA) based in Sub-Saharan Africa to read through all of the chats. We used a double-blind process so that 20% of all chats were reviewed by both assistants. Any discrepancies in how questions were being coded were flagged early in the process to streamline coding styles. We chose to hand code these outcomes as opposed to using natural language processing (which we attempted) for three reasons. First, the vast majority of interactions include many "polite" words such as "Thanks" or "Please", which meant many conversations were coded as friendly by the machine learning algorithm even if they were acrimonious. Second, the chats contain a large number of misspellings and chat shorthand, which are not included in language databases. Finally, we thought it best to have individuals who are familiar with the cultural context interpreting the tone of the conversation.
Any	Measures any non-neutral chat tone. Chats were coded neutral or non-neutral tone (including angry, sad, happy, ecstatic, impatient)
Harassment	Measures any harassment of the agent

Any negative	Measures whether any negative words or phrases were used by the customer.
Bargaining	Measures any bargaining with the agent. This includes asking for discounts, or better prices.