GENDER PRICE GAPS AND COMPETITION: EVIDENCE FROM A CORRESPONDENCE STUDY

2023

BANCO DE ESPAÑA
Eurosistema

Documentos de Trabajo
N.º 2333

Margarita Machelett

GENDER PRICE GAPS AND COMPETITION:
EVIDENCE FROM A CORRESPONDENCE STUDY

GENDER PRICE GAPS AND COMPETITION: EVIDENCE FROM A CORRESPONDENCE STUDY [*]

Margarita Machelett [**]

BANCO DE ESPAÑA

(*) I am very grateful for helpful comments and feedback from Emily Oster, John Friedman, Brian Knight, Anna Aizer, Vivek Bhattacharya, Pedro Dal Bó, Sebastián Fanelli, Bobby Pakzad-Hurson, Dan-Olof Rooth, Jesse Shapiro, Neil Thakral, Matt Turner, Lorenzo Adelco, Samuele Giambra, Florian Gunsilius, Ryan Kessler, Samsun Knight, Ted Lawless, Stefano Pollini and Catalina Ravizza.
(**) Email: margarita_machelett@alumni.brown.edu

Documentos de Trabajo. N.º 2333

November 2023

https://doi.org/10.53479/34747

ISSN: 1579-8666 (on line)

Abstract

This paper describes a large-scale field experiment conducted 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 2% (over 10 dollars) higher than those received by men. These differences disappear when women signal low search costs, suggesting statistical rather than taste-based discrimination. Price requests that appear to come from high-income households raise quotes for men but not women, eliminating the gender gap. The price gap also falls with the number of nearby repair shops, suggesting that market competition alleviates gender-based price discrimination.

Keywords: competition, discrimination, field experiment, gender.

JEL classification: C93, D4, J16, J18, J71.

Resumen

En este trabajo se lleva a cabo un experimento a gran escala para evaluar causalmente si existe discriminación de género en los precios de las reparaciones de automóviles en Estados Unidos. Las mujeres reciben presupuestos un 2 por ciento (10 dólares) mayores que los que reciben los hombres. Esta diferencia desaparece cuando las mujeres señalan tener bajos costes de búsqueda, lo que sugiere que el tipo de discriminación es estadística y no basada en preferencias. En cambio, cuando las peticiones de presupuesto proceden de clientes que señalan rentas altas, el precio aumenta en el caso de los hombres pero no en el de las mujeres, eliminando dicha diferencia. También se observa una reducción en la discriminación de precios en aquellas áreas donde los talleres de reparación tienen más competidores cercanos.

Palabras clave: competencia, discriminación, experimento de campo, género.

Códigos JEL: C93, D4, J16, J18, J71.

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. From an economic point of view, gender discrimination can manifest itself in many forms, affecting women both as workers and consumers.

Despite the existence of many studies on gender-based pay inequality, considerably less attention has been devoted to inequities in product and service markets. These can, however, have serious consequences for welfare and equity of women. For example, measures of consumer prices in the US typically reveal that products for women are more expensive than similar products for men. A study by the New York City Department of Consumer Affairs reports that women products cost on average 7% more than similar products for men (Bessendorf, 2015). A prior study by the State of California further claims that women paid an annual "gender tax" of approximately $1,351 for the same goods and services as men or about a $15 billion for all women in California[1].

Both the existence of gender-based discrimination and the motives behind it, however, are in fact 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 these difficulties by conducting one of the first large-scale field experiments building on the correspondence testing methodology (Bertrand and Mullainathan 2004, Jowell and Prescott-Clarke, 1970 and Riach and Rich 1991). I sent more than 57,400 emails to shops requesting an estimate for the same standard repair on the US car repair market to test both the extent and patterns of gender-based price discrimination.

[1] 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.

The experiment randomized two components of the email content observed to shops. To measure gender-price gaps, I varied the perceived gender of the consumer by using distinctively male and female names. To help distinguish between taste-based Becker (1957) and statistical discrimination (Arrow, 1973, and Phelps 1972), I also varied other customer attributes observed to shops, such as having low search costs, high implicit income or no car repair knowledge. These customer type signals sought, in particular, to shift shop's priors on customers willingness to pay. In a taste-based discrimination framework, women would always pay a premium to compensate for sellers' distaste from serving them, whereas price gaps would disappear with additional information under a statistical discrimination framework.

Overall, the experiment resulted in a a response rate of close to one quote for every five emails sent. The responses show clear evidence of gender-based price discrimination in the "baseline" group. Indeed, women in this group receive price estimates of roughly 487 dollars - 10.4 dollars more than men. Reassuringly, this statistically significant difference is robust to the inclusion of additional controls, such as geographic shop characteristics and elements of the field experiment design.

This gender price gap is not driven by differential response rates or changes in the composition of the provided quotes. There is no evidence of gender gaps in non-monetary outcomes. For instance, shops do not include additional parts (on top of those essential for the repair) depending on customer's gender. There is also no evidence of differential treatment in terms of discounts, price matching, free inspections, and other such offers, which is typically expected of credence good markets (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 under-treatment as well as overcharging (Darby and Karni, 1973, and Dulleck, Kerschbamer and Sutter, 2011).

Importantly, the theoretical relationship between competition and price discrimination is ambiguous, making this an empirical question. In this paper, I rely on the detailed cross-sectional nature of the data - where I gathered all car repair listings in the US - to construct measures of competition based on the number of other car repair shops that are

near each shop. I relate these measures to my experimental results and find a significant negative association between competition levels and gender price gaps. Car shops without nearby competitors give quotes to women roughly 18 dollars more than men, and this gap is on average 1.6 dollars lower for every additional competitor within a 1-kilometer radius. It is worth noting that Becker's taste-based model predicts that biased sellers would be competed away in a competitive market. As I show in a simple framework, this finding can also be rationalized in a statistical discrimination framework in which sellers are allowed to price discriminate based on customer characteristics but also with the ease in which customers can switch between shops. Consistent with this result, I find that shops with presumably more market power (franchises and dealerships) tend to increase gender price gaps relatively more than independent shops.

In contrast to the observed gender-price gap in the *baseline* group, customers with additional attributes receive price estimates that do not vary significantly by perceived gender - although price shifts vary by group. Women get lower estimates once they signal low search costs (*search* group customers), while men estimates remain the same, effectively closing the gender price gap. By contrast, signaling high income does hurt men as they receive higher price estimates, while female estimates do not increase. These results suggest 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 their implicit 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).

This paper contributes to the literature on discrimination in consumer markets.[2] It is related to a small but growing literature that focuses on discrimination in consumer transactions involving monetary outcomes, beyond exploring differentials in call-back rates. This paper allows the detection of discrimination beyond differences in reply rates to measuring the magnitude of discrimination's effects on prices. In this case, measuring how much of the differences in provided estimates between men and women is due to discrimination with new and robust evidence in favor of the hypothesis that price discrimination

[2]See Bertrand and Duflo (2017), Neumark (2018), Riach and Rich (2002), and Guryan and Charles (2013) for reviews of studies on discrimination.

is both extant and economically meaningful. It extends the findings of studies that show race- and gender-based discrimination in auto and credit markets (e.g., Ayres and Siegelman 1995, Burke and Sass 2011, Busse et al. 2017 Pope and Sydnor 2011 and Ravina 2007). In contrast to these prior studies, the observed price gap is lower, but still provides contrary evidence to non-experimental recent studies that have not found evidence of discrimination in online markets (Morton, Zettelmeyer and Silva-Risso, 2003).

This paper also contributes to the literature studying the mechanisms of discrimination. In line with the findings of Castillo, Petrie, Torero and Vesterlund (2013), Gneezy, List and Price (2012), and Busse, Israeli and Zettelmeyer (2017), price discrimination meaningfully changes in response to changes in the perceived customer attributes. Castillo, Petrie, Torero and Vesterlund (2013) send six paired men and women testers to negotiate taxi fares in Lima, Peru. Differences disappear once the testers reject the first taxi fare and proceed to the second taxi in line. Busse, Israeli and Zettelmeyer (2017) request price quotes for a radiator replacement from repair shops in heavily populated areas with four men and five women callers. They find that price gaps disappear once the average US repair price is referenced. Gneezy, List and Price (2012) send six disabled and six non-disabled testers to request price quotes at 36 auto body shops in Chicago. The differentials disappear once individuals mention they are getting other quotes. Given the scale of this experiment, 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. From a methodological point of view, the correspondence study approach also ensures more robust comparability across treatment groups , guaranteeing that any observed differences are caused solely by the gender and customer type trait manipulation.[3] This methodology also allowed contacting shops across the US, expanding on the previously more localized studies.

This paper also contributes to the empirical literature on price dispersion and competition, given the scale and rich geographical variation of the experiment. In contrast to a number of prior observational 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

[3]Many of the weaknesses of audit studies have been discussed in Heckman (1998) and Siegelman and Heckman (1993).

1991)this paper joins the studies finding that increases in competition decrease price dispersion (Barron, Taylor and Umbeck 2004, Gerardi and Shapiro 2009). Importantly, this paper setting differ from the previous literature by relating competition to experimental results, and by allowing sellers to directly price discriminate between observable consumers characteristics, rather than having consumers self-select into different products.[4]

Finally, this paper relates to the literature on credence good markets, in particular to Schneider (2012), who conducts visits to car repair shops. He finds evidence of poor diagnosis, which does not vary when posing as a one-time versus a repeat-business customer.[5] My analysis confirms these observations, in that recommended repairs do not vary with customer characteristics.

The remainder of the paper proceeds as follows. The next section gives an overview of the experimental design, implementation and data collection. Section 3 describes the experiment data and discusses its main concerns. Section 4 presents the empirical results. Section 5 concludes.

2 Experimental design

Building on the correspondence testing method, I conducted a large-scale field experiment in the US car repair market to test if there is a gender-based price gap and how it varies with additional information about customer and market characteristics.[6]

This approach has several advantages with respect to prior studies using audit and phone call testing methods. It ensures more robust comparability across treatment groups by eliminating differences in individual characteristics observed to shops but not to researchers. It also allows for a larger sample size due to its low marginal cost. Conducting this experiment at such large scale provides more precise estimates and allows for the combination

[4]Doleac and Stein (2013) and Nunley, Owens and Howard (2011) field experiments in online auctions examine how consumers discriminate based on sellers perceived race. Nunley et al. (2011) define competition by space display of shops products, Doleac and Stein (2013) by the bids consumers make. To the best of my knowledge, these are the only other studies relating experimental results to competition, but defining competition by demand measures.

[5]See Kerschbamer and Sutter (2017) for a survey of recent laboratory and field experiments on creedence goods

[6]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.

of experimental price results with larger variation in market characteristics, in particular with varying the degrees of market competition.

As detailed in this section, I constructed a comprehensive database of car repair businesses in the US. Then, between July and August 2018, I contacted approximately 57,400 car repair shops using customized scripts. Close to half sent any reply, and close to one out of five shops provided a valid quote (10,323).

In the scripted emails, I requested a quote for a radiator replacement for the same car. This ensures that the quote provided is always equally costly, regardless of customer gender. Additionally, I randomized the informational content of the emails while perfectly controlling all information about a customer observed to each shop. I randomized the gender of the customer, and I also varied additional information about the customers to signal low search costs, knowledge about car repairs and implicit income. These variations added to a standard (*baseline* type) script that customer is searching for other estimates (*search* type), a customer is not knowledgeable on car repairs (*uninformed* type), or a customer who is highly educated to proxy for income (*income* type). The signals intended to shift shops priors on customers willingness to pay and test whether this information becomes more relevant than the gender.

The remainder of the section reviews the experimental design details.

