Customer Discrimination in the Workplace: Evidence from Online Sales in Sub-Saharan Africa*

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

March 7, 2022

The current draft is preliminary and incomplete.
Please do not distribute.

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 in 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 do not find evidence of differential bargaining or harassment.

*Funding for this project was graciously provided by PEDL. Borui Sun and Victoria Yin 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.

†Development Impact Monitoring Evaluation Unit, World Bank

‡Department of Economics, American University

§Department of Economics, Brown University

¶Department of Economics, University of Oregon

1 Introduction

Women work less outside the home, earn less, and run less profitable businesses than men on average, especially in low-income countries. Recent literature has focused on understanding why this might be the case. One potential contributor that has received relatively less attention is customer-side discrimination, a phenomenon whereby customers prefer to work or interact with certain types of workers. If consumers are selectively biased against women by purchasing products less often or bargaining more forcefully with them, then female wages, promotions, and hiring prospects could be affected. This source of discrimination is potentially large and persistent (Becker et al., 1971), since discrimination emanating from consumers may not disappear with market competition. Disentangling the precise role of customer discrimination is challenging because of all the different factors that may account for a worker's productivity including the workers own behavior, customer behavior and the workplace environment.

We overcome these challenges through an experiment at an online sales company in Sub-Saharan Africa. We study workers, specifically sales agents, who chat with customers online in order to answer questions and increase sales. Two aspects of the context provide a novel framework for estimating the causal effect of customer discrimination. First, the names of workers, and hence their inferred gender, were randomized on a daily basis, providing plausible variation in customer beliefs about worker gender.[1] Information on names was salient as customers could only infer agents' gender by the name that appears on the chat. Second, while agents were aware that their names were changed, they were unaware of the particular name assignment and hence their behavior was not directly related to the name variation. Therefore, any change in consumer behavior towards sales agents can only be a product of consumers rather than the agents themselves.

This research design and setting marries the advantages of two common methods to study employer-based discrimination: audit and correspondence studies. Audit studies, which recruit people with similar general characteristics except one (e.g. gender) to apply or interview for a job in person, are able to collect detailed outcomes through the in-person interactions that take place. However, the chosen actors in these studies may actually differ in many ways, which makes it difficult to isolate the impact of that specific characteristic. Correspondence studies, in which fictitious applications are sent to employers, can create clear causal comparisons but are unable to measure the same detailed set of interactions. The name randomization in the context of online workplace interactions overcomes these challenges. Agents' names were randomized daily, generating clear counterfactuals: the comparison we make is of consumers chatting to an agent with a male-sounding name on one day relative to a consumer chatting with that same agent with a female-sounding name on another day. The setup also provides an opportunity to collect detailed outcomes, including overall chat outcomes, such as likelihood of a purchase, and specifics of the interaction,

[1] Changing the names of sales agents appears to be relatively common in online sales settings (e.g. LiveAgent (2022)). This happened prior to the study period as well, with some of the agents employing different names of their choosing.

such as 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). Second, in Sub-Saharan Africa the high prevalence of conservative social norms that favor men over women as economic agents and business owners may contribute to heightened gender discrimination by consumers. 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 than women and that women have no say in decisions on large household purchases (Jayachandran, 2015). Such norms may contribute to higher rates of customer discrimination. Finally, the service sector is growing across the continent, which makes issues related to gender-based differences in customer interactions increasingly relevant.

We find that being assigned to a female name reduces the likelihood that the customer makes any purchase, the number of purchases they make, and the value of those purchases.[2] Specifically, the likelihood of any purchase decreases by 2.1 percentage points in our preferred specification, a relatively large effect given the low purchasing rates at baseline (3.3%). Total number of purchases falls by 0.026 and value falls by 3.6 euros. This suggests that productivity differences between workers in this context are not just a reflection of worker attributes, but also of differential customer responses.

To confirm these treatment effects stem from changing the agents' assigned names, we investigate whether customers are aware of agents' names in the first place. We find that in 10% of chats customers mention the agent's assigned name, which is an indication that names are a salient feature of the interaction. We can also rule out a potential confound whereby customers are responding to a mismatch between assigned name and actual gender of the agent. In sum, our main results suggest that being assigned a female name has a negative effect on productivity.

