Gender price gaps and competition:
Evidence from a correspondence study

Margarita Machelett*

Brown University

Link to current version

November 11, 2018

Job Market Paper

Abstract
I implement a large-scale field experiment in the US auto repair industry to study the existence and structure of gender-based price discrimination in service markets. Women receive price quotes that are 1.9 percent (9 dollars) higher than men. These differences disappear once women signal low search costs, suggesting statistical rather than taste-based discrimination. Price requests that appear to come from high income households raise price quotes for men but not women, eliminating the gender gap. The price gaps between genders falls with the number of nearby repair shops, suggesting that market competition alleviates discrimination.

**JELL codes:** C93, D4, J16 , J18

**Keywords:** competition, discrimination, field experiment, gender

---

*I am very grateful to Emily Oster, John Friedman and Brian Knight for their continuous guidance and invaluable support. I would also like to thank Vivek Bhattacharya, Pedro Dal Bó, Sebastián Fanelli, Bobby Pakzad-Hurson, Jesse Shapiro, Neil Thakral, and Matt Turner for very helpful comments. I also benefitted greatly from interactions with Lorenzo Adelco, Samuele Giambra, Florian Gunsilius, Ryan Kessler, Samsun Knight, Ted Lawless, Stefano Pollini, and Catalina Ravizza. I thank Luciana Casiutti, Josefin Higgs and Santiago Olivieri for help in processing the data. All errors are my own. Email: margarita_machelett@brown.edu.
1 Introduction

In recent years, the fight for women’s rights in the US has expanded from a battle for access to political and economic spheres into a broader-based struggle for gender equity and equality. In part, this stems from a consensus that the failure to achieve equal treatment results from persistent discrimination based on gender. In the workplace, this recent change is most clearly manifest in the form of the #MeToo and Time’s Up movements, where advocates are pushing both for freedom from harassment and for fairness in terms of rights, benefits, obligations, and opportunities (ABC Of Women Worker’s Rights And Gender Equality, Christos, 2009). From an economic point of view, gender discrimination can manifest itself in many forms, affecting women both as workers and consumers.

There have been many studies on pay inequality but considerably less attention has been given to inequities in product and service markets. However, these can have serious consequences for welfare and equity of women. For example, measures of consumer prices in the US typically reveal that women products are more expensive than similar men products: A recent study by the New York City Department of Consumer Affairs (De Blasio and Menin 2015) reports that women’s product cost on average 7 percent more than similar products for men. A prior study by the State of California further claims that women paid an annual “gender tax” of approximately $1,351 for the same products services as men, or about a $15 billion for all women in California.\footnote{This study led California to become the first state to enact a bill to protect consumers from price discrimination for services (CA State Senate 1995, Gender Tax Repeal Act of 1995, AB 1100, Aug 31, 1995). Following suit, many states and counties also passed bills to prohibit businesses from charging different prices for products or services based solely on the customer’s gender.} Some argue that these price differences are evidence of seller’s prejudice against women or discrimination based on a person’s gender (Ferrel et al., 2016), but in fact this is very difficult to prove. In many cases, the products being compared are not identical, and, moreover, it is frequently impossible to disprove that differences in pricing arise from differences observable to sellers but not to researchers. In the case of service markets, this difficulty is even greater as the lack of information about prices paid by each consumer creates yet another obstacle to inference.

In this paper, I circumvent this difficulty by conducting a large-scale field experiment on the
US car repair market to test both the extent and structure of gender-based price discrimination. Building on the correspondence testing methodology (Bertrand and Mullainathan 2004, Jowell and Prescott-Clarke, 1970 and Riach and Rich 1991), I sent tens of thousand emails to shops requesting an estimate for a radiator replacement for the same car. I randomized two customer characteristics observed to shops. First, I varied the perceived customer’s gender using distinctively woman and man names. Second, I also varied the signals about customer “types”. These type signals sought, in particular, to shift shop’s priors on customers willingness to pay. I added pieces of information to a “baseline” script revealing that the customer is searching for other estimates (“quote” type), that the customer is not knowledgeable on car repairs (“uninformed” type), or that the customer is highly-educated to proxy for income (“income” type). Overall, I obtained a sample with 10,313 valid price estimates, corresponding to a response rate of close to 1 quote for every 5 emails sent.

I find that there is a significant gender price-gap only for customers in the “baseline” group. Indeed, women in this group receive price estimates of roughly 486 dollars - about 9 dollars more than men. By contrast, customers in the other groups receive price estimates that do not vary significantly by perceived gender. This result suggests that while being a woman customer may lead shops to charge a higher price, gender becomes less relevant once shops learn additional information, such as customer’s willingness to search or income. Hence, discrimination appears more likely to stem from signal extraction problems (Aigner and Cain, 1977; Arrow, 1973; Phelps, 1972), rather than purely “taste-based” motives (Becker, 1957). Once women signal low search costs (“quote” group customers), they get lower estimates, while men estimates remain the same. By contrast, signaling high income does hurt men, in that they receive higher price estimates. Employee gender, however, appears to have little salient effect. Following Antonovics and Knight (2009) and Anwar and Fang (2006), I find that employee gender does not explain the gender-based price differences, which provides additional evidence towards a statistical discrimination motive for price discrimination.

Furthermore, the gender price gaps are not driven by differential response rates or changes in the composition of the provided quotes. There is no evidence of a gender gap in non-price
outcomes. Using information on the detailed price components, I find that shops do not include more additional parts (on top of labor, a radiator and coolant) depending on customer’s gender. There is also no evidence of differential treatment in terms of offers of discounts, price matching, free inspections, and other such bargains. This non-price discrimination dimension is relevant in my setting, in particular, because car repair services are a typical example of credence goods (Darby and Karni, 1973; Dulleck and Kerschbamer, 2006). That is, consumers can observe utility they derive from the good ex-post but they cannot judge whether the type or quality of the good they have received is ex-ante needed, giving rise to potentially inefficient over-or undertreatment as well as overcharging (Dulleck, Kerschbamer and Sutter, 2011).

I also find a number of additional results relating discrimination patterns to shop types and market structure characteristics. Franchises and dealerships increase prices for women by more than independent shops. This is an unexpected finding, as non-independent shops may have more stringent guidelines and fixed pricing rules. Importantly, I also exploit the cross-sectional detailed nature of the data to construct measures of competition based on the number of competitors that are near each shop and correlate these to the experimental results. I find a significant negative association between competition levels and gender price gaps, and it is not explained by commuting zone fixed effects. Women receive estimates of roughly 1.5 fewer dollars for every additional competitor within a 1-kilometer radius. Noting that there are on average approximately 4 nearby competitors, this is equivalent to an average quote decrease of 6 dollars.

This paper’s contribution to the literature of discrimination in consumer markets is threefold. First, the paper contributes to a small but growing literature that studies discrimination in consumer transactions, and provides new and robust evidence in favor of the hypothesis that price discrimination is both extant and economically meaningful.2,3 This paper extends the findings of studies that show race- and gender-based discrimination in auto and credit markets (e.g., Ayres

While most of this literature focuses on differential call-back rates- predominantly in the rental markets (e.g., Ahmed and Hammarstedt 2008, Bosch, Carnero and Farré 2010 Edelman, Luca and Svirsky 2017, Michael Ewens, Bryan Tomlin and Liang Choon Wang 2014, and Hanson and Hawley 2011)- there is a reduced but growing literature that focuses on discrimination in consumer transactions (involving monetary outcomes). This paper particularly relates to this latter literature.

See Bertrand and Duflo (2017) and Riach and Rich (2002) for reviews of field experiments on discrimination.
and Siegelman 1995, Pope and Sydnor 2011 and Ravina 2007), and provides contrary evidence to non-experimental studies that have not found evidence of discrimination in online markets (Morton, Zettelmeyer and Silva-Risso, 2003).

Second, this paper contributes to the literature studying the mechanisms of discrimination. In line with the findings of Castillo et al. (2013), Gneezy, List and Price (2012), and Busse, Israeli and Zettelmeyer (2017), this paper finds that price discrimination meaningfully changes in response to changes in perceived customer profile. Castillo et al. (2013) send 6 paired men and women testers to negotiate taxi fares using similar scripts in Lima, Peru. Differences in fares disappear once the testers reject the first taxi fare and proceed to the second taxi in line. Busse, Israeli and Zettelmeyer (2017) implement a field experiment in which 4 man and 5 woman callers request price quotes for a radiator replacement from repair shops in heavily populated areas, and find that price gaps disappear once the average US repair price is referenced. Gneezy, List and Price (2012) send 6 disabled and 6 non-disabled testers to request price quotes at 36 auto body shops in Chicago, and find differentials also disappear once individuals mention they are getting other quotes. This paper further advances upon the findings of these studies by interacting gender with other customer characteristics and combining the experimental results with novel geographic, shop, and market characteristics.

The correspondence study approach also ensures more robust comparability across treatment groups than was available in prior papers, guaranteeing that any observed differences are caused solely by the gender and customer type trait manipulation. Furthermore, this methodology allowed me to reach shops across the US, expanding on the previously more localized studies.

This paper also relates to the empirical literature on price dispersion and competition. In contrast to a number of prior studies showing that price dispersion varies directly with the level of competition in a market (e.g., Borenstein and Rose 1994, Stavins 2001 and Shepard 1991)-in particular in the airline and retail gasoline market-, this paper joins the studies finding that increases in competition decrease price dispersion (Barron, Taylor and Umbeck 2004, Gerardi and

\footnote{Many of the weaknesses of audit studies have been discussed in Heckman (1998) and Siegelman and Heckman (1993).}
Shapiro 2009). In addition, my setting differs from the previous literature by relating competition to experimental results, and by allowing sellers to directly price discriminate between observable types of consumers, in particular woman versus man consumers.5

Finally, this paper also relates to the literature on credence good markets, where information asymmetries between buyer and sellers may lead to inefficient provision of the good as well as overcharging.6 My empirical setting relates closely to Schneider (2012), who conducts visits to car repair shops knowing in advance the set of car defects. He finds evidence of poor diagnosis, which does not vary when posing as a one-time versus a repeat-business customer. My analysis confirms these observations, in that recommended repairs do not vary with customer characteristics. In a broader perspective, I complement this literature by providing further evidence on how prices vary with other consumer characteristics, such as changes in search costs and income.

The remainder of the paper proceeds as follows. The next section gives an overview of the experimental design, implementation and discusses its main concerns. Section 3 provides descriptive statistics and empirical results. Section 4 concludes.