2.1 Auto repair shops

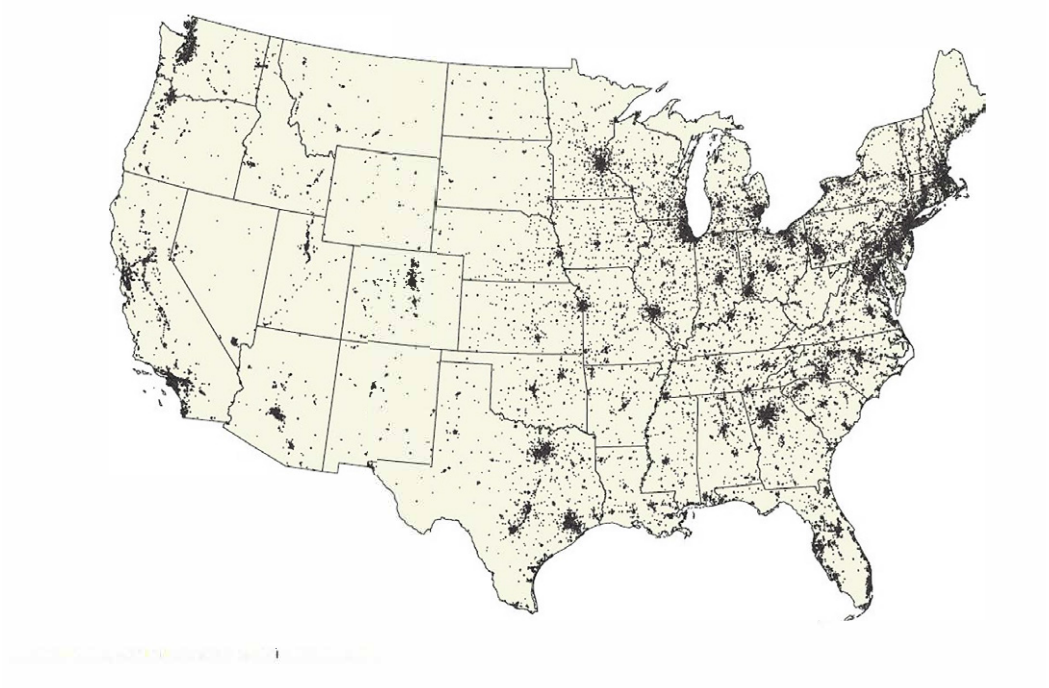
Car repair services provide an ideal setting to study price discrimination by gender as this market make it feasible (Stole, 2007). That is, this is a typical credence good market. There can be a suspicion that shops overcharge customers and can potentially treat customers differently. Thereby, implying that a shop might have market power and the ability to segment customers. Furthermore, it is very difficult to resell a service once it is provided, leaving arbitrage opportunities minimized.

The data were collected for all US auto repair shops listed in the online Yellow Pages (YP) (`www.yellowpages.com`). The YP offers a comprehensive listing of businesses nationwide and is used by nearly 60 million customers in the US each month, with auto repair shops

being one of the popular categories. Each listing provides information on the shop's name, location, website, and type of services provided.

The final sample reaches approximately 58,300 auto repair businesses, after sample restrictions were applied (e.g., excluding body paint shops). Figure 1 shows the location of the shops, which is dispersed across the US, with more facilities located in more popular areas.[7]

Figure 1: Distribution of shops



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

The information retrieved from the YP allows me to construct other variables of interest. First, I classified each repair shop as independent and non-independent (franchises and dealerships).[8]. This allows me 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 easier to reach the mechanic in charge of a repair directly.

Second, I built market competition measures using the YP geographic location of each shop and GPS software. With Google Maps geocoding API, I converted addresses into

[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.

[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 correspond to closed businesses that are still listed in the YP.

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. Finally, geographic shop locations are used to match each shop to their neighborhood.

2.2 Email design: Randomized components

As the second step of this experiment, I used this comprehensive database of car repair businesses to send emails inquiring about the cost of a standard car repair of one of US most popular cars (replacing a Honda Accord radiator). I built a baseline email script and randomized its informational content along two crucial dimensions. I varied the perceived customer's gender by using names that are distinctively recognized as female or male. I also varied the signals about customer *types,* distinguishing between customers with no car knowledge, searching for other quotes, or with implicit high income.

Using these additional variations to the email content can help distinguish the motives for discrimination. Following Becker (1957) taste-based discrimination framework, if shops have a distaste towards women customers, they would compensate for this disutility by providing women with higher quotes for the same type of repair. When the fraction of shops discriminating against women is large enough, a gender-based price differential would exist reflecting this distaste parameter. Importantly, a gendered price gap would be observed across all types of customers. In contrast, imperfect information could lead to gendered price gaps in a statistical discrimination framework (Aigner and Cain 1977, Arrow, 1973, and Phelps 1972). Discrimination would be the result of a signal extraction problem, where -absent of additional information- shops use customer's gender to infer how profitable it could be to sell to them. But, once shops have other valuable customer-specific information (such as the customer *type*), they may rely less on gender and more on these additional signals to provide quotes. In this second framework, for instance, revealing to have low search costs could help customers get more competitive prices. If shops prior were that women have higher search costs in this industry, then this signal would be more helpful for women than men. Under statistical discrimination, gendered price gaps would not be observed across all customer types.

2.2.1 Customer gender variation

The first variation I introduced refers to the perceived customer gender. I used eighteen distinctively female and male namesand created their respective email accounts. These names are observed by sellers in the email address, email account name, and email content signature.

The eighteen names are drawn from Bertrand and Mullainathan (2004) and Levitt and Dubner (2006), all of which are popular and are associated with a white-sounding name. I also paired these names with popular white-sounding last names from Bertrand and Mullainathan (2004).[9] Choosing first *and* last names that are predominantly associated with white individuals allows me to use these uniquely as signals of customer gender, ruling out the possibility that changes in price estimates are associated with changes in race.

2.2.2 Customer types variation

In addition to signaling customers gender, I added variations of customer attributes to a ***baseline*** email content (i.e., script). These variations to a ***baseline*** customer type are either a *search*, *income* or an *uninformed* type. They aim at shifting each seller's prior of consumers willingness to pay and are used to compare how gender price gaps vary, with the baseline type as the comparison group as specified in the Pre-Analysis-Plan (RCT ID: AEARCTR-0003279). While scripts vary the conveyed information on knowledge, income and quote search efforts, they are otherwise identical in structure and content.

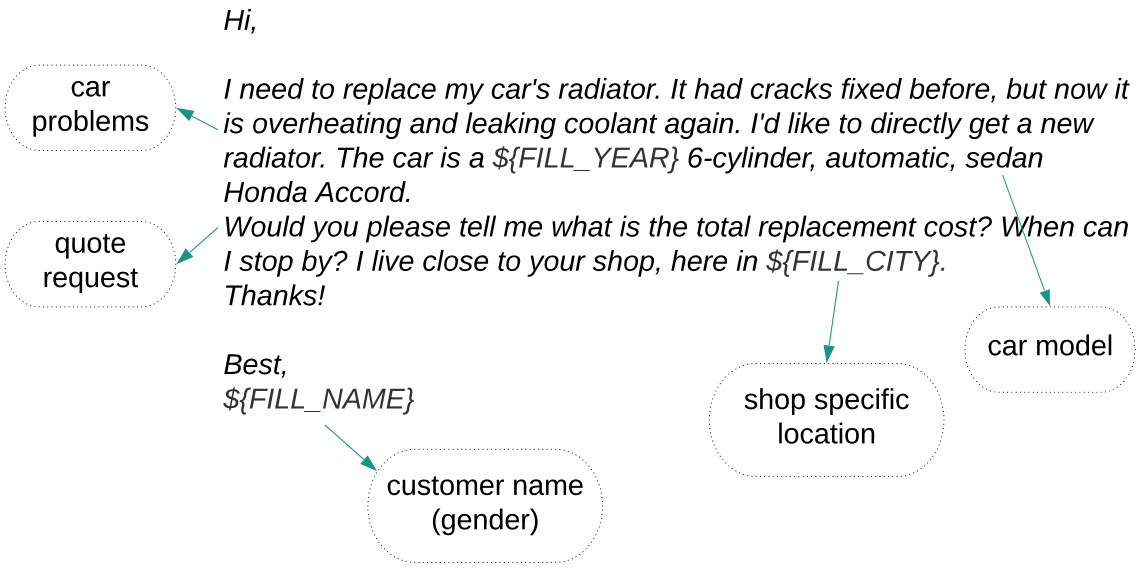
Figure 2 presents a template script used with the *baseline* customer. This customer script mentions a car model, radiator problems, asks for a total cost estimate for its replacement, and also mentions living close to the shop. The car is one of the most popular US cars - a Honda Accord. The problems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

[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 smaller 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 additional white-sounding names from Levitt and Dubner (2006). The selected names are consistent with Bertrand and Mullainathan (2004) study. That is, in addition to using popular female and male names in the US, each of these additional names is also among the 100 most popular baby boy and girls names as in Massachusetts between 1970 and 1986 (based on the Social Security Administration database on birth counts by year, state, name and gender). The last names were randomly matched to the first names.

Commission and RepairPal website, among others. The email also includes a reference to the residence near the shop to increase the apparent veracity of the email.

For each of the three additional customer types, I modified the baseline email script as follows. First, for the *search* type customers, I added a sentence where I state to be searching for other price estimates from nearby shops. Thus, with this quote type customer, I aim to signal that customers have low-search costs, making sure shops know with certainty that customers are exerting some effort to obtain more competitive estimates. Then, for the *income* type customers, I add the title "Ph.D." to the email signature. This high education signal intends to proxy for high income, signaling shops that this customer has a higher opportunity cost of time. Lastly, for *uninformed* type customers, I described the radiator problems using vague, non-technical terms - for example, I mentioned that the car "leaves green puddles" instead of "leaking coolant," and that the "temperature thing" is raising

Figure 2: Baseline customer template script



Note: This figure shows the script template used for a baseline-type customer. Text between "${}" is filled with specific randomized component.

instead of the "gauge." With this language I aim to signal that the customer has no car knowledge, giving more discretion to the seller to increasing prices or offering unnecessary additional repairs.

The script templates were pre-tested in a pilot carried out in June 2018, and all templates are available in Appendix 6.B.

2.2.3 Other email items variation

Finally, in addition to varying customer's gender and type, varied other email features used to contact each shop such as the email subject. The purpose of changing these additional email features is to avoid sending an identical email to shops while not affecting shop price estimates. All price estimate differences should arise due to the randomly assigned gender and customer type used to contact each shop. To ensure this, I include these additional email features as controls in regression specifications as specified in the Pre Analysis Plan. As expected, the discrimination pattern results are robust to their inclusion. Moreover, the coefficients of these control variables are never statistically significant.

2.3 Randomization

The randomization assignment of email scripts to car repair shops was done unconditionally on any further characteristics (e.g., geographic or shop characteristics). As a first step, half of the shops were assigned to be contacted by a woman customer and the other half by a man, using one of the created email accounts. Next, I randomly assigned one of the four customer types to each shop within each gender. I used optimal shares obtained from power calculations. These power calculations were obtained using pilot observations and weighted more the baseline and quote types, resulting in a distribution of shops to be contacted by 18 percent baseline, 16 percent quote, 8 percent uninformed, and 8 percent income scripts of each gender. Lastly, I reshuffled the sample and assigned one of each remaining email items (car year, script version and the subject of the email). Appendix 6.A details this randomization procedure.

2.4 Follow-up of email replies

Besides following a protocol to send emails, I also followed pre-specified email guidelines to follow up the shop's responses. I drafted scripted replies in the pilot in anticipation to shop's expected inquiries, and allowed for one reply to each shop when they did not provide a price estimate in their first reply. The guidelines aim to guarantee that all treatment groups behave identically, and thus, observed price estimates only vary with information

on customer's gender and types. Appendix 6.C details these rules and includes the reply template scripts.

2.5 Implementation and data collection

In this subsection, I describe the implementation and data collection used to guarantee comparability across groups.

I used automated tools to contact each shop via email during the mornings. Overall, I sent approximately 57,400 emails. This resulted in a response rate of one half and a valid price estimate in close to one out of five emails sent, yielding 10,323 price estimates.