The observed reductions in worker productivity could be explained by several possible mechanisms. These include general customer disinterest in working with female agents because of taste-based or statistical discrimination, differential bargaining, or openly negative interactions. Our results show that customer disinterest is the most salient. We find that some consumers only respond to female agents once they have received additional messages from the agent.[3] This suggests that some consumers are hesitant to engage with female agents unless these agents persist through additional messaging. We also find that consumers are less likely to express any tone (e.g. positive or negative words), which we interpret as another measure of engagement with the agent.

The other mechanisms we explore are not consistent with our results. We do not find any difference in the probability that consumers engage in differential bargaining or harass-

[2]Purchases are measured as occurring within one or two days after the time of the chat, and the results are robust to either definition.

[3]Agents always send the first message and the assigned name of the agent is revealed at that time.

ment.[4] This is potentially relevant as differences in bargaining by gender has featured 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 openly negative or harassing behavior, although any form of harassment is extremely rare in this context.[5]

The results from the experiment differ from those we would have found by analyzing the correlation between an agent's gender and their productivity. In a simple comparison between workers at the company, the probability of making a sale is the same for male and female agents. By contrast, the probability of making a sale is much lower for agents who are assigned female names in the well-controlled experimental sample. These differences suggest that while the company hired males and females with similar levels of productivity, these productivity metrics do not account for the fact that women face greater levels of consumer discrimination that dampen their ability to make sales. Removing these barriers could potentially result in higher productivity among female agents. The vast majority of agents in this context are female indicating that overall firm productivity could increase as well if consumer discrimination were eliminated.

This paper contributes in three main ways. First, this paper makes contributions to the existing literature on the effect of discrimination in the labor market. Most of these studies focus predominantly 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 the effect of customer discrimination has tended to test for racial and 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). More recently, Dupas et al. (2021a) find that economics seminar audiences ('the consumers') engage women and men speakers differently—asking them more questions that are patronizing and hostile.[6] While these studies benefit from studying diverse labor market settings, they have been in high income contexts and they are not able to fully control for differences in unobservable characteristics between consumers, thereby introducing the potential for bias. As in Doleac and Stein (2013), this paper circumvents that issue by using an online sales context where seller attribute information can be randomly assigned to consumers and remain unrelated to other characteristics. While that paper documents race-based customer discrimination among independent sellers in a high-income marketplace setting, we test for gender-based customer discrimination among workers in a low-income country context.

Secondly, this paper contributes to recent studies focused on identifying the barriers women face in the labor market in low-income countries. Recent work identifies a variety

[4]Differences in bargaining for females could occur due to differences in how females bargain (e.g. risk aversion) or how individuals bargain with them. We would expect the latter in this context because we are studying customer behavior.

[5]We cannot assess possible mechanisms such as homophily in gender-gender interactions because we do not have information about consumers' identity.

[6]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; while (Neumark et al., 1996) illustrate how restaurants discriminatory hiring practices could be driven by their patrons' preferences.

of constraints including 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 in work commutes (Borker, 2021), and employer discrimination (Jayachandran, 2015; Duflo, 2012; Sin et al., 2017). Our paper adds to this literature by focusing on the role of consumer preferences as a potentially important but understudied barrier to women's success in the labor market. It looks specifically at three ways these preferences could be manifest (bargaining, harassment and disinterest) and their potential impact on female worker productivity. Two related papers, Hardy and Kagy (2020); Delecourt and Ng (2021), investigate how customer preferences can lead to differential outcomes for male vs female-lead small businesses and find mixed evidence. Our results may be especially relevant in low-income countries where customers may be more hesitant to engage with a female labor force because of existing social norms.