2 Experimental design

Building on the correspondence testing method, I conducted a field experiment on the US car repair market to test if there is evidence of a gender-based price gap and how it varies with additional information about customer and market characteristics. With this aim, I scraped the online Yellow Pages (www.yellowpages.com) and created a comprehensive dataset on car repair businesses that details the email address of each shop together with information on shop location and type of service provided. Between July and August 2018, I sent emails to more than 57,000 repair shops inquiring about the cost of replacing the radiator of a Honda Accord, a popular car in the US. I randomized two customer characteristics observed to shops. First, I varied the

5 In an alternative setting, Doleac and Stein (2013) implement a field experiment to examine how consumers discriminate sellers that post online iPod advertisement based on their race. They define the degree of competition by the offers consumers make to each seller received.

6 See Dulleck and Kerschbamer (2006) for a review on this literature
perceived customer’s gender using distinctively women and men names. Second, I also varied signals about customer “types”. These signals are intended to shift shops priors on customers willingness to pay, and represent a customer who is searching for other estimates (“quote” type), a customer who is not knowledgeable on car repairs (“uninformed” type), or a customer who is highly educated to proxy for income (“income” type). These variations added to a standard (“baseline” type) customer allow me to test how prices vary with additional information about customers.

Overall, half of the shops replied to my emails (including all types of replies), and I obtained a sample with 10,313 valid price estimates, corresponding to a response rate of close to 1 quote for every 5 emails sent. The remainder of the section reviews the experimental details.

2.1 Auto repair shops data

I collected information on all US auto repair shops listed in the online Yellow Pages (YP) (www.yellowpages.com). The YP offer an comprehensive listing of businesses nationwide and are used by nearly 60 million customers in the US each month, with auto repair shops being one of the popular categories.

As of June 2018, YP had over 230,000 listings in the auto repair category. Each listing details information on the shop’s name, location, website, type of services provided. Importantly, the YP also includes the shop email address whenever available.

I imposed several restrictions on this collected sample. First, I used the information on type of services and shop names to keep shops specialized in mechanic repairs for cars, for instance, I dropped body paint shops, and truck and RV repair shops. I also restricted the sample to keep shops that do not share the email address or locations with other shops. This is meant to avoid the possibility of contacting the same repair shop twice. My final sample counts approximately 58,300 auto repair businesses. Figure 1 shows the location of the shops in my sample. Shops are dispersed across the US, with more facilities located in more popular areas. 

---

7 Alaska and Hawaii are included in the sample. Shops in Puerto Rico are excluded from the sample as they would expect to interact with their customers in Spanish.
Additionally, I classified each repair shop as independent or non-independent: Franchises, Honda and non-Honda dealerships using the shop name.\textsuperscript{8} This is to test if price patterns vary with shop types. Independent shops may have more discretion setting up prices as they do not have stringent pricing rules, and it is easier for a customer to directly reach the mechanic in charge of a repair.

Last, I built market competition measures using GPS software. With Google Maps geocoding API, I converted addresses into geographic coordinates (latitude and longitude), and then used QGIS to compute each shop’s distance to its competitors to obtain the number of nearby competitors in alternative distance radiuses. This is meant to explore the correlation between competition and price gaps.

2.2 E-mail design

As the second step of this experiment, I wrote a baseline email script to request a quote for a Honda Accord radiator replacement. I then created variations of this baseline script along two dimensions. First, I varied the perceived customer’s gender by using names that are distinctively recognized as female or male. Second, I varied the signals about customer types, distinguishing between customers with no car knowledge, searching for other quotes, and or highly educated (to signal income).

2.2.1 Customer gender variation

The first variation I introduced refers to the perceived customer gender. I used 18 popular names from Bertrand and Mullainathan (2004) and Levitt and Dubner (2006). All these names are distinctively female or male, and they are all white-sounding names to only vary gender perception.\textsuperscript{9}

\textsuperscript{8}I gathered listings of US franchises from online searches in websites such as entrepreneur.com and franchisegator.com. All franchises are available in the YP and a rough comparison suggests that the YP have in fact more shops than those currently open according to online searches. These likely corresponds to closed businesses that are still listed in the YP.

\textsuperscript{9}I use all the white-sounding first names in Bertrand and Mullainathan (2004) except for Meredith, Laurie, Brad, Brett, Jay and Todd as they have a small relative likelihood of being identified with the intended gender, according to online searches (e.g., gpeters.com). To extend the list of names, I chose six white-sounding names from Levitt
I created an email account for each name, of which half are associated with each gender. These names can be observed by the seller in the email address, its account name, and email signature.

### 2.2.2 Customer types variation

The second variation is the customer type, which is either a *baseline* type, an *uninformed* type, a *quote* type or an *income* type. These types are signaled through variations in the email content (i.e., scripts). The variations aim at shifting each seller’s prior of consumers willingness to pay. They are used to compare how gender price gaps vary, with the baseline type as the comparison group. While scripts vary the conveyed information on knowledge, income and quote search efforts, they are otherwise identical in structure and content.

Figure 2 presents a template script used with the *baseline* customer. The *baseline* customer script mentions a car model, radiator problems, asks for a total cost estimate for its replacement and also mentions living close to the shop. The car is one of the most popular US cars - a Honda Accord. The problems are leaking coolant and overheating. Both problems can be associated with a faulty radiator, according to the Federal Trade Commission and RepairPal website, among others. The reference to the residence near the shop is used to increase the apparent veracity of the email.

All remaining customer types include the baseline script with slight variations. The *quote* script adds a sentence to the baseline script where it states to be searching for other price estimates nearby. A subset of these *quote* types also mentions having checked a website with the area price references (RepairPal), signaling further market knowledge. Thus, shops know with certainty that customers in this group are exerting some effort to obtain more competitive estimates signaling low-search costs. The *income* type uses the baseline script and adds the title “Ph.D.”

---

and Dubner (2006). According to the Social Security Administration, each of the additional names is among the 100 most popular baby boy and girls names as in Massachusetts between 1970 and 1986, in consistency with Bertrand and Mullainathan (2004) study. The complete list of names is Allison, Amy, Anne, Brendan, Carrie, Colin, Dustin, Emily, Geoffrey, Greg, Heather, Kristen, Matthew, Neil, Sarah, and Scott, and the last names are Baker, Kelly, McCarthy, Murray, O’Brien, Ryan, Sullivan and Walsh. Using first and last names that are predominantly associated with white individuals allows me to use them uniquely as signals on gender. Finally, all accounts use the same gmail domain, and a number or letter randomly added to the email address due to unavailability of first-name last-name only e-mail combinations.
to the individual’s email signature. The Ph.D. title intends to proxy for high-income, signaling a higher opportunity cost from search. Finally, the uninformed type differs from the baseline script by describing the car problems using vague, non-technical terms (i.e., “the car leaves green puddles” instead of “it is leaking coolant”, and the “temperature thing rises” instead of the “gauge”). This type signals no knowledge of cars, presumably being easier to be overcharged or offered unnecessary additional repairs. The script templates were pre-tested in a pilot and all templates are available in Appendix A.2.

2.2.3 Other email items variation

Finally, in addition to varying customer’s gender and type, I also alternated some email features used to contact each shop. The purpose of this inclusion is to provide small variations to the emails while not affecting shop’s inference of each customer’s gender and type. These variations are the script version for each customer type, the email subject and the car year. There are two scripts for each customer type that provide the same information but are reorganized and slightly differently worded. The email subjects all refer to a radiator, but vary the words “new,” “replacement” or “change.” The car years are 2009 and 2010, which belong to the same car generation.

In summary, each shop is contacted by a customer equal in all aspects except for the treatment group (gender and customer type combinations) and these email characteristics. All email items are included as controls in regression specifications, and the results are robust to their inclusion. Furthermore, none of these controls are statistically different from zero.

2.3 Randomization

Once all email variations were determined, I randomly assigned one of the eight treatment groups to each shop, unconditional on any characteristics. As a first step, half of the sample was assigned to be approached by either gender, with one of the created email accounts. Then, I assigned customer types within each gender. I used optimal shares obtained from power calculations
which weighted more the baseline and quote types. This resulted in a distribution of shops to be contacted by 18.29% baseline, 15.73% quote, 8.24% uninformed, and 7.75% income scripts of each gender. Finally, I reshuffled the sample and assigned one of each remaining email items (car year, script number and email subject). Appendix A.1 further details this procedure.

2.4 Implementation and data collection

With all treatment groups assigned to contact each shop, I sent approximately 57,400 emails using automated tools. This wave of emails took place from July through August 2018, sending an average of 8,000 weekly emails between 10:30-11:30 AM Eastern time from Monday through Friday.\textsuperscript{10}

Besides following a protocol to send emails, I also followed pre-specified email reply guidelines. These guidelines guarantee that all treatment groups behave identically, and thus, price estimates only vary with information on customer’s gender and types. Appendix A.3 details these rules and includes the reply template scripts.

Once the interaction with shops finished, I collected information about each reply. The information can be divided into three main categories: reply type and employee first name, mentions to original email items, and price and service details.\textsuperscript{11}

The reply type information measures if, when and which type of answer is received (invalid, do not perform service, perform service, and perform service and give quote), as well as how much effort customers exert to receive a quote (measured by an indicator equal to one if a response is sent to a shop before receiving a quote). Employee first name is recorded to assign them their likely gender. Based on the first name, I assigned a gender to 80 percent of the em-

\textsuperscript{10}This project does not require IRB approval since it does not meet the definition of human subjects research as defined in Title 45 CFR 46.102(f), as verified with Brown Human Research Protection Program office.

\textsuperscript{11}With respect to the data collection, I divided emails into invalid email, not do service, do service and give quote, do service and not give quote gmail labels for each week. These are used to estimate reply rates by email reply content. I automated the weekly download of every email thread into a spreadsheet, with some variables automatically filled in based on the email’s text. These were later reviewed and completed manually. Three research assistants, unaware of this project hypotheses, helped with this last task. As much as possible, they reviewed give quote label emails, while I reviewed the remaining labels. Nevertheless, all collected information is not subject to an individual’s interpretation and is easy to verify (i.e., a shop either gives a price or not, mentions a discount or not, etc).
ployees, of which 18 percent correspond to woman first names. The gender is assigned based on the SSA National Data on relative name frequencies by gender in the population of U.S. births from 1940 through 2005. If a name is associated with women in more than seventy percent of the occurrences, then that name is classified as corresponding to a woman. By the same token, I identified man names and left undefined those replies without an employee name or those whose name’s relative frequency is above thirty percent but below seventy percent.

Regarding the mention to original email items, I collected indicator variables if the shop reply mentions the customer’s name, gender (e.g., Mr, Miss) or profession - in the case of high-income type customers. I also measured if they make a reference to RepairPal. These variables intend to capture whether these the experimental design features are salient to shops.