Once the interaction with the shops was completed, I process the information from each reply and then match it to the shop data, which contains each treatment group and email details assigned to a shop as well as shop characteristics.

The information collected can be grouped into three main categories: reply type and employee first name, the salience of experimental design features, and price and service information. The first category records if, when and which type of answer is received. The type of answer includes a classification of each email content according to whether the email was invalid, the shop does not perform the service, the shop provides the service and gives a quote, and the shop provides the service but does not provide a quote. This category also tracks how much effort customers exert to receive a quote, measured by an indicator equal to one if a second email (reply) is sent to a shop before receiving a quote. Employee first name is recorded to assign them their likely gender, of which 80 percent of the employees have an assigned gender and, 14 percent correspond to woman first names. The gender is assigned based on the SSA National Data on the 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 70 percent of the occurrences, then that name is classified as corresponding to a woman.

The salience of experimental design features measures if shops replies mention relevant content from the original email. This content intends to capture items from the experimental design, thus I constructed indicator variables for mentioning the customer's name, gender (e.g., Mr, Miss), and profession in the case of income type customers. I also created

an indicator variable to record when the shop's reply makes a reference to RepairPal. I restricted this sample to replies from shops that perform the requested radiator replacements, as these integrate the sample of interest.

Finally, the pricing information collected includes price and service details, as well as other offers customers receive. The total price estimate is the primary outcome of interest and was processed using pre-specified rules. I defined total price estimates as follows: the total discounted price provided in a quote, the average price whenever price ranges are provided, the first price provided by a shop whenever more the shop provided more than one quote in separate emails. I also kept the estimate provided by a shop regardless of additional price-match offers. I exclude estimates defined as invalid if shops explicitly excluded either the price for the labor or the radiator, prices were extremely low or high (below 100 dollars or above 2,000 dollars), or estimates had a price range ratio over twice the price range and no further explanation about price components was given. These thresholds are arbitrary but conservative values. Overall, 121 price estimates are dropped from the sample as they are considered invalid. In addition, I recorded price component details whenever available (i.e., the estimate includes a thermostat replacement, or additional radiator hoses).

Regarding the service details, I recorded shop offers such as price-match of competitor estimates, mentions of a warranty on their service, discounts, or any additional offers such as a car loaner, shuttle service, and financing options. These service details are collected to test whether shops change their behavior with each treatment group.

3 Final sample and design validation

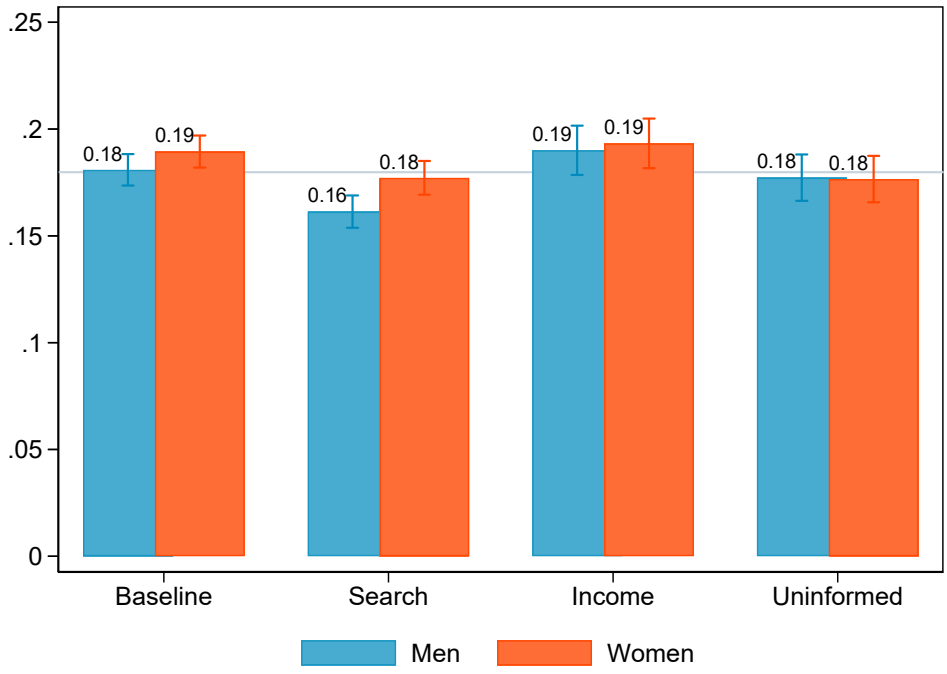
The experiment resulted in a response rate of one half, and a valid price estimate in close to one out of five emails, yielding 10,323 valid price estimates. While the aggregated follow-up rates are similar across groups, there are some minor differences in the rate of provided quotes, as illustrated in Figure 3. These small differences may introduce selection problems, which are discussed in the next section. Women in the *baseline* customer type groups receive 0.9 percentage point more quote replies than *baseline* men customers, however, this difference is not significant. The second difference across genders is that in the *search* type groups women also receive more replies (1.5 percentage points). Within gender groups,

women in the *uninformed* group receive an average of 1.2 percentage point fewer estimates than in the *baseline* group. One could speculate that this difference may be explained by shops being slightly less confident on the need to replace a radiator when an *uninformed* woman customer contacts them. Lastly, there is a 1.5 percentage point decrease in the reply rate for customers in the *search* group relative to the *baseline* group.

There are other three main issues of potential concern for the sample of shops that perform the requested repair and reply to my emails. 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 more than 90 percent of shop replies are received within two days since the first email inquiry, suggesting that email communication must be common enough for them to reply

Figure 3: Price estimate reply rates



Notes: This figure shows average shop replies with price estimates across gender and customer types: Baseline, search, income, and uninformed. The plot shows the 95% confidence intervals and the sample mean reply rate is marked by a grey line within the plot.

promptly to the inquiries. Accordingly, less than one 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 13 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 about the experimental design saliency, I find evidence suggesting that employees pay attention to the email content. Table 1 shows that half of the replies mention the customer's name, indicating that this feature is salient to shops. Furthermore, the name to whom the reply is addressed to corresponds to the assigned name found in the email signature, and not to that in the email address - which is a combination of the name and an additional letter or number (e.g., shops mention Geoffrey instead of Mgeoffrey observed in the email murray.mgeoffrey@gmail.com). About three

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

	Total (1)	Men (2)	Women (3)	*p*-value (4)	N all (5)
A. Reply detail					
On 1st day	0.77	0.77	0.77	0.49	10,323
On 2nd day	0.14	0.13	0.15	0.03	10,323
Not use email	0.003	0.002	0.003	0.31	10,323
Price on 1st reply	0.83	0.83	0.82	0.17	10,323
Price on attachment	0.13	0.12	0.14	0.06	10,323
B. Reply content					
Name	0.54	0.52	0.56	0.00	10.323
Gender	0.03	0.04	0.02	0.00	10,323
Profession	0.03	0.04	0.02	0.10	1,704
Repair Pal	0.02	0.03	0.01	0.00	1,436

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

percent of shops make a reference to the customer's gender, another three percent to the professional title (in the case of *income* customer types), and about two 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 include car details mentioned in the initial email when they send an attached document with an estimate. This is not a feature intended to be salient, but 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 both by gender and within gender-customer type groups (see Figures **??**). All groups are distributed similarly across the US, with more density in more populated markets. Reassuringly, there are no statistically significant differences by gender or gender-customer type with respect to the share of invalid emails, which further validates the randomization (See Figure A.3). Finally, 1.7 percent of emails originally assigned to a shop were not sent due to implementation issues such as having one email account disabled. This discrepancy between assigned and sent emails, however, did not modify the distribution of treatment groups.

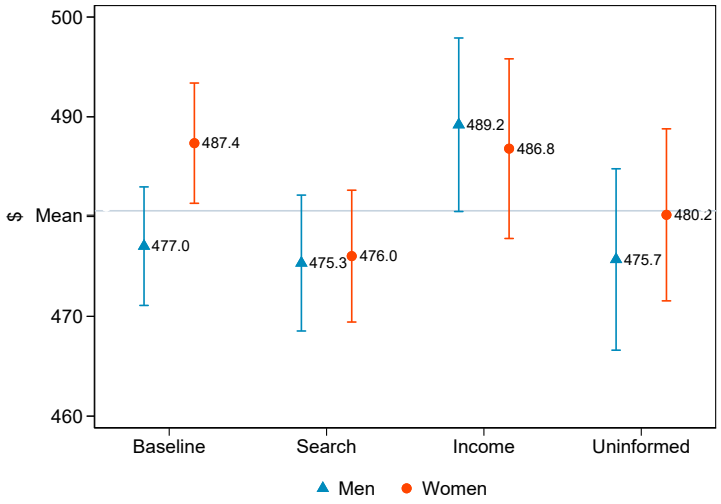
4 Results

4.1 Is there a gender price-gap?

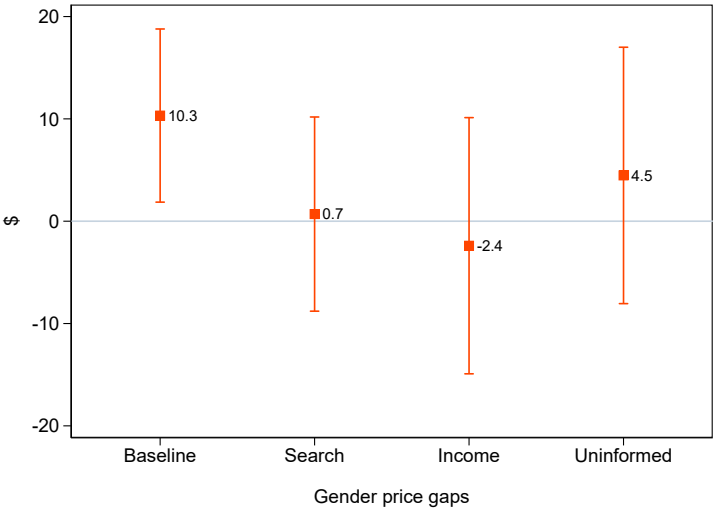
Figure 4 presents the main results. Panel (a) shows the mean price estimates by gender and customer type, while panel (b) plots the gender price-gap for each type. This figure shows that there is a gap for customers in the *baseline* group. Indeed, women in this group receive price estimates of roughly 487.4 dollars - about 10.4 dollars more than men. As later detailed, this gender price gap is not driven by differential response rates, changes in the composition of the provided quotes, and is robust to alternative specifications and alternative measures of total quotes provided to shops.

The implied gender gap in the baseline group is a small but a significant amount per repair. The 10.4 dollar gender differential in the *baseline* group represents over a 2 percent

Figure 4: Price estimates



A. Price Levels



B. Gender price gaps

Note: These figures show the price estimate differences across gender and customer type. Panel A plots raw average price estimates with their 95% confidence intervals for each group. The sample average price estimate is marked by a grey line. 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.

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 gender price

gap of almost six percentage dollar increase for women.[10] This difference may be due to the changes in discrimination over time, as Busse, Israeli and Zettelmeyer (2017) experiment took place six years before, 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.[11]

The results are robust to the inclusion of experimental design components as well as to state and commuting zone fixed effects. Appendix Section 8 replicates the analysis in a simple regression framework. Figure A.5 plots point estimates from a regression-adjusted specification of price estimates for each group relative to men in the baseline, and shows the main results are similar. Table 2 shows additional robustness checks.[12] 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. Indeed, the gender-based premium in the baseline group remains significant across specifications.

Reassuringly, the results are also robust to alternative definitions of total quotes, such as excluding discounts in quotes (see Appendix Figure A.6). There is also no significant

[10] 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.

[11] 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 et al. (2017) in a related setting.

[12] See Appendix A.6 for complete results

difference in the rate of replies received by women and men customers in the *baseline* group (see Figure 3), or a differential amount of components included in the quotes, on top of the radiator, antifreeze and labor (see Figure 5), as further detailed in next subsection.