Finally, we see our work as contributing to an important policy discussion about gender equality in the workplace in low-income countries (World Bank, 2013). Export-oriented customer-service and IT jobs are on the rise and, together with the emergence of new technologies, is likely to lead to more client-facing employment in the developing world. Increasing the number of client-facing jobs in the presence of consumer-side discrimination, without intervention, may very well exacerbate already-existing gender-based disparities. While forward-thinking companies have started to embrace new practices that equalize opportunities for men and women (better wages, more flexible hours), customer discrimination still has the potential to significantly affect women's productivity in the workplace. Additional methods to alleviate the role of consumer preferences may impact firm growth and break a cycle of discrimination whereby women are assumed to be less productive than men.

The paper is organized as follows. Section 2 provides background on the experiment and context. Section 3 explains the data, Section 4 describes the experimental strategy, and Section 5 discusses the results. Section 6 concludes.

2 Context

2.1 Service sector in developing countries

This paper studies consumers' discriminatory behavior when they engage with online sales agents in Sub-Saharan Africa. These types of customer-facing roles are growing across the continent as the number of service sector jobs increases. For example, the share of the working age population employed in services in Sub-Saharan Africa has risen by 12% from 2010-2019 (WDI). These trends are driven by women whose share of the working age population employed in services has increased by 16% over the same period, and currently stands at 39.7%.

This trend is likely to persist as internet connectivity continues to improve across Africa, and service sector jobs increasingly interface with clients online. In 2010 only 8.3% of the population in Africa had access to the internet in the last three months. By 2017, this number had increased to 22.3% (WDI). Online shopping, in particular, has increased by 18% annually between 2014 and 2017 (UNCTAD, 2018). The COVID-19 pandemic has likely accelerated these trends, as consumers have turned towards online purchasing.

2.2 Study details

We evaluate an experiment at an online sales company based in Sub-Saharan Africa. The company employed sales agents to answer customer questions and increase purchasing. Customers could initiate an interaction with a sales agent by clicking on a small pop-up window at the bottom of their webpage. The pop-up window would display the agents name, and feature a short greeting to the customer.

The company was keen to partner with the research team to investigate whether they should streamline the way their sales interface worked for customers. The goal of this particular test was to identify how customer behavior changed when agents were assigned to a male versus female sounding sales name. To this end, the company needed to be able to 1) randomize whether the name that appeared on the chat was a male or female sounding name; and 2) ensure that agents were unaware of the names assigned to them that particular day. To ensure the randomization was implemented correctly, a software program pulled one name per sales agent per day from an existing list, and assigned that name to the agent within the sales interface. This list of names was sourced by a local field team, who then assigned a gender to each name on the list.

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 sales interface. The company installed the plugin on each agent's computer with oversight from our field team. The plugin symbol appeared as a light grey square alongside the other extensions on the browser, albeit removed from the list of visible extensions. The plugin worked in the following way. If the agent's real name was James and they were assigned the name Pierre at the beginning of the day, James would only see "Agent" on his chat window, while the client would see "Pierre." This included any general references to the agents assigned name and any mentions of the assigned name in the chat transcript. The plugin was written in JavaScript and the underlying code instructed the plugin to monitor the information on the webpage in order to mask assigned names.

The vast majority of interactions occurred in English, which removes any concern about gendered identifiers. Another potential concern stems from the potential for chats to be transferred to the phone. This is possible but happens rarely; only 2% of interactions are transferred.

3 Data

The analysis relies on two sources of administrative data. The first dataset records every sale made and the amount of the sale. The second dataset contains the agent-chat interactions, which includes the full chat transcript, a timestamp for each message sent by the customers and the agent, and the country of the customer. The sales data were matched to the chat logs using the customers' IP address.[7]

From the chat transcripts, we can create objective and subjective outcome measures. Objective measures do not require human interpretation. These include outcomes such as whether a purchase occurred or the length of the conversation. Subjective measures refer to outcomes that require human interpretation of the chat such as the overall tone, whether the customer was bargaining with the agent, or whether any harassment occurred. Subjective outcomes were hand coded by enumerators, where 20% of the observations were double coded to ensure consistency throughout the sample.

Six female agents worked throughout the study. Agents are involved in various sales-related activities, including assisting customers over the phone and via chat. They are only involved in chats 5 week-days per month. On days when agents respond to chats they typically spend 2.5 hours on the online sales interface with customers, and engage in approximately 6 unique chats per day. Each chat lasts 11 minutes on average and contains 40 words.