Finally, price and service details refer to the information related to pricing and other offers. Regarding pricing, I collected the total price estimate, price ranges (when offered) and price component details. Total price estimates were defined using pre-specified rules. These include using the total discounted price provided in the estimate, the average price when price ranges are provided, the first price provided by a shop whenever the same shop provides more than one quote in separate emails, dropping prices that explicitly exclude either the labor or the radiator price in the estimate, extremely low or high prices (below $100 and above $2,000) and estimates where the price range is twice the range ratio.\textsuperscript{12} These thresholds are arbitrary but conservative values. The last rule with regards to total prices refers to cases where price match is offered. I keep the price provided by the shop, regardless of additional price-match offers. Finally, the service offers document if the shop price matches competitors estimates, mentions their service has a warranty, a discount or any additional offer such as a car loaner or shuttle service. The latter is used to test whether treatment groups receive more benefits. All the collected information is then matched to the shop data.

\textsuperscript{12}Less than 127 price estimates are considered invalid, of which 92% are incomplete quotes, 1 exceeds $2,000 and 9 price ranges are too wide without providing any explanation for this variation. Slightly more than 20 shops provide 2 different estimates in separate emails. As specified, the first estimate is reported.
2.5 Final sample and design validation

The experiment resulted in a response rate of one half, and a valid price estimate in close to 1 out of 5 emails, yielding 10,313 valid price estimates. While the aggregated follow-up rates are similar across groups, there are some differences in the rate of provided quotes, as illustrated in Figure 3. First, women in the baseline group receive 0.009 percentage points more replies than men in the baseline group. This difference is only marginally significant when comparing the raw data. As soon as I control for state fixed effects, this difference is no longer significant. Second, women in the uninformed group receive an average of 0.01 percentage point less estimates than in the baseline group. Lastly, there is a 0.015 percentage point decrease in the reply rate for customers in the quote-search group relative to the baseline group. These small differences may introduce selection problems, which are discussed in the next section.

For the sample with price estimates, there are 3 main issues of potential concern. First, are shops used to communicating by email and providing estimates? Second, is the experimental design valid - that is, do shops perceive as salient the variations by gender and customer type? And third, is the randomization balanced across treatments?

For the first concern, I document that shops performing Honda radiator replacements in the final sample use email communication on a frequent basis. Table 1 shows that approximately 90 percent of shop replies are received within 2 days since the first email inquiry. In addition, less than 1 percent of shops mention they do not usually correspond by mail with customers. Furthermore, more than 80 percent of quotes are obtained without the need of sending a follow-up email and almost 10 percent of replies also provide an estimate in an attached document (oftentimes through their internet quote system). Altogether, this suggests that shops in the sample use and monitor their email frequently, are used to communicating with customers by email, and do not seem to find the email price requests suspicious.

With respect to the second concern, I find evidence suggesting that employees pay attention to email information. Table 1 shows that half of the replies mention the customer’s name, indicating that this feature is salient to shops. The name mentioned corresponds to the assigned name
and not to that included in the email address - which includes an additional letter or a number. About 3 percent of shops make a reference to the customer’s gender and to the professional title (in the case of income customer types), and about 2 percent of shops explicitly mention RepairPal website after the initial email does so. In these latter cases, shops usually refer to where their estimates stand relative to RepairPal area price range. Finally, shops that send an attached estimate include car details mentioned in the initial email. This is not a feature intended to be salient, but also suggests that employees read emails carefully.

Regarding the third and last concern, I provide evidence that the design is valid. First, the geographical distribution of shops across groups seems balanced by gender, and within each customer type (see Figures A.2 and A.3). All groups are distributed similarly across the US, with more competitors in more populated markets. Reassuringly, there are no significant differences by gender or gender-customer type with respect to the share of invalid emails, which further validates the randomization (See Figure A.4). Finally, 1.7 percent of emails originally assigned to a shop were not sent due to implementation issues such as having one account disabled. This discrepancy between assigned and sent emails, however, did not modify the distribution of treatment groups.

3 Results

3.1 Is there a gender price-gap?

Figure 4 presents the main results. Panel (a) shows the mean price estimates by gender and customer type, while panel (b) plots the gender price-gap for each type. This figure shows that there is a gender price-gap only for customers in the “baseline” group. Indeed, women in this group receive price estimates of roughly 486 dollars - about 9 dollars more than men. By contrast, customers in the “uninformed”, “quote”, and “income” groups receive price estimates that do not vary significantly by perceived gender. This result suggests that being a woman customer may lead shops to charge a higher price, but gender becomes less relevant once shops learn additional
information, such as customer willingness to search or income. Indeed, Figure 4 shows that shops change their quotes based on perceived income and search costs. Signaling low search costs (i.e., “quote” group) helps women get an estimate close to 476 dollars, about 10 dollars lower than their baseline. men, by contrast, do not profit from this signal. On the other hand, signaling high income seems to particularly hurt men, who receive estimates of 488 dollars, similar to women in the baseline group. Finally, revealing ignorance on car repairs do not seem to have an impact on prices, although these estimates are noisier. Overall, these results could suggest that shops priors are that, absent additional information, woman customers have a higher willingness to pay. When information about search costs and income is provided, shops update their priors resulting in no price differentials by gender.

The implied gender gap in the baseline group is a small but a significant amount per repair. The 9 dollar gender differential in the baseline group represents a 1.9 percentage gap. This represents clear evidence that gender-based price discrimination still persists, but also that it is significantly lower than other estimates in the academic and policy literature. In a related setting, Busse, Israeli and Zettelmeyer (2017) find a larger gender price gap of almost 6 percentage dollar increase for women. This difference may be due to the changes in discrimination over time, as Busse, Israeli and Zettelmeyer (2017) experiment took place 6 years ago, in 2012. Another important feature that may be driving the difference is that I used email communication while they called shops. Using email communication may increase transparency and decrease shops discretionary behavior as estimates are written, and help protect disadvantaged groups (Morton, Zettelmeyer and Silva-Risso, 2003). Thus, my results suggest that internet communication is particularly beneficial to individuals whose characteristics disadvantage them in personal interactions, although it does not completely remove price differences.

---

13 The cleanest comparison between Busse, Israeli and Zettelmeyer (2017) and this paper is between my “baseline” group and their “uninformed” group, where agents request a quote without providing a reference price. This is the number I am reporting here.

14 One should be careful, however, of drawing too strong conclusions from this comparison, since there are important differences in the samples used in both studies. In particular, Busse, Israeli and Zettelmeyer (2017) focus on very populated areas, while I obtained estimates from 10,313 shops across the US, including less populated areas. Furthermore, since their study was implemented in 2012, one cannot rule out that the decrease in price discrimination is partially driven by changes over time, as suggested by Edelman, Luca and Svirsky (2017) in a related setting.
These results are robust to the inclusion of experimental design components as well as to state and commuting zone fixed effects. Appendix Section D replicates the analysis in a simple regression framework. Figure A.9 plots point estimates from a regression-adjusted specification of price estimates for each group relative to men in the baseline. The main results are similar. Table A.4 shows additional robustness checks. Column (1) controls for state fixed effects, additional varying email items (car year, subject controls and script version used), and day of week and week fixed effects indicating when each shop is contacted, as pre-specified in the analysis plan. Column (2) adds a non-independent shop type control (i.e., franchises and dealerships) since their price levels are usually higher, and column (3) adds Commuting Zone (CZ) fixed effects and an additional indicator dummy for the shops that were not matched to any commuting zone. These smaller geographic areas are included to control for economic diversity better. Overall, the regression-adjusted results are consistent with the raw price differences shown earlier. Indeed, the gender-based premium in the baseline group remains significant across specifications and in fact increases to an average of 9.6 dollars. The price decrease for women in the quote search group is marginally significant, while the price increase for men in the income group is statistically significant and varies between an average of 11.7 to 13.5 dollars relative to men in the baseline group. Reassuringly, results are robust to the inclusion of controls, and none of the additional experimental design features are relevant to explain variation in prices.

A potential concern may arise from the differences in the rate of provided quotes, although these differences are not statistically significant in all specifications, as discussed in Section 2.4. This may introduce a problem of sample selection bias. For instance, women in the baseline group are 0.009 percentage points more likely to receive a reply with an estimate. It is possible that any given shop does not price-discriminate between men and women, but that the most expensive shops only reply to women. To address this concern, I implement a correction for selection following Heckman’s two-step approach. I do this for the comparison groups that have significant differences in the rate of quote replies: women and men in the baseline group, quote-search relative to baseline, and uninformed women relative to their baseline. Table A.5 in the Appendix shows the results. The results suggest that gender-price gap in the baseline group is
not driven by selection bias. However, the inverse Mills ratio coefficient indicates that there is some evidence of sample selection.\textsuperscript{15}

An additional exercise tries to distinguish between taste-based and statistical discrimination motives, as empirically tested in the literature (see Antonovics and Knight 2009, Price and Wolfers 2010, among others). These studies compare the treatment by different groups of potential discriminators, which are distinguished by the same observable characteristic in which discriminatory behavior is suspected. In my setting, I can test if the employee’s gender matters to explain gender-based differences in prices. If statistical discrimination alone explained differences in treatment, then everything else equal, treatment differences should be independent of the group characteristics. For instance, woman employees should not help other woman customers get better deals. I test this hypothesis in a simple regression where I include a dummy for perceived woman employees and its interaction with a woman customer dummy. Table 2 presents the results. I find that there is no significant change in price-gaps when a woman employee gives a quote. This provides no support to a taste-based motive driven price-gap. In addition, the observed price differences across treatment groups are robust to the inclusion of these controls.

Finally, I explore whether the main results are associated with observed shop types. Is it a common practice for dealerships and franchises to price customers uniformly? Presumably, independent shops have more scope to update estimates more frequently. I study these questions in Table 2. With respect to pricing patterns by shop type, I observe that, as expected, non-independent shops provide men with estimates of about 90 dollars higher than independent shops. Unexpectedly, however, the relative gender difference increases about 28 dollars more among non-independent shops. The reason for this rather surprising result is not obvious, but I can rule out some explanations. First, there is no differential reply rates of non-independent shops. Second, I can rule out that it is driven by systematic differences in the composition of quotes. That is, franchises and dealerships do not include more additional replacement parts in the price

\textsuperscript{15}I also implemented a non-parametric bounds estimation following Lee (2009). Unfortunately, unlike Heckman’s parametric correction, I can only use a limited set of covariates to tighten the estimates of the bounds. As a result, the difference between both bounds is large. I cannot reject the null of no effect in the lower bound for any of the groups with significant differences in the rate of quote replies.
estimates provided to women than those provided to men.

3.2 Do increases in competition levels help close price-gaps?

The relationship between gender-price discrimination and competition is theoretically ambiguous. While a monopolist may discriminate across customers and a perfectly competitive firm cannot discriminate at all, the intermediate cases are much less clear (Borenstein 1985; Borenstein and Rose 1994; Holmes 1989). Therefore, the relationship between price discrimination and competition is an empirical question. While I do not have exogenous variation in the level of competition, I have rich detailed data on the location of each car repair shop in the US. This allows me to shed light on this question.