4.2 How do prices vary with additional information?

By contrast to the observed gender gaps in the *baseline* group, gender gaps disappear once additional information is observed by shops. Customers in the *search, income*, and *uninformed* groups receive price estimates that do not vary significantly by perceived gender but price shifts have different patterns. This 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., *search* group) helps women get an estimate of 476 dollars, 11.4 dollars lower than their *baseline*. Men, by contrast, do not profit from this signal. Their quotes remain the same, suggesting that being a male and having low-search costs in this market are considered substitutes. By contrast, signaling high income seems to particularly hurt men, who receive estimates of 489.2 dollars, similar to women in the baseline group. Women quotes in this group remain the same. 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 of 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.

Reassuringly, results are robust to the inclusion of controls, and none of the additional experimental design features are relevant to explain variation in prices. 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 12.7 to 14.2 dollars relative to men in the baseline group (See Appendix A.6).

Are there differences in reply rates?

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

Table 2: Robustness: Impact of gender and customer type on prices

	Dependent variable: Price estimates		
	(1)	(2)	(3)
Women	10.948**	10.415**	10.526**
	(4.252)	(4.129)	(4.192)
Search	-0.911	-0.443	0.102
	(4.557)	(4.439)	(4.535)
Search ×Women	-10.440	-11.705*	-11.352*
	(6.416)	(6.205)	(6.311)
Income	12.909**	12.722**	14.236***
	(5.302)	(5.183)	(5.363)
Income ×Women	-13.816*	-13.014*	-14.025*
	(7.585)	(7.372)	(7.608)
Uninformed	-1.518	-3.107	-1.972
	(5.453)	(5.266)	(5.372)
Uninformed ×Women	-5.919	-4.865	-5.124
	(7.624)	(7.339)	(7.540)
Controls	State FE, all items, DOW, week FE	+ Shop type	+ CZ FE
Observations	10,323	10,323	10,323

Note: This table reports coefficients from a regression of total price estimates on an indicator for woman customer, for each customer type, and their interaction with woman customer, with baseline customer type as the omitted category. 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 commuting zone fixed effects, with a dummy for observations not linked to a commuting zone. Robust standard errors in parenthesis, with significance levels: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

in Section 2.5. This may introduce a problem of sample selection bias. For instance, women in the search group are 1.2 percentage points less likely to receive a reply with an estimate than women in the baseline group. 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. Replies might vary with day of the week, and population nearby, whereas prices might not. For instance, shops might be more busy and find it harder to reply on Fridays, whereas quotes should not vary with this. For completeness, I do this for the comparison groups that have significant adjusted differences in the rate of

quote replies, or whose quote estimates vary between groups. Table A.7 in the Appendix shows the results, which suggest that gender-price gap in the baseline group is not driven by selection bias.[13]

Are there gender differences in diagnostic recommendations and customer service?

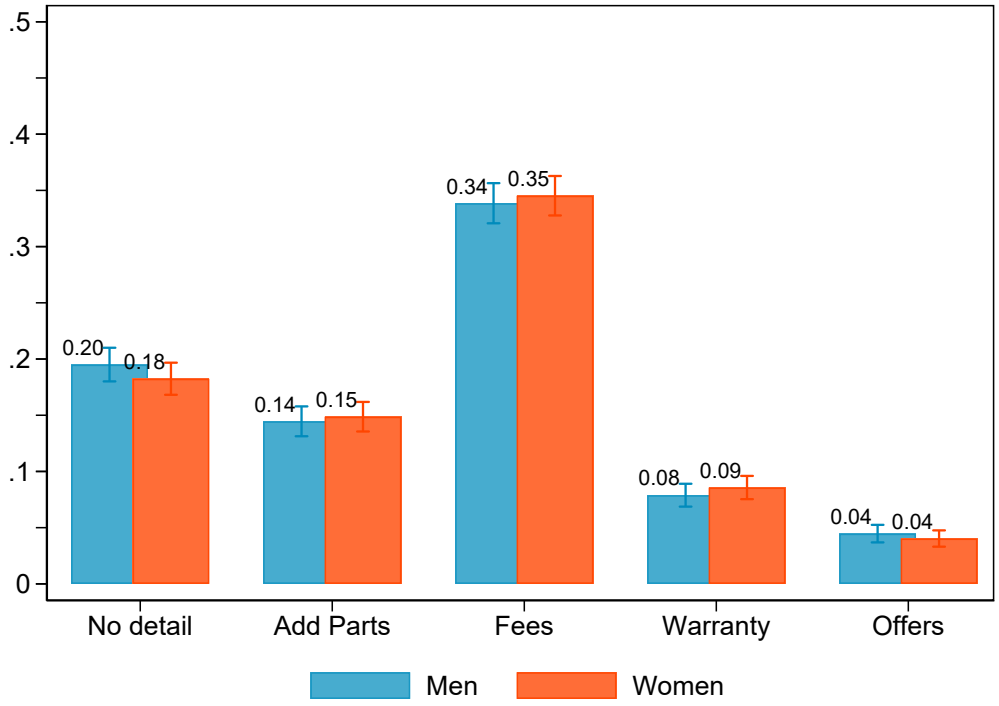
Shops can directly overcharge customers, but also vary prices by over-treatment of the repair. That is, by offering more than the necessary repairs. In my setting, customers specifically request a quote for a radiator but mechanics can also include the quote of additional parts, on top of labor, coolant and a new radiator in a quote. This brings another channel through which shops can treat customers differently. Prices might be similar, but would shops include more items when providing quotes to women? Shops can treat customer differently by also changing complementary services they offer. Overall, I do not find differential recommendations by gender. Figure 5 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.11 in the Appendix shows that shops are less likely to provide a detailed quote to *uninformed* men customers, while they are marginally more likely to offer these to high-*income* type customers and to women in the *baseline* group. This difference, however, does not translate into decreases (or increases) in additional recommended replacements relative to men 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 by customer type 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-

[13] 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 bounds for any of the groups with significant differences in the rate of quote replies.

income type 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

Figure 5: Additional parts and service offers



Notes: This figure shows share of estimates that do not provide any detail on the components included (*No detail*), include additional replacement parts on top of a radiator, coolant and labor (*Add parts*), include taxes and fees (*Add parts*), have a warranty (*Warranty*) or include additional offers such as discounts, financing options and price-matching (*Offers*) for baseline customers. The plot shows the 95% confidence intervals, obtained with robust standard errors.

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.

Are there quote differences by shop characteristics?

An additional exercise tries to distinguish between taste-based and statistical discrimination motives, as empirically tested in the literature (Fershtman and Gneezy 2001 Donohue and

Levitt 2001, Antonovics and Knight 2009, and 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. Group bias would imply that people treat members of their own group more favorably than they treat other people. In my setting, I 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

Table 3: Effects by shop characteristics

	Dependent variable: Price estimates		
	(1)	(2)	(3)
Women	10.960**	10.910**	8.674**
	(4.269)	(4.333)	(4.424)
Women employee	8.910*	8.852	8.415
	(5.298)	(5.416)	(5.422)
Women employee ×Women	-3.836	-2.223	-1.366
	(7.158)	(7.374)	(7.384)
Non independent shop	109.467***	110.606***	100.356***
	(4.425)	(4.552)	(6.397)
Non independent shop ×Women			19.725**
			(8.951)
Controls	State FE, mail items, DOW, week FE	+CZ FE	
Women employee (%)	14	14	14
Non-independent	11	11	11

Notes: This table reports coefficients from regressions of total price estimates on shop characteristics (an indicator for perceived women employees, and indicator for non-independent shops) 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. Robust standard errors in parenthesis, with significance levels: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

customer dummy. Table 3 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. Table 3 shows these results. As expected, non-independent shops (franchises and dealerships) provide higher estimates, baseline men customers receive increased estimates of over 100 dollars higher than those from independent shops. Perhaps unexpectedly, however, the relative gender difference increases about 19 dollars more among non-independent shops. The reason for this rather surprising result is not obvious, but I rule out some explanations. First, there is no differential reply rates among non-independent shops. Second, I can also rule out that this increase in price gaps is driven by systematic differences in the composition of quotes. That is, franchises and dealerships do not include relatively more additional replacement parts in the price estimates provided to women than those provided to men. Consistent with List (2004) study, where more experienced recruited sellers provided higher pries to minorities who on expectation where inexperienced at a sports card market, one could consider that dealerships and franchises are be more *experienced* and have larger market power and are, thereby, able to discriminate more. This brings into consideration how market competition levels could impact gender price gaps.

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

From a discrimination perspective, Becker's taste-based model predicts that shops charging a premium to women should not persist in competitive markets. This same conclusion applies in a statistical discrimination framework, as explained in Appendix 7.

From a general perspective, the relationship between gender-price discrimination and competition is theoretically ambiguous. On the one hand, a monopolist with market power can perfectly price discriminate across consumers. Opposite to this, when markets are

perfectly competitive and firms have no market power, then the law of one price prevails and price discrimination cannot exist. From these extreme cases, it would follow that the degree of price discrimination should decrease as markets become less concentrated. However, this theoretical prediction is not supported in intermediate cases. To the contrary, theoretical studies suggest that price discrimination actually can increase when moving from a monopoly to imperfect competition (Borenstein1985, Borenstein1994, and Holmes 1989).

Thus, the relationship between price discrimination and competition becomes an empirical question. This field experiment provides and ideal setting to study whether gender-based discrimination becomes more or less pervasive in more competitive markets. I combine experimental results of the effects of gender on prices with non-exogenous variation in market structure. Taking advantage of rich detailed data on the location of each car repair shop and the variation of geographic characteristics across the country, I define measures of competition based on the density of nearby competitors.

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 4 presents the main results, including several specifications to gain insights into the robustness of the results. Each column regresses price estimates on the number of competitors, the number of competitors interacted with a dummy for perceived women customers, a dummy for women customers, customer type controls, a dummy for non-independent nearby competitors to control for additional characteristics of the type of competitors, as well as the standard main controls from the experimental design (mail varying items, time, state and commuting zone fixed effects).

Overall, the results suggest that there is a robust negative association between gender based price discrimination and increases in the level of competition. When there are no nearby competitors, the gender price gap for baseline customers increases to approximately 18 dollars. This is a significant increase with respect to the 10 dollar gap shown in table 2. Each additional competitor is associated with a decrease in average prices. Importantly, it is associated with an additional significant decrease in estimates provided to women. Women receive estimates of roughly 1.6 fewer dollars with an additional competitor nearby. Taken at face value, this estimates imply that price gaps would disappear with eleven shops.

This result suggests that while consumer heterogeneity may lead shops to price discriminate, the extent of competition also matters. This supports the hypothesis that price discrimination is negatively associated with the degree of competition. Unlike the previous exercises comparing prices changes by employee gender and customer types, these results are consistent with taste-based discrimination. 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 (Becker, 1957). Nevertheless, this result can also be rationalized by a statistical discrimination framework, in which sellers are allowed to discriminate based on observable customer characteristics as well as the level of market competition. Appendix 7 illustrates this mechanism in the context of a simple model in which statistical discrimination is allowed. In my setting, the degree in which shops vary prices with perceived 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.