4 Empirical strategy

Estimating a causal effect of gender-based consumer discrimination is challenging. A simple comparison of worker productivity by gender is unlikely to measure a causal effect due to the many factors correlated with gender that may also affect worker productivity. A secondary challenge is that if the worker is aware that he or she has been assigned a name corresponding to the opposite gender, then the worker may behave differently. This would impede the estimation of a causal effect of consumer discrimination as it may be confounded by changes in worker behavior.

This study uses the randomization of the presentation of gender to consumers and a name-masking procedure to overcome both challenges. First, sales agents' names are randomized daily. In doing so, we ensure that customers who interact with agents are randomly matched to either a female or male sounding name over time. The randomization separates factors correlated with gender from the gender perceived by consumers. Second, agents are not aware of the name consumers see—any revelation of the agent's name during the chat is masked automatically by a computer program and is not seen by the agent. Specifically, any mention of the agent's name within the chat is changed to display "agent", as discussed in Section 2.2. These measures limit the potential for worker behavior to change as a result of the random assignment of gender.

The randomization occurs as follows. Agents are randomly assigned to be either male or female each day with replacement. Their particular name is then randomly assigned from a dictionary of female- and male-sounding names. This occurs for all days of the study period. We restrict our sample to weekdays, when agents have regular schedules. The number of

[7] We restrict the data to include observations with fewer than five previous purchases given that some users may access the site using non-unique locations such as public areas or at businesses. This retains 98% of the observations.

agents working daily varies. On some days only one agent is operating the chat, and on other days there are multiple agents interfacing with customers via chat.

To estimate the effect of customer discrimination on worker productivity, we exploit this randomization procedure. 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 consumer i, working with agent a, in month m. The indicator $[\text{Assigned female}]_{ia}$ is a 1 if individual i is matched with agent a, who is assigned female in that period. The fixed effect γ_{am} are agent-month fixed effects, restricting comparisons within this grouping. Therefore, the simplified research design is a comparison between a consumer who chats with agent a in some month m on a day when the agent is assigned female compared with a consumer chatting with the same agent in the same month when the agent is assigned male. We further control for consumer characteristics, X_i, including country, past purchase history, and past chat history.

Customers may have multiple interactions with agents: customers may be disconnected or return to ask more questions. Therefore, the randomization is plausibly a series of lottery events, as an individual is assigned back to the same agent or another agent. We include all of these observations in the sample, analogous to a number of recent papers (Cellini et al., 2010; Gelber et al., 2015; Pecenco et al., 2019). To account for this, we two-way cluster our standard errors at the agent-day and customer-day level.[8]

One concern with this approach is that customers may try to switch which agent they interact with by starting a new conversation. In other words, they may not 'comply' with the treatment status by exiting one conversation and starting a new one, potentially with an agent of their preferred gender. We deal with this by assigning any additional chats the customer has in a particular day to the same name assignment as the first chat they had. For example, if a customer interacts with a female agent during the first chat but a male during their second chat of the day, we code both chats as having an agent with a female sounding name.[9]

There are two potential external validity concerns with this approach more generally. First, the name-masking procedure could also affect agent productivity. For example, a customer who thinks they are interacting with a female might engage in flirtatious behavior which could confuse a male sales agent who is not used to these types of interactions. This does not threaten our identification, but it may create a set of interactions that are no longer reflective of what you might see in reality. However, this situation is less likely in our context because the agent does not have any information about the customer and may be able to justify their behavior more easily as a result. Second, agents have certain gender-specific lan-

[8]Clustering standard errors at the agent-by-day level assumes no effect of the agent's treatment assignment on subsequent days. We test for dynamic treatments 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 ($p = .377$).