In my setting, I define competition levels by the number of competitors in each market and each market for a radiator repair by 1-km radius circles centered around each shop. Table 3 presents the main results. Each column regresses price estimates on the number of competitors, and its interaction with a woman dummy, controls for the type of competitors nearby (i.e., whether they are independently-owned or not) and the same controls as in the previous subsection analysis. I find that an additional competitor is associated with a significant decrease in gender price gaps. Indeed, women receive estimates of roughly 1.5 fewer dollars with an additional competitor nearby. Noting that there are on average approximately 4 nearby competitors, this is equivalent to an average quote decrease of 6 dollars. Furthermore, the gender-gap in the baseline group increases to 15 dollars.

This result suggests that while consumer heterogeneity may lead shops to price discriminate, the extent of competition also matters. This supports the hypothesis that the price discrimination is negatively associated with the degree of competition. Unlike our previous results comparing prices changes by employee gender and customer types, these results are also consistent with Becker (1971) model; discrimination against women is less present in markets with more competition, as sellers for whom it is more costly to interact with women will be competed out by sellers who do not discriminate. In my setting, the degree in which shops vary prices with per-
ceived gender also depends on the ease with which customers can substitute stores. As it becomes easier for customers to switch to a competitor shop, increasing prices based on gender becomes less profitable. Appendix B illustrates this mechanism in the context of a simple model in which statistical discrimination is allowed.

I show the robustness of the association between price-gaps and competition in Table A.6. First, I verify that the observed pattern is not driven by outliers and influential points by running a robust regression, which reweights observations based on how well behaved they are (Fox, 1997). Column (2) shows that results are similar. Second, I show that the results are qualitatively similar if I restrict the sample of shops used to compute the competition measures to those that at least have an email or website in their Yellow Pages listing. Third, I show that results are robust to alternative definitions of the threshold used to compute the competition measure. Column (1)-(5) re-run the regression in Table A.7 by using a threshold of 1, 1.5, 2, 2.5, and 3 kilometers, respectively. Results are qualitatively similar, but quantitatively the results get dampened with larger distances. I explore this further by defining mutually exclusive rings of these same circumferences and regressing prices on them. Column (6) shows that the results are driven by the innermost ring.

Another potential venue to explore is how discrimination varies with other geographic characteristics. For instance, are women more discriminated against in areas where more single parents live? Are more ethnically diverse places associated with more equal prices? Or maybe gender price gaps are higher in places where commuting to work takes longer. A first glance at this suggests that geographic characteristics are not strongly correlated with gender price gaps. Figure 5 shows the price differential for each commuting zone, which has no evident pattern to other commonly studied characteristics. The price differential is obtained from a regression of commuting zone fixed effects interacted with a woman dummy, controlling for commuting zone. Figure A.7 shows this differential aggregated at the state level.
3.3 Are there gender differences in diagnostic recommendations and customer service?

Repair services are defined as credence goods. That is, they have the characteristic that consumers can observe the utility they derive from the good ex-post, they cannot judge whether the type or quality of the good they have received is the ex-ante needed (Dulleck, Kerschbamer and Sutter, 2011). Car repair services are therefore markets characterized by asymmetric information between sellers and consumers. This can give rise to inefficiencies, such as overtreatment of services, i.e., to cases where more than necessary parts are offered. While this paper experiment only requests a quote for a radiator replacement, mechanics can also recommend and quote additional parts. For instance, besides quoting labor, coolant, and a new radiator, they may also include a new thermostat, hoses or a radiator cap. This brings another channel through which shops can treat customers differently. Do shops include more items when providing quotes to women? Does signaling no market knowledge also result in more replacement recommendations? Does signaling effort to obtain other quotes result in less additional replacements? They can treat customer differently by also changing complementary services they offer. Do women get more quotes with warranties on the service? Do they receive more offers?

Overall, I do not find differential recommendations by gender. Figure 6 shows that there is no difference in rates in which taxes and fees, additional replacement parts, warranties, and additional offers are provided. This suggests that shops do not provide additional components to women, nor do they seem to compete more aggressively to earn their business. There are also no differences in the amount of other recommended replacements by customer types. Table A.8 in the Appendix shows that shops are less likely to provide a detailed quote to uninformed man customers, while they are marginally more likely to offer these to high-income customers and to women in the baseline group. This difference, however, does not translate into decreases (or increases) in additional recommended replacements relative to man customers in the baseline group. In the case of uninformed customers, it could suggest more skepticism in the customer self-diagnosis or suggest a smaller commitment from shops to keep effective prices close to
estimated prices. The observed lack of differences in quoted parts resembles Schneider (2012) findings. In a different but related setting, he does not find evidence that mechanics suggest more or fewer repairs among one-time and repeat-business type customers.

Though customers do not receive additional parts differently, there are a few small but significant differences in the inclusion of taxes and fees, and additional services. High-income customers receive more quotes that include shop fees and taxes than customers in the baseline group. This increase does not explain the observed differences in price estimates. The second difference is that men searching for nearby estimates are more likely to be quoted a service with warranty and additional offers, which include discounts, price matching, and financing options. This could suggest that when man customers signal the potential to substitute stores shops also try to compete more in the quality of customer care. Importantly, these non-monetary findings indicate the discrimination in prices is not driven by differential responses or additional repairs included in the quotes.

4 Conclusion

In this study, I present evidence that woman customers receive worse estimates than their man counterparts when no information other than gender is provided. Specifically, women receive price estimates of roughly 486 dollars - about 9 dollars more than men. These effects, however, disappear once additional information about the customer is revealed. Once customers additionally signal to be in the “uninformed”, “quote”, or “income” groups, they receive price estimates that do not vary significantly by perceived gender. women particularly benefit when they mention they are searching for other price quotes, they receive estimates about 10 dollars lower. men, by contrast, do not profit from this signal. On the other hand, signaling high income seems to particularly hurt men, who receive estimates of 488 dollars, similar to women in the baseline group. Overall, these results could suggest that shops priors are that, absent additional information, woman customers have a higher willingness to pay. Once additional information is revealed, gender becomes less relevant.
By considering the extent to which price gaps vary with shop and market characteristics, I provide further evidence on what other features explain discrimination patterns. I find that results are not affected by employee gender, further suggesting that discrimination arises through a signal extraction problem, rather than simple prejudice against women. That is, employee’s preferences for a specific customer gender are not driving the results. Consistent with previous literature of discrimination in car sales, I find that dealerships and franchises discriminate against women, and this bias is greater than in independent shops. Importantly, the disadvantages faced by women are reduced when markets are more competitive, providing some support for Becker’s test-based discrimination motive.

As with all such correspondence studies there is an important consideration to these findings. The outcome observed is only an intermediate outcome. That is, I compare price estimates and not the final prices paid once a repair is done. It is possible that price estimates are not equivalent to transacted prices, which is ultimately the value we are most interested in. However, while mechanic shops have some incentive to provide low prices to earn a new client, they also have an incentive to provide an estimate close to the final price. If they surprised their clients with much higher quotes, then they might risk losing them. In addition, this may be more difficult in my setting, as estimates are written and not given over the phone call. A discrepancy between estimates and final prices would be problematic in one case. Discrimination against women would disappear - or be reversed - if we thought that shops increased prices by more to men than to women. Discrimination against women, on the other hand, would increase if we thought that shops increased final prices by more to women. One may speculate that the results are a lower bound on the extent of price discrimination when only in-person visits occur. The internet facilitates information search, and has been found to be particularly beneficial to those whose characteristics disadvantage them in negotiating (Morton, Zettelmeyer and Silva-Risso, 2003).

In addition, there are important considerations about the generality of this paper findings. One should interpret with caution the representativeness of this paper results. While observed estimates come from shops throughout the US, they represent less than 10 percent of all repair shops in the US. Only 40 percent of the yellow pages listings have an available email address,
and close to 1 out of 5 emails sent replied with an estimate. This motivates further comparisons between shops, and an in-depth exploration on discrimination patterns and geographic characteristics.

Overall, this paper shows that gender-based price discrimination still exists despite increased progress towards gender equity and equality. Such discrimination patterns seem to arise through shops’ prior beliefs about men and women willingness to pay rather than a distaste for serving women customers. Gender becomes less relevant once additional information on search costs or household income is observed. Furthermore, revealing that a customer is searching for other price quotes helps customer get more competitive prices and closes the gender-based premium, and the gender premium also decreases with increased competition. In this setting, competition is associated with lower consumer
References


Ravina, Enrichetta. 2007. “Beauty, personal characteristics and trust in credit markets.”


Figures

Figure 1: Distribution of shops

Notes: This figure shows the distribution of the 58,324 shops included the final sample. Source: Online Yellow Pages.

Figure 2: Baseline customer template script

Hi,

I need to replace my car’s radiator. It had cracks fixed before, but now it is overheating and leaking coolant again. I’d like to directly get a new radiator. The car is a $(FILL\_YEAR)$ 6-cylinder, automatic, sedan Honda Accord.

Would you please tell me what is the total replacement cost? When can I stop by? I live close to your shop, here in $(FILL\_CITY)$. Thanks!

Best,
$(FILL\_NAME)$

Customer name (gender)

Shop specific location

Car model
Figure 3: Price estimate reply rates

Notes: This figure shows average shop replies with price estimates across gender and customer types: Baseline, uninformed, quote and high-income. The plot shows the 95% confidence intervals. The sample mean reply rate is marked by a grey line within the plot.
Notes: These figures show the price estimate differences across gender and customer type. Panel A plots raw average price estimates with their 95% confidence intervals separately for each group. The sample average price estimate is marked by a grey line within the plot. Panel B plots the average gender price gaps with their 95% confidence intervals for each customer type. Positive values imply a higher price estimate for women with respect to men in that group, and the case with no gender price gaps (y-axis = 0) is marked by a grey line within the plot.
Figure 5: Price differential against women by commuting zone

Notes: This map presents coefficient estimates from a regression of price estimates on commuting zone fixed effects interacted with a woman dummy, controlling for commuting zone fixed effects. The sample is restricted to commuting zones with at least two estimates by gender. Larger coefficients (blue) represent higher prices for women relative to men. Bins have equal counts of observations.