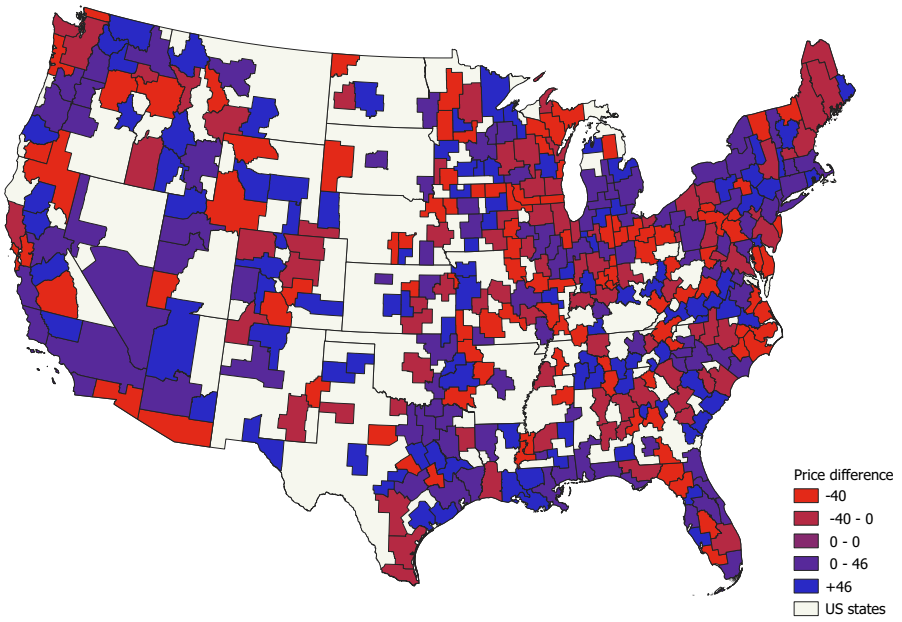
Table 4: Effects on price gaps with competition

	Dependent variable: Price estimates		
	(1)	(2)	(3)
Competitors	-0.857**	-0.455	-0.741**
	(0.371)	(0.365)	(0.369)
Competitors ×Women	-1.637***	-1.584***	-1.614***
	(0.483)	(0.473)	(0.473)
Women	18.548***	17.705***	17.999***
	(4.757)	(4.640)	(4.692)
Controls	State FE, mail items, DOW, week FE	+ Shop type	+ CZ FE

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: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

As an additional exercise, I show the results are robust to alternative samples and market definitions in Tables A.9 and A.10. First, I verify that the observed pattern is not driven by outliers and influential points by running a robust regression, which re-weights observations based on how well behaved they are (Fox, 1997), see Column 2, Table A.9. Second, I show that the results are qualitatively similar if I restrict the sample of competitor shops to those that at least have an email in their YP listing. These could be considered the shops with "online" presence that would compete for customers using email communication. Finally, 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.10 by using radiuses thresholds that increase by 0.5 km (1, 1.5, 2, 2.5, and 3 km, respectively). Results are qualitatively similar, but quantitatively the results gets dampened with larger distances.

Figure 6: 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.

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 demographic characteristics are not strongly correlated with gender price gaps. Figure 6 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.

5 Conclusion

This study presents evidence that women customers receive worse estimates than their male counterparts when no information other than gender is provided to shops. Women in the *baseline* group receive price estimates of roughly 487 dollars - about 10 dollars more than men. This price gap remains once we include controls for other experimental design features or geographic characteristics. Furthermore, this gender price gap is not driven by differential response rates or changes in the composition detail of the provided quotes. These effects, however, disappear once additional information about the customer is revealed. Once customers additionally signal to be in the *uninformed*, *search*, or *income* groups, they receive price estimates that do not vary significantly by perceived gender. Women benefit when they mention they are searching for other price quotes, since they receive estimates about 11 dollars lower, while men quotes do not vary with this signal. Thus, this signal effectively closes the gap by lowering female price quotes. On the other hand, signaling high income also closes this gender price gap, but by hurting men. Men whose implicit income is high receive estimates close to 490 dollars, similar to women estimates in the baseline group.

Overall, these results suggest that discrimination patterns are explained by a statistical discrimination model. Shops priors are that absent additional information, women have a higher willingness to pay but once additional information is revealed, gender becomes less relevant. This results may be consistent with shops beliefs that women have higher search

costs than men, thereby helping women get more competitive prices once they signal the opposite, and hurting men by getting higher prices once they signal higher search costs.

This paper also examines the extent to which discrimination patterns vary with shop and market characteristics. I provide further evidence suggesting that discrimination arises through a signal extraction problem, rather than simple distaste towards women. I find that results are not affected by employee's gender, which indicates that employee's preferences for a specific customer are not driving the results. Consistent with previous literature on discrimination in car sales, I find that dealerships and franchises discriminate against women, and interestingly, this bias is greater than that of independent shops. Importantly, the disadvantages faced by women are reduced when markets are more competitive.

As with all such correspondence studies, there is an important consideration to these findings. Ultimately, the value we are most interested in is the final price paid once a repair is done, and I observe an intermediate outcome, the price estimate. Car repair shops have both an incentive to provide low price estimates to attract a new customer, but also to provide an estimate close to the final price to avoid surprising customers with higher quotes at the risk of losing them. This paper setting makes it difficult to provide lower estimates than final, as they are written and not given over a phone call. Furthermore, a discrepancy between estimates and final prices would be problematic if we thought that shops increased prices by more to men than to women, vanishing - or reversing - the discrimination price gaps against women. On the other hand, discrimination against women 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 with respect to a setting in which only in-person visits occur. The internet facilitates information search and has been found to be particularly beneficial for individuals 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. The representativeness of this paper results should be interpreted with caution. 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 YP listings have an available email address, and close to one out of five emails sent replied with an estimate. This

motivates further comparisons between shops, and an in-depth exploration of the relation between discrimination patterns and geographic characteristics.

Overall, this paper shows that gender-based price discrimination still exists despite increased progress toward 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 customers are searching for other price quotes helps them get more competitive prices and closes the gender-based premium, and the gender premium also decreases with increased competition. In this setting, competition may be a powerful tool to reduce gender-based discrimination.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens and Jeffrey M. Wooldridge. (2022). "When should you adjust standard errors for clustering?". *The Quarterly Journal of Economics,* 138(1), pp. 1-35. https://doi.org/10.1093/qje/qjac038

Aigner, Dennis J., and Glen G. Cain. (1977). "Statistical theories of discrimination in labor markets". *ILR Review,* 30(2), pp. 175-187. https://doi.org/10.1177/001979397703000204

Anderson, Simon P., and André De Palma. (2005). "Price dispersion and consumer reservation prices". *Journal of Economics & Management Strategy,* 14(1), pp. 61-91. https://doi.org/10.1111/j.1430-9134.2005.00034.x

Antonovics, Kate L., and Brian G. Knight. (2009). "A new look at racial profiling: Evidence from the Boston Police Department". *The Review of Economics and Statistics,* 91(1), pp. 163-177. https://doi.org/10.1162/rest.91.1.163

Arrow, Kenneth. (1973). "The Theory of Discrimination". In Orley Ashenfelter and Albert Rees (eds.), *Discrimination in labor markets.* Princeton University Press, pp. 3-33. https://www.jstor.org/stable/j.ctt13x10hs.5

Ayres, Ian, and Peter Siegelman. (1995). "Race and gender discrimination in bargaining for a new car". *The American Economic Review,* 85(3), pp. 304-321. https://www.jstor.org/stable/2118176

Barron, John M., Beck A. Taylor and John R. Umbeck. (2004). "Number of sellers, average prices, and price dispersion". *International Journal of Industrial Organization,* 22(8-9), pp. 1041-1066. https://doi.org/10.1016/j.ijindorg.2004.05.001

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

Bertrand, Marianne, and Esther Duflo. (2017). "Field experiments on discrimination". In Abhijit Vinayak Banerjee and Esther Duflo (eds.), *Handbook of economic field experiments,* vol. 1. North-Holland, pp. 309-393. https://doi.org/10.1016/bs.hefe.2016.08.004

Bertrand, Marianne, and Sendhil Mullainathan. (2004). "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination". *The American Economic Review,* 94(4), pp. 991-1013. https://doi.org/10.1257/0002828042002561

Bessendorf, Anna. (2015). *From cradle to cane: The cost of being a female consumer. A study of gender pricing in New York City.* New York City Department of Consumer Affairs. https://www.nyc.gov/assets/dca/downloads/pdf/partners/Study-of-Gender-Pricing-in-NYC.pdf

Borenstein, Severin. (1985). "Price discrimination in free-entry markets". *The RAND Journal of Economics,* 16(3), pp. 380-397. https://www.jstor.org/stable/2555565

Borenstein, Severin, and Nancy L. Rose. (1994). "Competition and price dispersion in the US airline industry". *Journal of Political Economy,* 102(4), pp. 653-683. https://doi.org/10.1086/261950

Busse, Meghan R., Ayelet Israeli and Florian Zettelmeyer. (2017). "Repairing the damage: The effect of price knowledge and gender on auto repair price quotes". *Journal of Marketing Research,* 54(1), pp. 75-95. https://doi.org/10.1509/jmr.13.0291

Castillo, Marco, Ragan Petrie, Maximo Torero and Lise Vesterlund. (2013). "Gender differences in bargaining outcomes: A field experiment on discrimination". *Journal of Public Economics,* 99, pp. 35-48. https://doi.org/10.1016/j.jpubeco.2012.12.006

Darby, Michael R., and Edi Karni. (1973). "Free competition and the optimal amount of fraud". *The Journal of Law and Economics,* 16(1), pp. 67-88. https://doi.org/10.1086/466756

Doleac, Jennifer L., and Luke C. D. Stein. (2013). "The visible hand: Race and online market outcomes". *The Economic Journal,* 123(572), pp. F469-F492. https://doi.org/10.1111/ecoj.12082

Donohue III, John J., and Steven D. Levitt. (2001). "The impact of race on policing and arrests". *The Journal of Law and Economics,* 44(2), pp. 367-394. https://doi.org/10.1086/322810

Dulleck, Uwe, and Rudolf Kerschbamer. (2006). "On doctors, mechanics, and computer specialists: The economics of credence goods". *Journal of Economic Literature,* 44(1), pp. 5-42. https://doi.org/10.1257/002205106776162717

Dulleck, Uwe, Rudolf Kerschbamer and Matthias Sutter. (2011). "The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition". *American Economic Review,* 101(2), pp. 526-555. https://doi.org/10.1257/aer.101.2.526

Edelman, Benjamin, Michael Luca and Dan Svirsky. (2017). "Racial discrimination in the sharing economy: Evidence from a field experiment". *American Economic Journal: Applied Economics,* 9(2), pp. 1-22. https://doi.org/10.1257/app.20160213

Fershtman, Chaim, and Uri Gneezy. (2001). "Discrimination in a segmented society: An experimental approach". *The Quarterly Journal of Economics,* 116(1), pp. 351-377. https://doi.org/10.1162/003355301556338

Fox, John. (1997). *Applied regression analysis, linear models, and related methods.* SAGE Publications.

Gerardi, Kristopher S., and Adam Hale Shapiro. (2009). "Does competition reduce price dispersion? New Evidence from the airline industry". *Journal of Political Economy,* 117(1), pp. 1-37. https://doi.org/10.1086/597328

Gneezy, Uri, John List and Michael K. Price. (2012). "Toward an understanding of why people discriminate: Evidence from a series of natural field experiments". NBER Working Paper Series, 17855, National Bureau of Economic Research. https://doi.org/10.3386/w17855

Guryan, Jonathan, and Kerwin Kofi Charles. (2013). "Taste-based or statistical discrimination: The economics of discrimination returns to its roots". *The Economic Journal,* 123(572), pp. F417-F432. https://doi.org/10.1111/ecoj.12080

Heckman, James J. (1998). "Detecting Discrimination". *Journal of Economic Perspectives,* 12(2), pp. 101-116. https://doi.org/10.1257/jep.12.2.101

Holmes, Thomas J. (1989). "The effects of third-degree price discrimination in oligopoly". *The American Economic Review,* 79(1), pp. 244-250. https://www.jstor.org/stable/1804785

Jowell, Roger, and Patricia Prescott-Clarke. (1970). "Racial discrimination and white-collar workers in Britain". *Race,* 11(4), pp. 397-417. https://doi.org/10.1177/030639687001100401

Kerschbamer, Rudolf, and Matthias Sutter. (2017). "The economics of credence goods — a survey of recent lab and field experiments". *CESifo Economic Studies,* 63(1), pp. 1-23. https://doi.org/10.1093/cesifo/ifx001

Lee, David S. (2009). "Training, wages, and sample selection: Estimating sharp bounds on treatment effects". *The Review of Economic Studies,* 76(3), pp. 1071-1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Levitt, Steven D., and Stephen J. Dubner. (2006). *Freakeconomics.* William Morrow Paperbacks. Revised and expanded edition.