[9]The results are qualitatively and quantitatively similar if we only include the first instance of a consumer chat in a day.

guage 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 characteristics of the customer and agent prior to the chat on whether the agent is assigned to female and agent-month fixed effects. Female assignment is not correlated with any of the customer and agent characteristics at the 5% level, and we cannot jointly reject the null hypothesis that all of these effects are zero ($p = 0.65$). We additionally include a row in the table to identify whether the customer mentioned the agent's true name. This happens negligibly (mean is 0.00) and likely only occurs because of overlap between the agents' name and the topic of the conversation (e.g. Rose Hotel with an agent named Rose).

5 Results

5.1 Effect of name assignment

The experiment is designed to identify the impact of gender on consumer behavior. This strategy relies on consumers paying attention to the gendered names we assign agents. We can confirm that agent names are salient to customers by measuring how often consumers say the assigned name of the agent in the chat. This provides a lower bound for consumers awareness of agents' name and their associated gender. In our study sample, this occurs in 11% of all chats and 17% of chats where the consumers ever responds to the agent. We interpret this as a relatively high share of customer awareness given that many chats are relatively short. Therefore, agent names are indeed a salient feature of the chat interaction and could affect customers' future behavior.

Table 2 presents the effects of female name assignment on outcomes related to consumer purchases. We measure purchases within 24 or 48 hours of when the chat occurred in order to capture behavior that is plausibly related to the chats. We measure purchases in three ways: the number of distinct purchases; the probability of making any purchase; and, the total price of purchases. We find that consumers assigned to agents with female names are less likely to purchase products on the website. Specifically, column 1 shows that female agent assignment decreases the likelihood of any purchase within 48 hours of the chat by 2.1 percentage points ($p = 0.035$). The likelihood that a chat results in any purchase in the control group is only 3.3%. This point estimates translate into a 63% reduction in the likelihood of making a sale.[10] Column 2 shows that consumers also purchase 0.026 fewer total products ($p = 0.013$), and

[10] Our confidence intervals show that the data is consistent with more modest sales reductions as well.

that the total value of their purchases falls by 3.6 Euros (column 3). Columns (4-6) show the same outcomes measured within 24 hours of the chat—our results are not sensitive to the time window. Our preferred estimates pertain to outcomes measured within 48 hours in order to capture all related consumer behavior.

Our results highlight the importance of customer-side discrimination in explaining productivity differences between men and women in the workplace (for consumer facing roles). Research on the gender wage gap finds that women receive lower pay partly because they are less productive (Sin et al., 2020; Gallen et al., 2017; Blau and Kahn, 2017; Caliendo et al., 2017). We show that these differences in productivity may themselves be partially explained by discriminatory behavior on the part of consumers. In our context, this suggests that for men and women to have similar productivity levels, women will have to overcome more difficult workplace conditions created by consumers.

The repeated randomization design allows us to test whether there are heterogeneous treatment effects across agents. This is possible because each agent is randomized into both treatment statuses. These estimates may reflect differences in agent characteristics and/or the types of consumers that agents encounter. Specifically, we interact treatment with agent identity to calculate the agent-specific impact of being assigned a female sounding name on sales outcomes within 48 hours. While all agents have negative point estimates from female name assignment (except for one whose positive coefficient is not statistically different from zero), we can reject that the treatment effects are the same across all agents at the 10% level ($p = 0.070$). This suggests that the impact of being perceived as a male/female on productivity (sales) may differ across agent and consumer types rather than there existing an overall constant effect.

5.2 Mechanisms

There are a number of reasons why purchases may fall when consumers chat with female agents. Our data allows us to explore three potential mechanisms. First, customers may be hesitant to engage with female sales agents because of taste-based or statistical discrimination: customers may have an aversion to working with online female sales agents, or may believe women are less efficient at helping with purchases. Second, recent literature suggests that women are more likely to be subject to harassment and verbal abuse on the job, which would also harm their productivity. Finally, a large literature suggests that women may be subject to differential bargaining processes than men (e.g. Ashraf (2009); Rousille (2021); Vesterlund (2018); Castillo et al. (2013); Card et al. (2016))—a fact customers may internalize and attempt to leverage 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 of whether to engage with an agent at all, some consumers may be hesitant to chat with female agents or prefer not to altogether. On the intensive margin of chat interactions, consumers may use different tones when they are speaking to female agents.