Figure 6: Additional parts and service offers

Notes: This figure shows share of estimates that include: i) taxes and miscellaneous shop, ii) additional parts: radiator cap, radiator hoses and thermostat, iii) a warranty, iv) any additional offer, including a discount, a free inspection, price-matching, and others such as financing options or a shuttle, for men and women separately. The plot shows the 90% confidence intervals, obtained with robust standard errors.
### Tables

**Table 1: Summary statistics: Shops email use and salient design items**

<table>
<thead>
<tr>
<th></th>
<th>(1) Total</th>
<th>(2) men</th>
<th>(3) women</th>
<th>(4) p-value</th>
<th>(5) N all</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shop replies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On 1st day</td>
<td>0.76</td>
<td>0.76</td>
<td>0.75</td>
<td>0.17</td>
<td>14,220</td>
</tr>
<tr>
<td>On 2nd day</td>
<td>0.14</td>
<td>0.13</td>
<td>0.14</td>
<td>0.07</td>
<td>14,220</td>
</tr>
<tr>
<td>Not use email</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.48</td>
<td>14,220</td>
</tr>
<tr>
<td>Price on 1st reply</td>
<td>0.83</td>
<td>0.83</td>
<td>0.82</td>
<td>0.16</td>
<td>10,313</td>
</tr>
<tr>
<td>Price on attachment</td>
<td>0.10</td>
<td>0.09</td>
<td>0.10</td>
<td>0.04</td>
<td>14,220</td>
</tr>
<tr>
<td><strong>Shop mentions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>0.51</td>
<td>0.49</td>
<td>0.53</td>
<td>0.00</td>
<td>14,220</td>
</tr>
<tr>
<td>Gender</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>14,220</td>
</tr>
<tr>
<td>Profession</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.05</td>
<td>2,269</td>
</tr>
<tr>
<td>Repair Pal</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>1,936</td>
</tr>
</tbody>
</table>

*Note:* This table reports the shares of responses from shops that do Honda radiator replacements for all customers (col 1), men (col 2), and women (col 3). Column 4 reports the p-value from each statistic means comparison test by gender. Col. 5 reports the total number of observations. Price on 1st reply row restricts observations to shops that provide valid quotes, Profession row restricts to observations where “income” treatment is assigned to each shop, and Repair Pal row restricts to observations where initial email uses the “quote-search” customer type treatment with the “repair pal” search reference.
Table 2: Effects by shop characteristics

<table>
<thead>
<tr>
<th>Dependent variable: Price estimates</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-independent shop</td>
<td>93.550***</td>
<td>95.587***</td>
<td>95.934***</td>
</tr>
<tr>
<td></td>
<td>(6.919)</td>
<td>(7.049)</td>
<td>(7.052)</td>
</tr>
<tr>
<td>Non-independent shop × Female</td>
<td>28.438***</td>
<td>27.675***</td>
<td>27.671***</td>
</tr>
<tr>
<td></td>
<td>(9.748)</td>
<td>(9.976)</td>
<td>(9.981)</td>
</tr>
<tr>
<td>Female employee</td>
<td>7.160</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.389)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female employee × Female</td>
<td>-0.167</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.388)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>7.211*</td>
<td>7.487*</td>
<td>7.581*</td>
</tr>
<tr>
<td></td>
<td>(4.199)</td>
<td>(4.257)</td>
<td>(4.410)</td>
</tr>
<tr>
<td>Controls</td>
<td>State FE</td>
<td>+ CZ FE</td>
<td></td>
</tr>
<tr>
<td>Mail items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOW, Week FE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.083</td>
<td>0.104</td>
<td>0.104</td>
</tr>
</tbody>
</table>

Notes: This table reports coefficients from regressions of total price estimates on shop characteristics and their interaction with an indicator for woman customers, an indicator for woman customers, customer types and their interaction with woman customer (with baseline customer type as the omitted category). Column (1) shop characteristics refers to non-independent shop types, and includes state fixed effects, other email varying items: car year, subject and script number controls, adds day of week and week fixed effects - indicating when the email was sent to a shop. Column (2) adds Commuting Zone fixed effects with an additional dummy for shops not matched to a Commuting Zone. Column (3) adds an indicator for identified woman employees, and it’s interaction with providing a quote to a woman customer. Robust standard errors in parenthesis, with significance levels: *p < 0.10, **p < 0.05, ***p < 0.01.
Table 3: Effects on price gaps with competition

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Price estimates</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitors</td>
<td>-0.682**</td>
<td>-0.291</td>
<td>-0.669**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.343)</td>
<td>(0.339)</td>
<td>(0.336)</td>
<td></td>
</tr>
<tr>
<td>Competitors x Female</td>
<td>-1.522***</td>
<td>-1.462***</td>
<td>-1.501***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.464)</td>
<td>(0.459)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>15.329***</td>
<td>15.135***</td>
<td>15.594***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.576)</td>
<td>(4.483)</td>
<td>(4.538)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mail items</td>
<td>State FE + Shop type + CZ FE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOW, Week FE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.044</td>
<td>0.087</td>
<td>0.110</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This reports coefficients from a regression of total price estimates on the number of competitors within a 1-km radius, its interaction with a woman indicator, a woman indicator, and each customer type control, and their interaction with woman customer, with baseline customer type as the omitted categories. Column (1) includes state fixed effects, other email varying items: car year, subject and script number controls, and day of week and week fixed effects - indicating when the email was sent to a shop. Column (2) includes a non-independent shop control (non-independent shops: dealership and franchises). Column (3) adds CZ fixed effects, with a dummy for observations not linked to a CZ. Robust standard errors in parenthesis, with significance levels: *p < 0.10, **p < 0.05, ***p < 0.01.
Appendix

A Experimental design details

A.1 Randomization: Components and implementation steps

This project used a simple randomization method to assign an email account (perceived gender variation) and the email content (customer type variation) to each shop. The randomized component is visualized below.

First, I randomized gender and customer types. I randomly assigned one of the eighteen account IDs to each shop, matching the same number of emails to each account ID, and therefore, assigning half of the sample to women. Then, I randomly assigned each of the four customer type treatments within each gender. Thus, 18.3% of shops will be contacted by woman customers using standard scripts, 15.7% will be contacted by low-search woman customers, 8% by high-educated women and 8% by uninformed women. Equivalent shares are assigned to each man-type combination.

Then, I randomized additional components of each email; the car year, script design and email subject. Each of these variations is included to avoid sending a second identical email to each shop. The car years are 2009 and 2010, both belong to the same car generation. Within each customer type, there are two scripts; script 1 and script 2. The content is the same, but they are reordered and slightly differently worded. Across types, each script number mirrors one another. That is, the high-education type script 1 is the same as the standard script 1 except for the signature, which adds the Ph.D. title. All the subjects transmit the need for a radiator replacement, these are “change radiator,” “radiator replace” and “new radiator”.

Finally, the order and date in which each shop is contacted is also randomly assigned. The assignment of shops to treatments was done in weekly waves while sending a similar number of emails by day. Along the process, some shops were identified as not repairing Honda radiators, thereby dropped from the sample. For instance, collision shops were dropped as they do not
perform repairs unless caused by an accident.

Figure A.1: Randomization Steps

![Diagram of randomization steps]

Note: This figure shows each of the randomized components used to contact each facility. The red arrows indicate new reshuffles in the sample. The blue dotted arrows indicate the components that were randomly assigned.

A.2 Customer type scripts

This section provides both scripts used for each customer type. In each script, the terms ${FILL\_CITY}$, ${FILL\_YEAR}$ and ${FILL\_NAME}$ are included. At the time of sending the script, those terms are automatically replaced with the corresponding shop’s city, car year and user account name. As can be observed, script 1 is identical across type’s except for the information revealing content. Likewise, script 2 is identical across types. Script 1 and Script 2 within types aim to
convene the same information and structure, in a slightly reorganized and re-worded mail.

**Standard script 1**

*Hi,*

Do you replace radiators? Would it be possible to know what is the total estimate to replace mine and when I could take it? I live in ${FILL_CITY}, near your service center.

I have a 6 cylinder ${FILL_YEAR} automatic, sedan Honda Accord. It has been leaking and overheating. I already had some leaks fixed, now I’d like to replace the radiator.

Thank you!

Best,

${FILL_NAME}”

**Standard script 2**

*Hi,*

I need to replace my car’s radiator. It had cracks fixed before, but now it is overheating and leaking coolant again. I’d like to directly get a new radiator. The car is a ${FILL_YEAR} 6-cylinder, automatic, sedan Honda Accord.

Would you please tell me what is the total replacement cost? When can I stop by? I live close to your shop, here in ${FILL_CITY}.

Thanks!

Best,

${FILL_NAME}”
Low-Search script 1 and 1b

“Hi,

Do you replace radiators? Would it be possible to know what is the total estimate to replace mine and when I could take it? I live in ${FILL_CITY}, near your service center.

I have a 6 cylinder ${FILL_YEAR} automatic, sedan Honda Accord. It has been leaking and overheating. I already had some leaks fixed, now I’d like to replace the radiator.

“I am getting price estimates nearby”

or “I’ve checked Repairpal website for price references in the area. I will try to get a price estimate nearby.”

Thank you!

Best,

${FILL_NAME}”

Low-search script 2

“Hi,

I need to replace my car’s radiator. It had cracks fixed before, but now it is overheating and leaking coolant again. I’d like to directly get a new radiator. The car is a ${FILL_YEAR} 6-cylinder, automatic, sedan Honda Accord.

Would you please tell me what is the total replacement cost? When can I stop by? I live close to your shop, here in ${FILL_CITY}.

“I am getting price estimates nearby”

or “I’ve checked Repairpal website for price references in the area. I will try to get a price estimate nearby.”
Thanks!

Best,

${FILL_NAME}”

High-Education script 1

“Hi,

Do you replace radiators? Would it be possible to know what is the total estimate to replace mine and when I could take it? I live in ${FILL_CITY}, near your service center.

I have a 6 cylinder ${FILL_YEAR} automatic, sedan Honda Accord. It has been leaking and overheating. I already had some leaks fixed, now I’d like to replace the radiator.

Thank you!

Best,

${FILL_NAME}, Ph.D.”

Standard script 2

“Hi,

I need to replace my car’s radiator. It had cracks fixed before, but now it is overheating and leaking coolant again. I’d like to directly get a new radiator. The car is a ${FILL_YEAR} 6-cylinder, automatic, sedan Honda Accord.

Would you please tell me what is the total replacement cost? When can I stop by? I live close to your shop, here in ${FILL_CITY}.

Thanks!

Best,
Uninformed script 1

“Hi,

Do you replace radiators? Would you let me know what is the total cost to replace mine and when I could take my car? I live in ${FILL_CITY}, near your service center.

My car is leaving *these green liquid puddles* and the *temperature thing* is going up a lot. I had the radiator fixed before but I think it is time to replace it. The car is a Honda Accord. Not sure if it helps, it is an automatic, 4-door, "6-cyl." car from ${FILL_YEAR}.

Thank you!