List, John. (2004). "The nature and extent of discrimination in the marketplace: Evidence from the field". *The Quarterly Journal of Economics,* 119(1), pp. 49-89. https://doi.org/10.1162/003355304772839524

Morton, Fiona Scott, Florian Zettelmeyer and Jorge Silva-Risso. (2003). "Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities?". *Quantitative Marketing and Economics,* 1(1), pp. 65-92. https://doi.org/10.1023/A:1023529910567

Neumark, David. (2018). "Experimental research on labor market discrimination". *Journal of Economic Literature,* 56(3), pp. 799-866. https://doi.org/10.1257/jel.20161309

Nunley, John M., Mark F. Owens and R. Stephen Howard. (2011). "The effects of information and competition on racial discrimination: Evidence from a field experiment". *Journal of Economic Behavior & Organization,* 80(3), pp. 670-679. https://doi.org/10.1016/j.jebo.2011.06.028

Phelps, Edmund S. (1972). "The statistical theory of racism and sexism". *The American Economic Review,* 62(4), pp. 659-661. https://www.jstor.org/stable/1806107

Pope, Devin G., and Justin R. Sydnor. (2011). "What's in a picture? Evidence of discrimination from Prosper.com". *Journal of Human Resources,* 46(1), pp. 53-92. https://doi.org/10.1353/jhr.2011.0025

Price, Joseph, and Justin Wolfers. (2010). "Racial discrimination among NBA referees". *The Quarterly Journal of Economics,* 125(4), pp. 1859-1887. https://doi.org/10.1162/qjec.2010.125.4.1859

Ravina, Enrichetta. (2007). "Beauty, personal characteristics and trust in credit markets". *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.972801

Riach, Peter A., and Judith Rich. (1991). "Testing for racial discrimination in the labour market". *Cambridge Journal of Economics,* 15(3), pp. 239-256. https://doi.org/10.1093/oxfordjournals.cje.a035168

Riach, Peter A., and Judith Rich. (2002). "Field experiments of discrimination in the market place". *The Economic Journal,* 112(483), pp. F480-F518. https://doi.org/10.1111/1468-0297.00080

Schneider, Henry S. (2012). "Agency problems and reputation in expert services: Evidence from auto repair". *The Journal of Industrial Economics,* 60(3), pp. 406-433. https://doi.org/10.1111/j.1467-6451.2012.00485.x

Shepard, Andrea. (1991). "Price discrimination and retail configuration". *Journal of Political Economy,* 99(1), pp. 30-53. https://doi.org/10.1086/261739

Siegelman, Peter, and James J. Heckman. (1993). "The Urban Institute audit studies: Their methods and findings". In Michael Fix and Raymond J. Struyk (eds.), *Clear and convincing evidence: Measurement of discrimination in America.* Urban Institute Press, pp. 187-258.

Stavins, Joanna. (2001). "Price discrimination in the airline market: The effect of market concentration". *The Review of Economics and Statistics,* 83(1), pp. 200-202. https://doi.org/10.1162/rest.2001.83.1.200

Stole, Lars A. (2007). "Price discrimination and competition". In Michael Armstrong and Robert H. Porter (eds.), *Handbook of industrial organization,* vol. 3. North-Holland, pp. 2221-2299. https://doi.org/10.1016/S1573-448X(06)03034-2

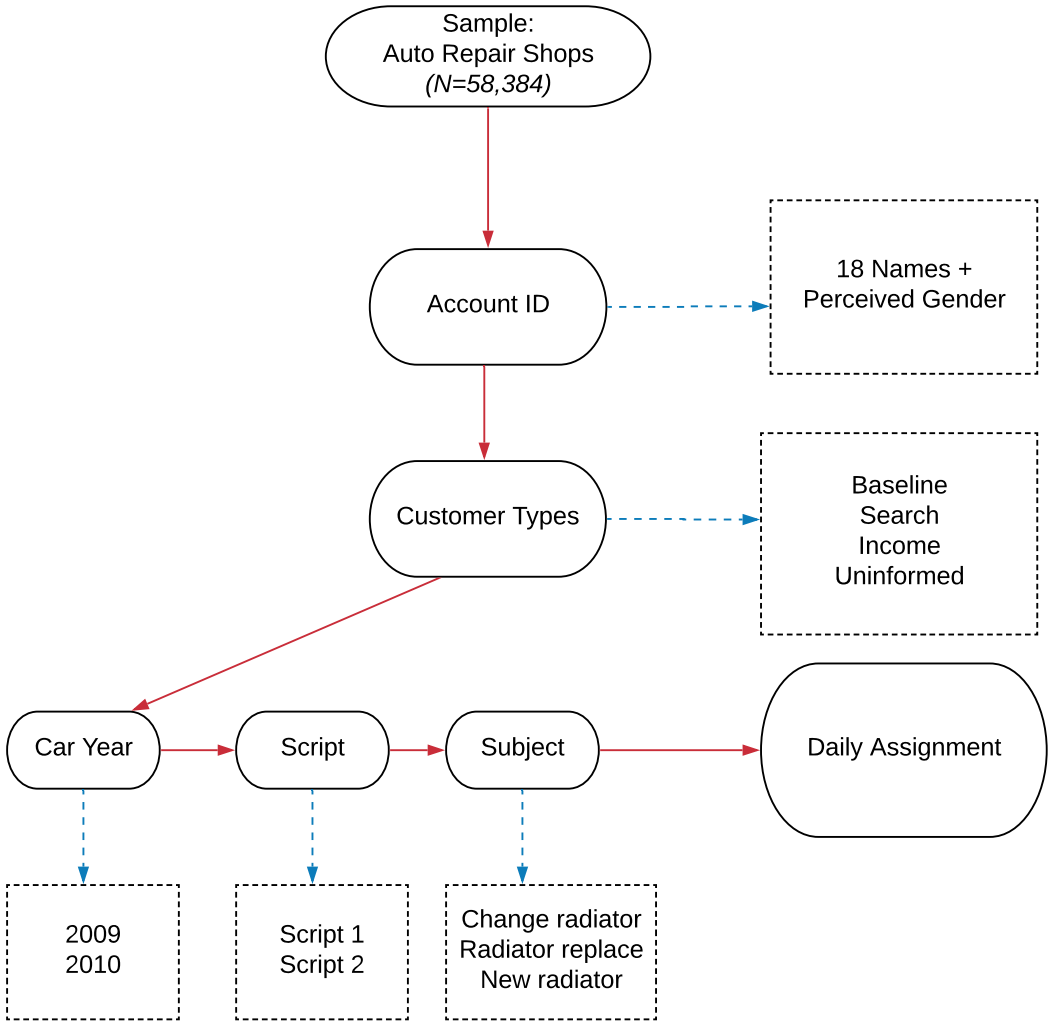
6 Appendix: Experimental design details

6.A 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

Figure A.1: 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.

therefore, assigning half of the sample to women. Then, I randomly assigned each of the four customer type treatments within each gender. Thus, 18. percent of shops will be contacted by woman customers using standard scripts, 16 percent will be contacted by low-search woman customers, 8 percent by high-educated women and 8 percent by uninformed women. Equivalent shares are assigned to each man-type combination.

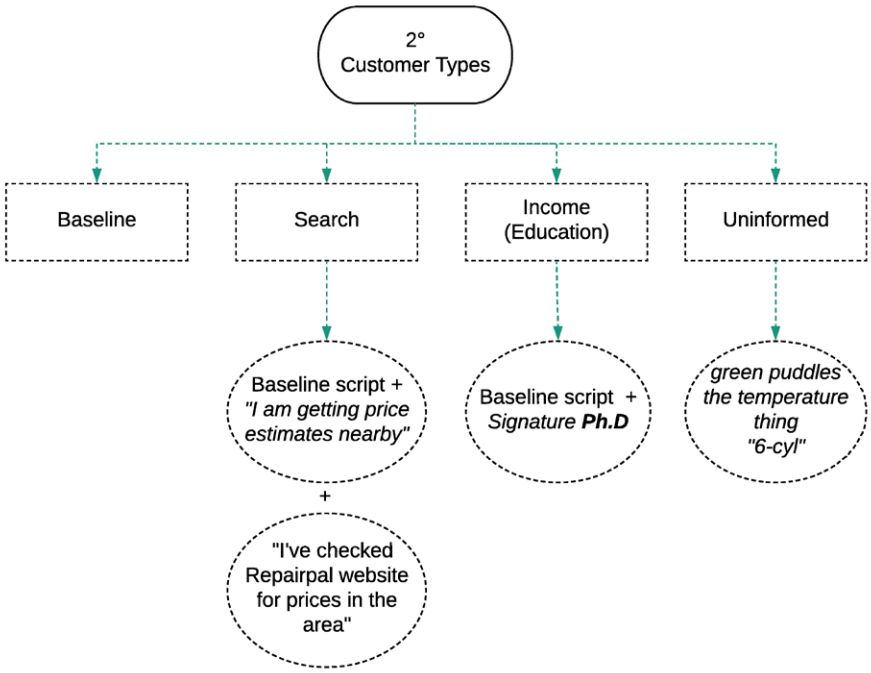
Then, I randomized additional components of each email; the car year, script design and email subject. Each of these variations is included to avoid sending a second identical email to each shop. The car years are 2009 and 2010, both belong to the same car generation. Within each customer type, there are two scrips; 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.

6.B Customer gender and type scripts

This section provides details of customer type manipulation trait. Figure A.2 summarizes the main changes added to the baseline script. Complete scripted details used for each customer type are also presented. 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.

Figure A.2: Customer type variations to Baseline type script



Note: This figure shows each of the main variations to the baseline script used to signal a low-search cost, a high-income, and a uninformed type customer.

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,

${FILL_NAME}, *Ph.D.*"

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}"

6.C 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.

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 it 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}

6.D Data collection rules

With respect to the data collection, I created Gmail labels each week to group emails into *invalid email, not do service, do service and give quote, do service and not give quote*. 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 easily verifiable (i.e., a shop either gives a price or not, mentions a discount or not, etc.).

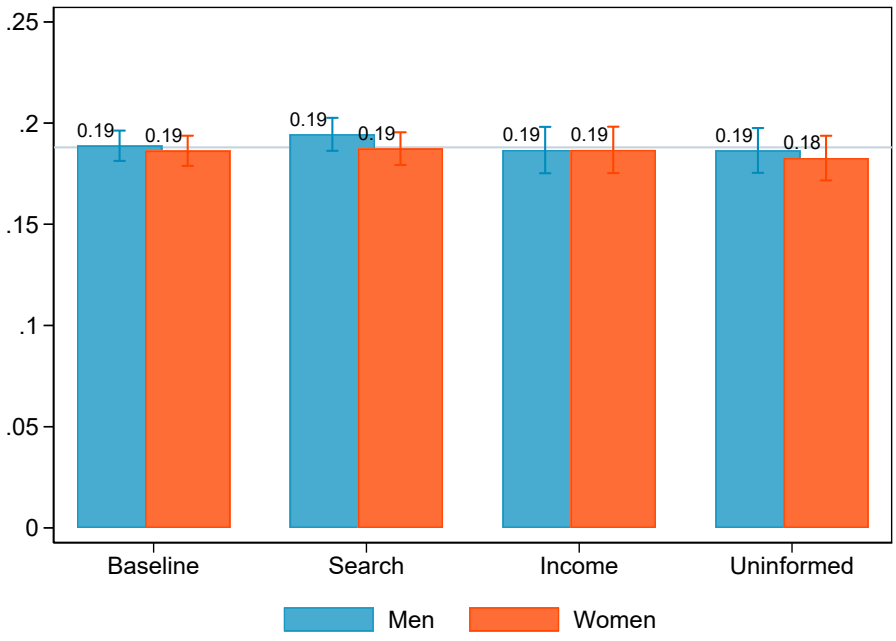
The most relevant pre-specified rules used to translate email replies into a spreadsheet are the following: keep one observation for each contacted shop, the reported time and date of each reply corresponds to the first reply received by each shop. Identify if a reply is sent before obtaining a quote 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.