Columns (1-3) of Table 3 show the effect of female name assignment on *extensive mar-*

gin consumer interactions. Agents always sends the first message, and the conversations continue from there. In column (1), female assignment leads to a negative but statistically insignificant effects on the likelihood the customer ever responds ($p = 0.166$). 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 prior to receiving a response by consumers ($p = 0.004$), which we interpret as an indication of lower consumer engagement. Finally, we test for engagement by looking at the number of messages the customer sends in their initial response to the agent. Column (3) shows that consumers send fewer messages when initiating a conversation with an agent with a female sounding name. Taken together, these three measures provide suggestive evidence that consumers may be hesitant to engage with an agent with a female sounding name.

To investigate customer hesitancy along the *intensive margin*, we measure the tone of the conversation. While specific tones are a relatively imperfect proxy for true emotions, whether a tone is detected at all may reflect a customer's level of engagement with the agent. To this end, we construct a measure for any non-neutral tone detected in the conversation. All chats were manually reviewed and tagged for whether the tone was neutral or not, and for whether any bargaining or harassment was detected. Column (1) of Table 4 demonstrates a 2.8 percentage point reduction in any measured tone when customers engage with female-assigned agents, a 33% reduction relative to the control group mean ($p = 0.048$). This suggests that consumers exhibit weaker levels of engagement overall with female-assigned agents. These results echo our findings on the extensive margin.

We turn next to investigating the second possible mechanism—customers are more abusive when they believe they are talking to female agents (Dupas et al., 2021b). This is motivated by a growing literature that documents high rates of harassment for women in the workplace (Georgieva, 2018). Columns (2-3) of Table 4 show this mechanism is unlikely to explain the differences in sales that we observe. Column (2) directly measures language classified as harassment while column (3) measures any negative language used within the chat. We find little evidence of these types of interactions in this context: 0.3% of conversations for the male-assigned sample contain measurable harassment while 0.9% have any negative language at all. We also do not find that female name assignment matters differentially for these outcomes.

Finally, we investigate whether customers are more likely to bargain with sales agents when speaking with a female agent. If consumers believe women may be less effective at bargaining, they may bargain differentially. Column (3) of Table 4 shows the effect of name assignment on bargaining behavior. While 15% of chats exhibit some bargaining behavior (by asking for deals for example), we find no differential effect based on female name assignment. At least in this context, consumers do not respond differently when bargaining with agents with male or female sounding names.

Taken together, our results suggest that customers interact differently with women and men in ways that can meaningfully reduce their productivity on the job. This is particularly important to document in the service industry, which is dominated by customer-facing roles. Moreover, as service sector jobs continue to increase, this problem could become even more important in explaining persistence in the gender wage gap. An exploration of mechanisms suggests that our consumers engage less with female agents—we document this along the extensive (any engagement) and intensive (tone used) margins. We find no evidence of differential rates of harassment and bargaining in our setting. We are also not able to test for other mechanisms such as homophily (a customer's preference to interact with an agent of the same gender) because we lack of information on customers.

5.3 Comparison to correlational effect

We compare the results from our experimental research design to a simpler non-experimental comparison of male and female agents to investigate the direct of selection effects. The non-experimental results measure correlations between chat purchases and agent gender using chats *outside* the experimental sample.[11] We include similar controls in both estimates but we cannot control for agent fixed effects in the non-experimental comparison. If the firm paid a constant wage, simple economic theory suggests that the likelihood of sales should be similar across agents of different genders. This would be an equilibrium outcome whereby agents are being paid their marginal productivity of labor, although we were never informed of agents' wages.

Table 5 shows the correlation between agent gender and sales. We analyze the same set of sales variables, including total sales, any sales, and total price within 24 and 48 hours from the time of the chat. We find no statistically significant differences between male and female agents across these outcomes. In contrast to our experimental results, the point estimates for the effect on sales are positive and we can rule out negative associations beyond 1.1 percentage points, less than half the observed experimental treatment effect.