Best,

${FILL_NAME}”

Uninformed script 2

“Hi,

My car is leaving *these green liquid puddles* and the *temperature thing* is rising too much. I had the radiator fixed before, now I think I should get a new one placed. The car is a Honda Accord. In case it is useful, it is an automatic, 4-door and "6-cyl." car from ${FILL_YEAR}.

Would you let me know what is the total cost estimate to replace my radiator? What would be a good time to take it? I live close to your shop, here in ${FILL_CITY}.

Thanks!

Best,

${FILL_NAME}”
A.3 Reply rules and template scripts

The following guidelines are used to reply to shops, regardless of the email account and customer type used to contact each shop.

The main two rules are to reply only once to shop’s follow-up emails, and second, to used an available reply template whenever possible. These templates are listed below and anticipate shop’s common requests, such as whether the customer has a phone or would be able to take the car for an inspection. Each request reply has two alternative templates, which are rotated on a weekly basis. If further car information is asked, the first template “car information” should be used, adding either the information asked. The car is 6-cylinder, Honda Accord, sub-model EX, 4 door, automatic car.

Another rule used was to try to reply to every email. So, I replied “thanks” in each of the following cases: when a shop mentions it is their policy not to give estimates without seeing the car, when they suggest contacting other person in the shop or, if they reply the following week of the experiment. This last rule is designed to avoid sending the script corresponding to the new week to a shop contacted in the previous week. Finally, I replied “thanks” when a shop provides a quote. If the shop mentions that the part should be ordered or offers to set an appointment, I replied “thanks, I will let you know” in order to avoid any confusion. There are some cases in which these shop emails remained unanswered, this should in no way affect this paper results, as no quote is given.

Finally, when an employee mentioned being out of office, I left this email unanswered, as I only wanted to recruit estimated from employees while they are working.

As will be noted in the reply scripts, they have the text ${name} and ${sender_name}. Each of these is manually replaced. ${name} refers to the employee’s name, and is highlighted in yellow so I minimized the chance of skipping this text replacement step. ${sender_name} is replaced by the account id being used. The email setting from each account had a different background theme, the name of the user in the gchat window, and I distributed accounts in each browser - allowing me to also familiarize which account is being used based on the explorer (chrome,
explorer and Firefox). Additionally, every account allowed the “undo” option after sending an email for 20 seconds.

**Reply scripts 1**

**Additional car information**

Hi ${name},

It’s a 3.5-liter car. Just in case, it also has AC.

Thank you!

${sender_name}

**VIN**

Hi ${name},

I am sorry, I won’t be close to my car until tomorrow. If it helps towards an approximate estimate, it is a 3.5-liter car, sedan, and with AC.

Thank you!

Best,

${sender_name}

**VIN + CALL**

Dear ${name},

I am sorry, I am not close to my car now. If it helps towards an approximate estimate, it is a 3.5-liter car, sedan, and with AC.

If it’s OK, I’d prefer to talk by over email.

Thank you!
Best regards,

${sender_name}

Call

Hi ${name},

If it’s OK, I’d prefer to talk over email. In case it is helpful, it is a 3.5-liter car, with air conditioner. In case it is the radiator, I would like to have a ballpark figure, if possible.

Thank you!

Best regards,

${sender_name}

Inspect

Hi ${name},

I realize the final estimate may change upon inspection, but I was hoping to have a sense of total repair costs before bringing it in (in case it is the radiator). If it is, I would not want to fix it but to have a new one installed.

Just in case, the engine is 3.5 liters and it is a car with air conditioner. Thank you, I appreciate all your help!

Best,

${sender_name}

Call + inspect

Hi ${name},
If it is OK, I’d like to talk over email before taking my car. I realize the final estimate may change upon inspection, but I was hoping to have a sense of total repair costs before bringing it in (in case it is the radiator). If it is, I would not want to fix it but to have a new one installed.

Just in case, the engine is 3.5 liters and it is a car with air conditioner. Thank you, I appreciate all your help!

Best,

${sender_name}$

**Info on service, no price**

Dear ${name},

I realize the final estimate may change upon inspection, but I was hoping to have a sense of total repair costs before bringing it in (in case it is the radiator). If it is, I would not want to fix it but to have a new one installed. Just in case, the engine is 3.5 liters and it is a car with air conditioner.

Thank you, I appreciate all your help!

Best,

${sender_name}$

**Are you bringing the radiator?**

Oh no, I would just bring the car in. In case it helps, the engine is 3.5 liters and it is a car with air conditioner.

Thank you!

${sender_name}$
Previous repairs

Hi ${name},

I had these same car issues a while ago. I had it repaired, not replaced. I think some cracks were closed, but I cannot say for sure. That is why, if the radiator is having problems I will just want a new one. That is why I am asking for an approximate estimate for this, to have an idea beforehand. In case it helps, the car has air conditioner and it’s 3.5 liters.

Thank you!

Best,

${sender_name}

Still need repair?

Yes, thank you. In case it helps, the car has air conditioner and it’s 3.5 liters. Just want to have a sense of approximate costs if it is the radiator that has issues (would like a new one, not repairing mine).

Best regards,

${sender_name}

Others – Radiator type:

I have no preference between radiator brands. I would just want something that I won’t have to replace again soon.

Thank you!

${sender_name}
**Reply Scripts 2**

**Additional car information**

Hi ${name},

Thanks for your reply! It is a 3.5-liter car, with AC.

Best,

${sender_name}

**VIN**

Dear ${name},

Thanks for your reply. Unfortunately, I cannot check the VIN until tomorrow night. Is there any other information I could give you now instead of waiting for the VIN number?

Just in case, the car has AC and it is 3.5 liters.

Best,

${sender_name}

**VIN + CALL**

Hi ${name},

Thanks for your reply. If you would not mind, I would rather correspond by email. Unfortunately, I cannot check the VIN until tomorrow night. Is there any other information I could give you now instead of waiting for the VIN number?

Just in case, the car has AC and it is 3.5 liters.

Best,

${sender_name}
Call

Hi ${name},

Thanks for your reply. If you would not mind, I would rather correspond by email. I’d appreciate having an approximate estimate if it is indeed the radiator.

Just in case, the car has AC, 3.5 liters.

Best,

${sender_name}

Inspect

Dear ${name},

If possible, could you let me know approximately how much it would cost? I just want to have a ballpark figure if it is indeed the radiator. Of course, I understand that it is an estimate before an inspection. But, it is it the radiator that has issues, I want to get a new one.

If it helps, the car has AC, 3.5 liters.

Thanks for your reply!

Best,

${sender_name}

Call + inspect

Dear ${name},

Thanks for your reply. If you would not mind, I would rather correspond by email first. If possible, could you let me know approximately how much it would cost? I just want to have a ballpark figure if it is indeed the radiator. Of course, I understand
that it is an estimate before an inspection. But, it is the radiator that has issues, I want to get a new one.

If it helps, the car has AC, 3.5 liters.

Best,

${sender_name}$

Info on service, no price

Hi ${name}$,

If possible, could you tell me a ballpark figure in case it is the radiator? I wanted to have an idea in case it is the radiator. I would like to get a new one and not fix the one I have. Of course, I understand that the final estimate may vary after an inspection. Let me know if there is any other information I could give you.

The car has AC, and it’s 3.5-liter.

Best,

${sender_name}$

Are you bringing the radiator?

No, I would just take my car. Just in case, it is a 3.5-liter car, with AC.

Best,

${sender_name}$

Previous repairs

Dear ${name}$,

Thanks for your reply. I am not 100% certain on what specific repair I had on the radiator. I know it had some cracks that were leaking, and they added something to
close them (but this may not be entirely accurate). I had this done a while ago but now I have the same issues. So, if it is indeed the radiator that has issues, I’d like to have an idea of how much it would cost to replace it instead of fixing it. I want to get a new one.

Just in case, it is a 3.5-liter car, with AC.

Best,

${sender_name}$

**Still need repair?**

Thank you for your reply. Yes, I still need to get my car fixed. Let me know if you could give me a ballpark figure in case it is the radiator. I don’t want it fixed but I’d like getting a new one.

Just in case, the car has AC, and it is 3.5-liter.

Best,

${sender_name}$

**Others – Radiator type:**

I’d want one that I won’t need to fix again in the near future, but I am indifferent between brands.

Thanks!

${sender_name}$

**A.4 Data collection rules**

The most relevant pre-specified rules used to translate email replies into a spreadsheet are the following: keep one observation for each contacted shop, the reported time and date of each
reply corresponds to the first reply received by each shop. Identify if a reply is sent before obtaining a quote from a shop, record the employee first name used to sign the email, and if missing, the name of the email account. Keep the first observed employee name when more than one employee reply to an email. If available, keep track of the price composition and which parts are explicitly not included (if any), the parts are labor, parts, radiator, coolant (includes and indicator for coolant flush services), hoses and clamps, thermostat, other (includes shop fees and other offers), taxes, quantity of estimated labor and estimated coolant. Also keep track of the radiator brand offered in the estimate (in particular, if it is a factory Honda part). Whenever more than one estimate is sent by the same shop, but in separated emails, keep the first estimate and add a comment to identify these cases.

A.5 Design validation

This section shows further validations of the experimental design implementation. First, I plot the geographic distribution of auto-repair shops assigned to be contacted by each treatment group. Figure A.2 shows the assignment of shops to men and women separately, and Figure A.3 further divides the assignment of shops by treatment group (combining each customer type and gender). As can be observed, both maps present a similar distribution of geographical locations contacted by each group.

In addition, Figure A.4 shows the distribution of invalid emails by treatment group, which reassuringly shows not significant differential across groups.
Figure A.2: Geographic distribution of contacted shops

(a) Men customers

(b) Women customers

Notes: These figures show the location of shops contacted by woman and man fictitious customers
Figure A.3: Distribution of shops assigned by gender and customer type
Figure A.4: Distribution of invalid-emails

Notes: This figure shows average invalid emails for women and men by their customer type used to contact each shop: Baseline, uninformed, quote and high-income. The plot shows the 95% confidence intervals, obtained with robust standard errors. The sample mean invalid emails is marked by a grey line within the plot.

B Theoretical Framework

This section provides a simple theoretical framework to help guide the empirical exercise in Section 3. I posit two dimensions of consumer heterogeneity to capture the main sources of variation in the empirical section.