6.E Design validation

This section shows further validations of the experimental design implementation. Figure A.3 shows the distribution of invalid emails by treatment group, which reassuringly shows not significant differential across groups. Table A.5 complements these results showing the reply rates by gender across customer types and withing customer type across genders for all invalid email replies, as well as replies of shops not performing Honda radiator replacements, performing the service and restricting to those that perform the service and provide an estimate.

Figure A.3: 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.

7 Theoretical Framework

This section provides a simple theoretical framework to help guide the empirical exercise in Section 4. 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 θ_i. This parameter θ_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 θ_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.[14] 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^ks 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 λ_i of shopping in an alternative store. This parameter λ_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 λ_i. The counterpart of λ_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 λ_i to be typically lower. In this sense, the correct interpretation is that

[14] First-order stochastic dominance is implied but not enough to guarantee this result.

Table A.5: Email replies: Impact of gender and gender-type on replies

	Invalid (1)	Not repair (2)	Repair (3)	Give quote (4)	Give quote (5)
A. Men customers					
Search	0.006	-0.004	-0.034***	-0.020***	-0.020***
	(0.006)	(0.003)	(0.006)	(0.005)	(0.005)
Income	-0.002	-0.002	-0.002	0.009	0.009
	(0.007)	(0.004)	(0.008)	(0.007)	(0.007)
Uninformed	-0.002	-0.004	0.010	-0.004	-0.005
	(0.007)	(0.004)	(0.008)	(0.007)	(0.007)
Constant	0.189***	0.055***	0.252***	0.181***	
	(0.004)	(0.002)	(0.004)	(0.004)	
A. Women customers					
Search	0.001	-0.006*	-0.023***	-0.012**	-0.013**
	(0.006)	(0.003)	(0.006)	(0.006)	(0.006)
Income	0.000	-0.003	-0.000	0.004	0.004
	(0.007)	(0.004)	(0.008)	(0.007)	(0.007)
Uninformed	-0.004	0.003	0.006	-0.013*	-0.012*
	(0.007)	(0.004)	(0.008)	(0.007)	(0.007)
Constant	0.186***	0.056***	0.258***	0.189***	
	(0.004)	(0.002)	(0.004)	(0.004)	
C. Gender differential by customer type					
Search×Women	-0.007	-0.002	0.018***	0.016***	0.015***
	(0.006)	(0.003)	(0.006)	(0.006)	(0.006)
Income×Women	0.000	-0.001	0.007	0.003	0.003
	(0.008)	(0.005)	(0.009)	(0.008)	(0.008)
Uninformed×Women	-0.004	0.007	0.002	-0.001	0.000
	(0.008)	(0.005)	(0.009)	(0.008)	(0.008)
Baseline×Women	-0.002	0.000	0.006	0.009	0.008
	(0.005)	(0.003)	(0.006)	(0.005)	(0.005)
Main Controls					Yes

Notes: This table reports coefficients from regressions of invalid email reply rates (Column 1), not do service reply rates (Column 2), do service reply rates (Column 3), and give quote estimate reply rates (Columns 4 and 5) on customer types by gender (Panel A shows results for men, Panel B for women) with baseline type customers as the omitted category. Panel C includes both genders and adds the interaction coefficients between each customer type and women customers. Column 5 controls for state fixed effects, non-independent shop and email varying itemsRobust standard errors in parenthesis, with significance levels: $^{*}p < 0.10, ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$.

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. [15] Importantly, in the model λ_i is not an informational cost. Indeed, in the equilibrium I describe below there is a pure

[15]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.

strategies equilibrium so agents fully anticipate the distribution of prices at other stores in the Nash equilibrium.[16] However, the credible threat of switching stores is key for the properties of the Nash equilibrium.

Consumer i utility function is given by

$$u_i = \left\{ \begin{array}{c} \theta_i - p_s \text{ if } i \text{ buys in the closest shop} \\ \theta_i - p_{s'} - \lambda_i \text{ if } i \text{ buys in a different shop} \end{array} \right\}$$

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

$$\pi_j^{kl} = \underbrace{p_j^{kl}\left(1-G^k(p_j^{kl})\right)}_{\text{initial customers}} - \underbrace{1_{p_j^{kl}>p_{-j}^{kl}} p_j^{kl}\left(F(p_j^{kl}-p_{-j}^{kl})\right)}_{\text{customers lost if price is high}} + \underbrace{1_{p_j^{kl}<p_{-j}^{kl}} p_j^{kl}\int_0^{p_{-j}^{kl}-p_j^{kl}} f^l(\lambda)(1-G^k(p_j^{kl}+\lambda))d\lambda^l}_{\text{customers gained if price is low}}$$

When $p_j^{kl} \to p_{-j}^{kl}$, one can show the first-order condition converges to[17,18]

$$\frac{1-G^k(p_j^{kl})}{g^k(p_j^{kl})}(1-p_j^{kl}f^l(0)) - p_j^{kl} = 0. \tag{1}$$

If consumers were immobile, i.e., if $f^l(0) \to 0$, then p_j^{kl} is the monopoly price. Each p_j^{kl} varies with customer observable characteristics, as stores engage in statistical discrimination

[16] 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, as I show below, this "physical" cost is very well-behaved in the sense that the Nash equilibrium features a number of desirable properties, while the predictions regarding informational frictions and search costs depend significantly on details of the model and may give rise to a number of counter-intuitive predictions, such as the Diamond paradox (for a survey see (Stole, 2007)).

[17] 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.

[18] 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.

to maximize profits.[19] Proposition 7 shows that this is still the case when switching costs are nontrivial. By contrast, if $f^l(0) \to \infty$, $p_j^{kl} \to 0$. 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 7 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. In a symmetric Nash equilibrium, the following is true:

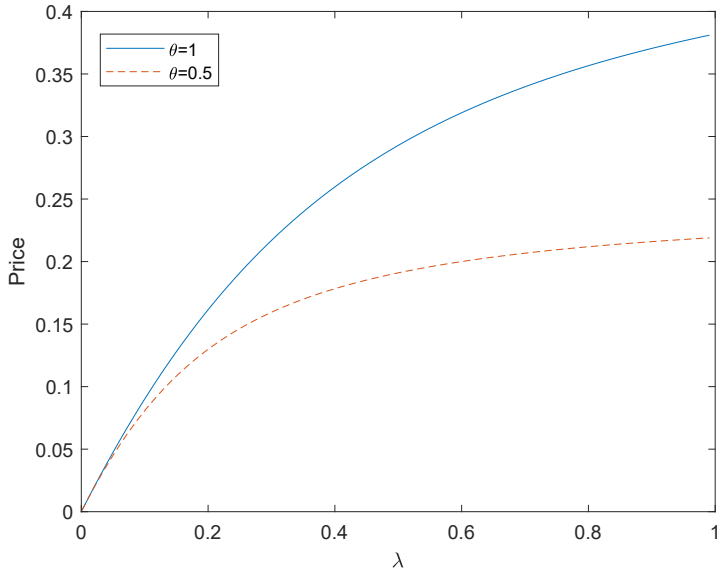
(a) Suppose $G^k(\cdot)$ satisfies the monotone likelihood ratio property with respect to $G^{k'}(\cdot)$. Then, everything else equal, agents of type k are charged a higher price, i.e., $p^{kl} > p^{kl'}$.

(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 $G^k(x) = \frac{\bar{\theta}^k - x}{\bar{\theta}^k - \underline{\theta}^k}$, 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 $p^{kl} - p^{k'l} > p^{kl'} - p^{k'l'}$.

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.4: Price Discrimination and Competition



[19] 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.

Figure A.4 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.

8 Price differences by treatment group: Robustness

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

$$p_{jk} = \alpha_0 + \beta_1 Women_k + \beta_2 Type_k + \beta_3 (Type_k \times Women_k) + \omega_s + \theta X_{jk} + \varepsilon_{jk} \qquad (2)$$

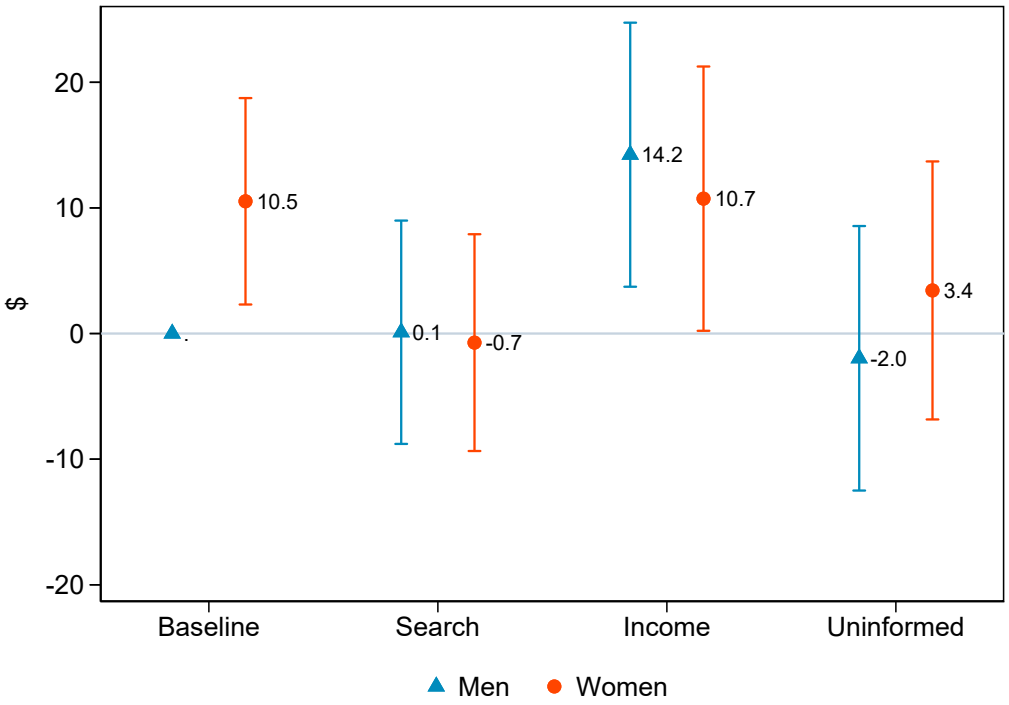
where p_{jk} is the estimate that shop j provides to k customer (gender and type customer combination), $Women_k$ is an indicator variable equal to 1 when a woman customer contacts shop j, $Type_k$ is a vector with indicator variables for each customer type. The omitted category is always the baseline type. And, $Type_k \times Women_k$ is the vector of customer type interactions with the woman indicator. ω_s are state fixed effects, X_{jk} correspond to controls of experimental design items that vary including the timing in which the emails were sent. That is, day of the week and week of the experiment in which shop j is contacted, and script number, car year and email subject covariates. Additional specifications include a control for non-independent shops and commuting zone fixed effects. Finally, ε_{jk} is the unobserved component, which has mean zero and is uncorrelated with our treatment groups.[20]

Table A.6 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

[20] I do not cluster standard errors given this field experiment's sampling design and experimental design Abadie et al. (2017)

commuting zone. Figure A.5 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 10.5 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 12.2 to 14.2 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.5: 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.6: Impact of gender and customer type on prices: Robustness