We compare the effect of female-name assignment in the experimental sample to female agents in the correlational sample using seemingly unrelated regression with any sale within the 48 hours as our outcome. The coefficients reject the null hypothesis of equality at the 5% level ($p = 0.043$). This highlights the importance of running a randomized control trial to identify how customer preferences change worker productivity – a fact that is not observable when looking at the cross section.

6 Conclusion

This paper demonstrates that customer discrimination can lead to negative effects on female productivity. When sales agents are assigned a female sounding name, the likelihood of consumers making any purchase falls by 2.1 percentage points. Consumers also purchase fewer total products, and the total value of their purchases falls. An exploration of mech-

[11]The experimental sample was not representative of all sales at the company, therefore this exercise is only suggestive.

anisms suggests these results are most consistent with customer disinterest with working with female agents, rather than differential bargaining or openly negative behavior.

The results have a number of implications. First, these results suggest that female sales agents in our context may be more productive than their male counterparts if we hold fixed customer behavior. This 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, workplace policies that give equal pay for equal work may not fully correct for the discrimination women face in light of customer discrimination. This is especially relevant in the service industry where employee pay is often tied to output (the number of sales, for example) through piece-rate wages. These results also imply that employers/institutions may have to take additional measures so that women don't feel compelled to create an online presence that obscures their identity.

Finally, these results speak to barriers that women may face in the labor market in Sub-Saharan Africa. Like many regions around the world, significant gender disparities exist in the formal sector in Sub-Saharan Africa, where the share of women who work full-time for an employer is below 15 per cent (Klugman and Twigg, 2016). At the same time, there exists a variety of models and empirical studies suggesting that improvements in gender parity can result in significant economic growth (World Economic, 2017). Governments across the continent are recognizing this and have spearheaded a number of initiatives to address these issues, including powerful provisions supporting gender equality (eg.**?**). Our paper provides additional evidence on an understudied barrier that institutions should address when designing policy responses.

References

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.

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 (2017): "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.

CELLINI, S. R., F. FERREIRA, AND J. ROTHSTEIN (2010): "The value of school facility investments: Evidence from a dynamic regression discontinuity design," *The Quarterly Journal of Economics*, 125, 215–261.

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. S. MODESTINO, M. NIEDERLE, J. WOLFERS, ET AL. (2021a): "Gender and the dynamics of economics seminars," Tech. rep., National Bureau of Economic Research.

DUPAS, P., A. SASSER MODESTINO, M. NIEDERLE, J. WOLFERS, AND T. S. D. COLLECTIVE (2021b): "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.

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

GELBER, A., A. ISEN, AND J. B. KESSLER (2015): "The effects of youth employment: Evidence from New York city lotteries," *The Quarterly Journal of Economics*, 131, 423–460.

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.

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, W. (2022): "Live Agent," 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," .

PECENCO, M., C. SCHMIDT-PADILLA, AND H. TAVERAS (2019): "Opportunities and Entrepreneurship: Evidence on Advanced Labor Market Experience," .

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.

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

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

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_web.pdf.

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

7 Tables

Table 1: Placebo tests for female assignment

	N	Var. Mean	Female
Customer mention agent true name	2657	.00	-.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	-.021**	-.026**	-3.6***	-.019**	-.021**	-3.5***
	(.0098)	(.01)	(1.4)	(.0087)	(.0091)	(1.2)
Control Mean	.033	.037	4.476	.029	.030	4.004
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. 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) Ever respond (C)	(2) Msgs to response (A)	(3) Msgs to response (C)
Female	-.028	.0095**	-.11*
	(.02)	(.0041)	(.059)
Control Mean	.660	1.010	1.255
N	2655	2655	2655

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	-.028**	.00022	.0063	.0098
	(.014)	(.0028)	(.004)	(.018)
Control Mean	.086	.003	.009	.148
N	1744	1744	1744	1744

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) Any	(2) Total	(3) Total price	(4) Any	(5) Total	(6) Total price
Female	.0037	.0051	-3.7	.0023	.0018	-2.8
	(.0076)	(.0084)	(2.8)	(.0074)	(.0077)	(1.9)
Control Mean	.02	.03	7.01	.02	.02	5.47
N	8866	8866	8866	8866	8866	8866

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. 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.