On the one hand, consumers - indexed by $i$ - differ in their willingness to pay for the car repair service $\theta_i$. This parameter $\theta_i$ is assumed to be distributed according to $G^k$ where $k \in K$ is an observable characteristic. In the model, stores are allowed to discriminate based on observable $k$, which they use to infer $\theta_i$. In the experiment there are 8 observable characteristics (i.e., $\#K = 8$),
which differ by gender and customer type characteristics. While customers within group are heterogeneous, one may expect that some agents, such as those with high-income or high-search costs, to “typically” have a higher willingness to pay. In the model, I say that agent $k$ is “typically” more willing to pay than agent $k'$ if $G^k$ satisfies the monotone likelihood ratio property (MLRP) with respect to $G^{k'}$. This implies that the sensitivity of demand of agents $k$ relative to $k'$ is relatively larger when the price is relatively high (i.e., for $p > p'$, $\frac{g^k(p)}{g^k(p)} > \frac{g^{k'}(p')}{g^{k'}(p')}$. As I will show later, this implies that if $G^{females}$ satisfies the MLRP with respect to $G^{males}$, then there will be a positive gender gap.\(^{16}\) Note that, in principle, one can enrich the set of observable characteristics by interacting these with geographic characteristics. For example, if women living in area $A$ earn much less than men in area $A$, then one would expect the $G^k$s to be closer (or even reversed) to one another, leading to a lower (or opposite) price gap.

On the other hand, I assume that consumers differ in their ability to substitute the current shop with another potential shop. Each consumer $i$ is randomly assigned to an initial shop with equal probability and faces a switching cost $\lambda_i$ of shopping in an alternative store. This parameter $\lambda_i$ is assumed to be distributed according to $F^l$ where $l \in L$ is an observable characteristic. Again, stores in the model are allowed to discriminate based on the observable ($l$), which they use to infer $\lambda_i$. The counterpart of $\lambda_i$ in the experiment is the extent of competition in any given market, i.e., the ease with which customers may substitute stores. For example, if there are many stores within a 1 kilometer radius, then one may expect this $\lambda_i$ to be typically lower. In this sense, the correct interpretation is that the consumer lives near this store (or happens to “passively” shop in it as in Anderson and De Palma, 2005), and finds it costly to visit another store.\(^{17}\) Importantly, in the model $\lambda_i$ is not an informational cost. Indeed, in the equilibrium I describe below there is a pure strategies equilibrium so agents fully anticipate the distribution of prices at other stores in the Nash equilibrium.\(^{18}\) However, the credible threat of switching stores is key for the properties

\(^{16}\)First-order stochastic dominance is implied but not enough to guarantee this result.

\(^{17}\)The model presented here is a two-store model. I conjecture that having many stores will behave similarly to lowering switching costs in the context of the two-store model.

\(^{18}\)The interaction with informational frictions is an interesting avenue for research, but it is outside the scope of this paper. I decided to focus on the physical cost for two reasons. First of all, given the rise of the internet, the cost of sending emails to figure out quotes is likely to be small relative to the cost of actually going to the store. Second,
of the Nash equilibrium.

Consumer \( i \) utility function is given by

\[
    u_i = \begin{cases} 
        \theta_i - p_s & \text{if } i \text{ buys in the closest shop} \\
        \theta_i - p_s' - \lambda_i & \text{if } i \text{ buys in a different shop}
    \end{cases}
\]

There are 2 stores. Each store is assigned 1/2 “initial” consumers of each type \( \{k, l\} \). Stores may set price differently across types. For simplicity, I assume stores face a constant marginal cost, making the problem with respect to each type \( \{k, l\} \) separate from one another. Both shops have the same marginal cost, which I set without loss of generality to 0.

Consider the problem of shop \( j \) selling to a customer of type \( \{k, l\} \). The shop’s profits are given by

\[
    \pi^{kl}_j = p^{kl}_j \left( 1 - G^k(p^{kl}_j) \right) - 1_{p^{kl}_j > p^{kl}_{-j}} p^{kl}_j \left( F(p^{kl}_j - \min\{p^{kl}_{-j}\}) \right) + 1_{p^{kl}_j < p^{kl}_{-j}} \int_{0}^{p^{kl}_j} f^l(\lambda) (1 - G^k(p^{kl}_j + \lambda)) d\lambda
\]

When \( p^{kl}_j \to p^{kl}_{-j} \), one can show the first-order condition converges to \(^{19,20}\)

\[
    \frac{1 - G^k(p^{kl}_j)}{g^k(p^{kl}_j)} (1 - p^{kl}_j f^l(0)) - p^{kl}_j = 0. \tag{1}
\]

If consumers were immobile, i.e., if \( f^l(0) \to 0 \), then \( p^{kl}_j \) is the monopoly price. Each \( p^{kl}_j \) varies with customer observable characteristics, as stores engage in statistical discrimination to maximize profits. \(^{21}\) Proposition 1 shows that this is still the case when switching costs are nontrivial.

---

19. Despite the indicator functions, the objective is smooth at \( p_{-j} \) since you only lose/gain an infinitesimal number of consumers from an infinitesimal change in price. This smoothness precludes many of the standard problems of inexistence of pure strategy equilibria in this kind of models.

20. One naturally also needs to impose a condition on \( f \) to show this (local) solution to the first-order conditions is also a global maximum. I assume this is the case in the exposition.

21. There is no taste-based discrimination in the model. One could potentially incorporate this as a difference in the marginal cost of serving different kinds of customers. I will include this extension in a future version.
By contrast, if \( f^l(0) \to \infty \), \( p^{kl}_j \to 0 \). This is the perfect competition outcome: Stores cannot discriminate based on observable characteristics since they would lose all their customers to the other shop. Proposition 1 shows that this intuition is more general: the price increases monotonically with switching costs. One may then wonder if statistical discrimination and switching costs are “complements”, i.e., whether statistical discrimination is stronger when switching costs are larger. While the previous results imply this is obviously true in the \( f^l(0) \to 0 \) and \( f^l(0) \to \infty \) cases, a stronger result can be shown when \( G \) is uniform.

**Proposition 1.** In a symmetric Nash equilibrium, the following is true:

(a) Suppose \( G^k(\cdot) \) satisfies the monotone likelihood ratio property with respect to \( G^k(\cdot) \). Then, everything else equal, agents of type \( k \) are charged a higher price, i.e., \( p^{kl} > p^{kl'} \).

(b) Suppose agents of type \( l \) have a higher switching cost, i.e., that \( f^l(0) < f^l(0) \). Then, everything else equal, agents of type \( l \) are charged a higher price, i.e., \( p^{kl} > p^{kl'} \).

(c) Suppose \( G^k(x) = \frac{\bar{\theta}^k - x}{\bar{\theta}_k - \bar{\theta}^k} \), i.e., \( G^k \) is uniform. Then, agents with high willingness to pay are discriminated against more strongly when switching costs increase, i.e., if \( \bar{\theta}^k > \bar{\theta}^k \) and \( f^l(0) < f^l(0) \), then \( p^{kl} - p^{kl'} > p^{kl'} - p^{kl''} \).

**Proof.** (a) and (b) follows from a straightforward application of the implicit function theorem on 1. (c) follows from computing the cross-derivative using the results in (a) or (b).

Figure A.5 illustrates the results. I assume \( f = (\bar{\lambda}^l)^{-1} \) and vary \( \bar{\lambda}^l \) in the x axis. The solid line shows the equilibrium price where agents have a high willingness to pay (\( \bar{\theta}^k = 1 \)) while the dashed line corresponds to a case with a lower willingness to pay (\( \bar{\theta}^k = 1/2 \)). As switching costs increase, stores compete less with one another and are able to extract more surplus from agents with a high willingness to pay.
Figure A.5: Price Discrimination and Competition

C Replies and price estimates descriptive statistics

Figure A.6: Distribution of shops providing estimates
Figure A.7: Price differential against women by commuting zone

Notes: This map presents coefficient estimates from a regression of price estimates on state fixed effects interacted with a woman dummy, controlling for state fixed effects. Larger coefficients (blue) represent higher prices for women relative to men. Bins have equal counts of observations. In this sample, shops from 32 states quote higher estimates to women.

Figure A.8: Distribution of price estimates

Notes: This figures present the distribution of the 10,313 price estimates distribution, with estimates above 1,000 dollars included in the 1,000 dollar bin. The top panel plots the observed prices for men and the bottom for women, irrespective of customer types. The vertical grey line represents the mean estimate for each group.
D Price differences by treatment group: Robustness

Regression framework

This section re-does the main analysis in a simple regression framework. The OLS estimation is:

\[ p_{jk} = \alpha_0 + \beta_1 \text{Female}_k + \beta_2 \text{Type}_k + \beta_3 (\text{Type}_k \times \text{Female}_k) + \omega_s + \theta X_{jk} + \epsilon_{jk} \] (2)

where \( p_{jk} \) is the estimate that shop \( j \) provides to \( k \) customer (gender and type customer combination), \( \text{Female}_k \) is an indicator variable equal to 1 when a woman customer contacts shop \( j \), \( \text{Type}_k \) is a vector with indicator variables for each customer type. The omitted category is always the baseline type. And, \( \text{Type}_k \times \text{Female}_k \) is the vector of customer type interactions with the woman indicator. \( \omega_s \) are state fixed effects, \( X_{jk} \) correspond to email variation controls and timing variation controls. That is, day of the week and week of the field experiment in which shop \( j \) is contacted, and script number, car year and email subject controls. Additional specifications include a control for non-independent shops and commuting zone fixed effects. Finally, \( \epsilon_{jk} \) is the unobserved component, which has mean zero and is uncorrelated with our treatment groups.\(^{22}\)

Table A.4 shows these results. As pre-specified in the analysis plan, column (1) controls for state fixed effects, additional varying email items (car year, subject controls and script version used), and day of week and week fixed effects indicating when each shop is contacted. Column (2) adds a non-independent shop type control (i.e., franchises and dealerships) since their price levels are usually higher, and column (3) adds Commuting Zone (CZ) fixed effects and an additional indicator dummy for the shops that were not matched to any commuting zone. Figure A.9 complements these last results, directly plots adjusted point estimates for each treatment group with respect to men in the baseline group. Overall, the regression-adjusted results are consistent with raw price differences. Indeed, the gender-based premium in the baseline group remains significant across specifications and increases to an average of 9.6 dollars. The price decrease for women in the quote search group is marginally significant, while the price increase for men in

\(^{22}\)I do not cluster standard errors given this field experiment’s sampling design and experimental design Abadie et al. (2017)
the income group is statistically significant and varies between an average of 11.2 to 12.2 dollars with respect to men in the baseline group. Reassuringly, results are robust to the inclusion of controls, and none of the additional experimental design features is relevant to explain variation in prices.

Figure A.9: Effects of treatment groups on price

Notes: This figure shows average price gaps for each gender and customer type relative to male-baseline customers. The coefficient estimates are obtained from a specification with indicator variables for each customer type interacted with an indicator for female, an indicator for man customers, respectively. The omitted category is male-baseline customer, marked by a triangle with no confidence intervals in the plot. Additional controls include state and commuting zone fixed effects, with a dummy variable for observations not linked to a commuting zone, car year, subject and script number, day of the week and week fixed effects, and a non-independent shop indicator (i.e., franchise, dealership). The scatter shows each average price differential with their 95% confidence intervals. A null price gap estimate (y-axis= 0) is marked by a grey line within the plot.
Table A.4: Impact of gender and customer type on prices: Robustness

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Price estimates</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>9.637**</td>
<td>9.628**</td>
<td>9.837**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.235)</td>
<td>(4.137)</td>
<td>(4.202)</td>
<td></td>
</tr>
<tr>
<td>Uninformed</td>
<td>-1.955</td>
<td>-3.329</td>
<td>-1.875</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.445)</td>
<td>(5.288)</td>
<td>(5.377)</td>
<td></td>
</tr>
<tr>
<td>Uninformed \times Female</td>
<td>-5.559</td>
<td>-4.668</td>
<td>-5.659</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.598)</td>
<td>(7.351)</td>
<td>(7.518)</td>
<td></td>
</tr>
<tr>
<td>Quote</td>
<td>-1.019</td>
<td>0.025</td>
<td>0.472</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.557)</td>
<td>(4.482)</td>
<td>(4.570)</td>
<td></td>
</tr>
<tr>
<td>Quote \times Female</td>
<td>-9.591</td>
<td>-11.217*</td>
<td>-10.893*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.397)</td>
<td>(6.241)</td>
<td>(6.354)</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>12.229**</td>
<td>11.718**</td>
<td>13.471**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.298)</td>
<td>(5.200)</td>
<td>(5.379)</td>
<td></td>
</tr>
<tr>
<td>Income \times Female</td>
<td>-13.237*</td>
<td>-12.552*</td>
<td>-13.754*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.561)</td>
<td>(7.372)</td>
<td>(7.612)</td>
<td></td>
</tr>
<tr>
<td>Subject: Change radiator</td>
<td>-3.170</td>
<td>-3.061</td>
<td>-2.273</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.206)</td>
<td>(3.127)</td>
<td>(3.203)</td>
<td></td>
</tr>
<tr>
<td>Subject: Radiator replace</td>
<td>-2.465</td>
<td>-0.832</td>
<td>-0.430</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.302)</td>
<td>(3.222)</td>
<td>(3.281)</td>
<td></td>
</tr>
<tr>
<td>Car year: 2010</td>
<td>1.143</td>
<td>1.111</td>
<td>1.246</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.590)</td>
<td>(2.523)</td>
<td>(2.566)</td>
<td></td>
</tr>
<tr>
<td>Script 1</td>
<td>-3.434</td>
<td>-2.640</td>
<td>-2.802</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.608)</td>
<td>(2.537)</td>
<td>(2.585)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>State FE</td>
<td>+ Shop Types</td>
<td>+ CZ FE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mail items</td>
<td>DOW, Week FE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.031</td>
<td>0.082</td>
<td>0.103</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports coefficients from a regression of total price estimates on an indicator for woman customer, and indicator for each customer type, and their interaction with woman customer, with baseline customer type as the omitted categories. Column (1) includes state fixed effects, other email varying items: car year, subject and script number controls, and day of week and week fixed effects - indicating when the email was sent to a shop. Column (2) includes a non-independent shop control (non-independent shops: dealership and franchises). Column (3) adds CZ fixed effects, with a dummy for observations not linked to a CZ. Robust standard errors in parenthesis, with significance levels: *p < 0.10, **p < 0.05, ***p < 0.01.
Selection bias correction

Table A.5: Heckman two-step correction: Robustness

<table>
<thead>
<tr>
<th>Comparison group</th>
<th>Give price estimate</th>
<th>Price estimate</th>
<th>Inverse Mills Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women vs. men (baseline)</td>
<td>0.009*</td>
<td>8.854**</td>
<td>36.8*</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(4.294)</td>
<td>(20.54)</td>
</tr>
<tr>
<td>Uninformed vs. Baseline (women)</td>
<td>-0.012*</td>
<td>-7.124</td>
<td>51.387*</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(5.494)</td>
<td>(25.728)</td>
</tr>
<tr>
<td>Quote vs. Baseline (all)</td>
<td>-0.015***</td>
<td>-6.083*</td>
<td>10.539</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(3.235)</td>
<td>(18.034)</td>
</tr>
</tbody>
</table>

Notes: This table reports the difference in the rate of shops providing quotes (Column 1), the corrected price estimate differential using Heckman’s two-step correction method (Column 2), and the associated inverse Mills ratio coefficient (Column 3), for each comparison group that showed a significant difference in the rate of price estimates provided. The coefficient obtained in Column (2) controls for state fixed effects, email varying items, a dummy for non-independent shops, and day of week and week fixed effects. Standard errors in parenthesis, with significance levels: *p < 0.10, **p < 0.05, ***p < 0.01.
### Table A.6: Price gaps with number of competitors within a 1-km radius: Robustness

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitors</td>
<td>-0.669</td>
<td>-0.600</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td>(0.331)</td>
<td>(0.488)</td>
</tr>
<tr>
<td>Competitors x Female</td>
<td>-1.501***</td>
<td>-1.360***</td>
<td>-2.134***</td>
</tr>
<tr>
<td></td>
<td>(0.459)</td>
<td>(0.442)</td>
<td>(0.670)</td>
</tr>
<tr>
<td>Av. Competitors</td>
<td>3.92</td>
<td>3.92</td>
<td>2.59</td>
</tr>
<tr>
<td>Observations</td>
<td>10,313</td>
<td>10,302</td>
<td>10,313</td>
</tr>
</tbody>
</table>

**Notes:** This table reports coefficients from a regression of total price estimates on market concentration measures and their interaction with woman customers. Each column used an alternative definition to measure competitors within a 1-km radius. Column (1) repeats the benchmark specification, with competitors defined as all mechanic shops within a 1-km radius. Column (2) uses the same measure and performs a robust regression - which weights observations differently based on how well behaved these are. Column (3) counts nearby competitors that have a listed website or email address available in the YP. All specifications control for woman indicator, a woman indicator, and each customer type control, and their interaction with woman customer, with baseline customer type as the omitted category, state fixed effects, other email varying items: car year, subject and script number controls, and day of week and week fixed effects - indicating when the email was sent to a shop, non-independent shop control, whether a non-independent shop is within the market concentration measure, CZ fixed effects, with a dummy for observations not linked to a CZ. Robust standard errors in parenthesis, with significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 
Table A.7: Price gaps with alternative distance radius: Robustness

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitors</td>
<td>-0.669**</td>
<td>-0.614***</td>
<td>-0.725***</td>
<td>-0.691***</td>
<td>-0.644***</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td>(0.233)</td>
<td>(0.180)</td>
<td>(0.146)</td>
<td>(0.117)</td>
<td>(0.398)</td>
</tr>
<tr>
<td>Competitors × Female</td>
<td>-1.501***</td>
<td>-0.971***</td>
<td>-0.667***</td>
<td>-0.425**</td>
<td>-0.305**</td>
<td>-1.388**</td>
</tr>
<tr>
<td></td>
<td>(0.459)</td>
<td>(0.304)</td>
<td>(0.223)</td>
<td>(0.175)</td>
<td>(0.141)</td>
<td>(0.573)</td>
</tr>
<tr>
<td>Av competitors</td>
<td>3.92</td>
<td>6.45</td>
<td>9.40</td>
<td>12.54</td>
<td>15.95</td>
<td>3.92</td>
</tr>
<tr>
<td>Observations</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
<td>10,313</td>
</tr>
</tbody>
</table>

Notes: This table reports coefficients from a regression of total price estimates on number of competitors and its interaction with a dummy for female. Each column defines the number of competitors within radius thresholds increasing by 0.5 km, starting with a 1 km radius in column (1), and ending with a 3-km radius in column (2). Column (1) defines competitors as the count within a 1-km radius, and the specification includes controls for competitors in the 1.5, 2, 2.5 and 3 km outer rings. All specifications control for woman indicator, a woman indicator, and each customer type control, and their interaction with woman customer, with baseline customer type as the omitted categories, state fixed effects, other email varying items: car year, subject and script number controls, and day of week and week fixed effects - indicating when the email was sent to a shop, non-independent shop control, whether a non-independent shop is within the market concentration measure, CZ fixed effects, with a dummy for observations not linked to a CZ. Robust standard errors in parenthesis, with significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01.
### E Are effects driven by other service offers?

Table A.8: Effect of additional part and service offers

<table>
<thead>
<tr>
<th></th>
<th>Give detail (1)</th>
<th>Add Parts (2)</th>
<th>Fees (3)</th>
<th>Warranty (4)</th>
<th>Offers (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.025*</td>
<td>-0.000</td>
<td>0.008</td>
<td>0.009</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.016)</td>
<td>(0.009)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Uninformed</td>
<td>-0.052***</td>
<td>-0.014</td>
<td>-0.029</td>
<td>-0.007</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.015)</td>
<td>(0.020)</td>
<td>(0.011)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Uninformed × Female</td>
<td>0.041</td>
<td>0.009</td>
<td>0.002</td>
<td>-0.007</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.021)</td>
<td>(0.028)</td>
<td>(0.016)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Quote</td>
<td>0.004</td>
<td>-0.005</td>
<td>0.017</td>
<td>0.021**</td>
<td>0.023**</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.017)</td>
<td>(0.010)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Quote × Female</td>
<td>-0.018</td>
<td>0.000</td>
<td>-0.004</td>
<td>-0.010</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.017)</td>
<td>(0.024)</td>
<td>(0.015)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Income</td>
<td>0.031*</td>
<td>-0.002</td>
<td>0.048**</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.015)</td>
<td>(0.020)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Income × Female</td>
<td>-0.023</td>
<td>0.014</td>
<td>-0.029</td>
<td>-0.004</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.021)</td>
<td>(0.029)</td>
<td>(0.017)</td>
<td>(0.019)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Av Mean</th>
<th>Observations</th>
<th>Adjusted R2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.79</td>
<td>10,313</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>10,313</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>10,313</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>10,313</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>10,313</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Notes:** This table reports coefficients from a regression of give price detail components (Column 1), include additional car parts in estimate: radiator cap, radiator hoses and thermostat (Column 2) Include tax and fees in estimate (Column 3), mention service has a warranty (Column 4), and additional offers: discount, a free inspection, price-matching, and others such as financing options or a shuttle (Column 3) on an indicator for woman customer, and indicator for customer types (excluding baseline type) and female, and indicators for each customer type, with baseline customer type as the omitted categories, commuting zone and state fixed effects, email varying items: car year, subject and script number controls, and adds day of week and week fixed effects, and a dummy for non-independent shop (dealership and franchises). Every variable takes a value equal to 0 when the shop does not mention its inclusion. Robust standard errors in parenthesis, with significance levels: *p < 0.10, **p < 0.05, ***p < 0.01.