	Dependent variable: Price estimates		
	(1)	(2)	(3)
Women	10.948**	10.415**	10.526**
	(4.252)	(4.129)	(4.192)
Search	-0.911	-0.443	0.102
	(4.557)	(4.439)	(4.535)
Search ×Women	-10.440	-11.705*	-11.352*
	(6.416)	(6.205)	(6.311)
Income	12.909**	12.722**	14.236***
	(5.302)	(5.183)	(5.363)
Income ×Women	-13.816*	-13.014*	-14.025*
	(7.585)	(7.372)	(7.608)
Uninformed	-1.518	-3.107	-1.972
	(5.453)	(5.266)	(5.372)
Uninformed ×Women	-5.919	-4.865	-5.124
	(7.624)	(7.339)	(7.540)
Subject: Change radiator	-2.836	-2.176	-1.503
	(3.217)	(3.111)	(3.188)
Subject: Radiator replace	-2.156	0.135	0.406
	(3.312)	(3.211)	(3.270)
Car year 2010	1.079	1.329	1.542
	(2.598)	(2.515)	(2.560)
Script 1	-3.174	-2.120	-2.153
	(2.617)	(2.532)	(2.575)
Controls	State FE DOW fe, week FE	+ Shop type	+ CZ FE
Observations	10,323	10,323	10,323

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

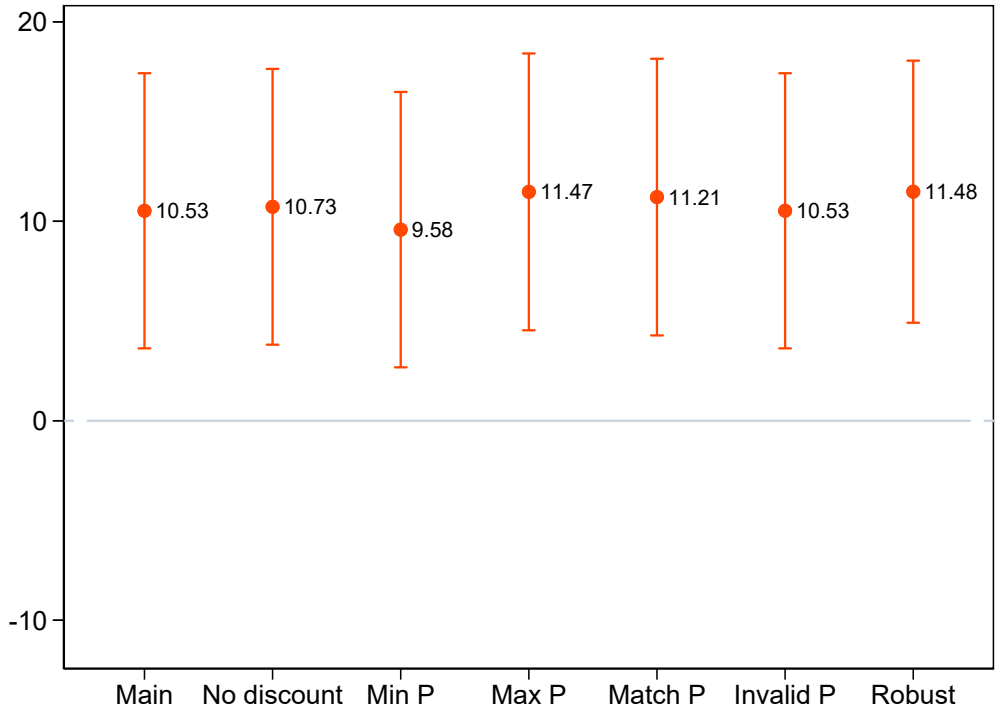
Table A.7: Robustness: Selection in replies

		Two-step correction		Non-parametric bounds	
	Give quote (1)	Price estimate (2)	Inverse MR (3)	Lower bound (4)	Upper bound (5)
Baseline: Women vs men	0.008	9.415**	-31.232	5.253	16.606*
	(0.005)	(5.098)	(70.901)	(14.621)	(10.051)
Search: Women vs men	0.016***	0.668	29.440	-20.502	15.543
	(0.006)	(5.053)	(43.947)	(14.407)	(11.779)
Women: Search vs baseline	-0.012**	-12.897***	6.946	-22.935**	1.905
	(0.006)	(4.592)	(52.707)	(10.705)	(22.631)
Men: Search vs baseline	-0.020***	0.344	-16.869	-20.740*	24.686*
	(0.005)	(5.595)	(50.641)	(10.008)	(12.188)
Men: Income vs baseline	0.008	12.561**	-33.854	4.425	18.621
	(0.007)	(6.113)	(51.902)	(16.757)	(14.304)
Women: Uninformed vs baseline	-0.012*	-8.625**	10.223	-20.808*	9.761
	(0.007)	(6.737)	(64.496)	(10.789)	(13.487)

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: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

Alternative quote definitions

Figure A.6: Robustness: Alternative quote definitions



Notes: This figure shows average price gaps differential for baseline customers using alternative quote definitions. 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.8: Robustness: Alternative quote definitions

	(1) Prices	(2) No disc.	(3) Min P	(4) Max P	(5) Match P	(6) Invalid P	(7) Robust
Women	10.526**	10.726**	9.582**	11.474***	11.021***	11.469***	11.995***
	(4.192)	(4.202)	(4.196)	(4.218)	(4.218)	(4.251)	(4.041)
Search	0.102	0.096	-0.729	0.947	-0.265	-0.223	1.414
	(4.535)	(4.554)	(4.521)	(4.586)	(4.574)	(4.618)	(4.373)
Search ×*Women*	-11.352*	-11.073*	-9.690	-13.033**	-12.591**	-12.213*	-13.939**
	(6.311)	(6.341)	(6.299)	(6.369)	(6.369)	(6.414)	(6.080)
Income	14.236***	14.645***	13.446**	15.036***	14.506***	14.584***	18.691***
	(5.363)	(5.382)	(5.352)	(5.409)	(5.393)	(5.470)	(5.224)
Income ×*Women*	-14.025*	-14.852*	-12.948*	-14.998*	-13.681*	-15.216**	-19.116***
	(7.608)	(7.623)	(7.597)	(7.676)	(7.644)	(7.733)	(7.335)
Uninformed	-1.972	-1.928	-2.764	-1.173	-0.968	-1.205	-0.661
	(5.372)	(5.392)	(5.370)	(5.413)	(5.401)	(5.422)	(5.233)
Uninformed ×*Women*	-5.124	-5.462	-4.493	-5.787	-6.212	-4.945	-4.825
	(7.540)	(7.566)	(7.550)	(7.602)	(7.594)	(7.606)	(7.404)
Changed P	0	116	239	239	55	121	
Observations	10323	10323	10323	10323	10327	10414	10407

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. Columns (1)-(6) redefine total quote measure. Column (1) uses the main pre-specified quote definition. Column (2) excludes discounts from quotes. Column (3) and (4) use minimum and maximum quote whenever a range of prices is provided, Column (5) selects minimum price provided by other shop to a customer during that week whenever a shop offers to price match quotes, Column (6) includes invalid quotes. Column (7) uses the main quote definition in a robust regression. Robust standard errors in parenthesis, with significance levels: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

Alternative competition measures

Table A.9: Robustness: Alternative competitors definitions

	Main sample	Robust regression	Email shops
	(1)	(2)	(3)
Competitors	-0.743**	-0.706**	0.055
	(0.369)	(0.356)	(0.741)
Competitors ×Women	-1.615***	-1.410***	-3.187***
	(0.472)	(0.455)	(0.954)
Women	17.991***	17.967***	17.192***
	(4.693)	(4.498)	(4.602)
Controls	Yes	Yes	Yes
Observations	10,323	10,316	10,323

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 email address available in the YP. All specifications control for woman indicator, each customer type and their interaction with women indicator 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.10: Price gaps with alternative distance radius: Robustness

	1-km (1)	1.5-km (2)	2-km (3)	2.5-km (4)	3-km (5)
Competitors	-0.737**	-0.828***	-1.064***	-1.030***	-0.854***
	(0.369)	(0.272)	(0.220)	(0.164)	(0.129)
Competitors×Women	-1.608***	-0.978***	-0.675**	-0.334*	-0.199
	(0.472)	(0.337)	(0.268)	(0.196)	(0.154)
Women	17.975***	18.066***	17.891***	15.404***	14.128***
	(4.694)	(4.830)	(4.967)	(4.931)	(4.947)
Main controls	Yes	Yes	Yes	Yes	Yes

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: $^{*}p < 0.10, ^{**}\ p < 0.05,\ ^{***}\ p < 0.01$.

9 Are effects driven by other service offers?

Table A.11: Effect of additional part and service offers

	(1)	(2)	(3)	(4)	(5)
	No detail	Add parts	Fees	Warranty	Offers
Women	-0.014	0.000	0.005	0.009	-0.002
	(0.013)	(0.012)	(0.016)	(0.009)	(0.007)
Search	-0.003	-0.005	0.018	0.022**	0.001
	(0.014)	(0.013)	(0.017)	(0.010)	(0.007)
Search ×Women	0.010	0.000	-0.002	-0.011	0.007
	(0.020)	(0.017)	(0.024)	(0.014)	(0.010)
Income	-0.033**	-0.003	0.046**	0.002	-0.001
	(0.016)	(0.015)	(0.020)	(0.012)	(0.009)
Income ×Women	0.020	0.014	-0.026	-0.004	-0.006
	(0.023)	(0.021)	(0.029)	(0.017)	(0.012)
Uninformed	0.048***	-0.014	-0.029	-0.006	0.015
	(0.018)	(0.015)	(0.020)	(0.011)	(0.010)
Uninformed ×Women	-0 -0.042*	0.008	0.003	-0.007	0.002
	(0.024)	(0.021)	(0.028)	(0.016)	(0.014)
Controls	Yes	Yes	Yes	Yes	Yes
Av Mean	0.19	0.15	0.35	0.08	0.05

Notes: This table reports coefficients from a regression of not giving detail of total quote components (Column 1), including additional car parts in quote (radiator cap, radiator hoses, thermostat or others) (Column 2), include tax and fees in quote (Column 3), providing a warranty (Column 4), and providing additional offers such as discount, a free inspection, price-matching, and 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 in parenthesis, with significance levels: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

2215 JOSÉ MANUEL CARBÓ and SERGIO GORJÓN: Application of machine learning models and interpretability techniques to identify the determinants of the price of bitcoin.

2216 LUIS GUIROLA and MARÍA SÁNCHEZ-DOMÍNGUEZ: Childcare constraints on immigrant integration.

2217 ADRIÁN CARRO, MARC HINTERSCHWEIGER, ARZU ULUC and J. DOYNE FARMER: Heterogeneous effects and spillovers of macroprudential policy in an agent-based model of the UK housing market.

2218 STÉPHANE DUPRAZ, HERVÉ LE BIHAN and JULIEN MATHERON: Make-up strategies with finite planning horizons but forward-looking asset prices.

2219 LAURA ÁLVAREZ, MIGUEL GARCÍA-POSADA and SERGIO MAYORDOMO: Distressed firms, zombie firms and zombie lending: a taxonomy.

2220 BLANCA JIMÉNEZ-GARCÍA and JULIO RODRÍGUEZ: A quantification of the evolution of bilateral trade flows once bilateral RTAs are implemented.

2221 SALOMÓN GARCÍA: Mortgage securitization and information frictions in general equilibrium.

2222 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Accuracy of explanations of machine learning models for credit decisions.

2223 JAMES COSTAIN, GALO NUÑO and CARLOS THOMAS: The term structure of interest rates in a heterogeneous monetary union.

2224 ANTOINE BERTHEAU, EDOARDO MARIA ACABBI, CRISTINA BARCELÓ, ANDREAS GULYAS, STEFANO LOMBARDI and RAFFAELE SAGGIO: The Unequal Consequences of Job Loss across Countries.

2225 ERWAN GAUTIER, CRISTINA CONFLITTI, RIEMER P. FABER, BRIAN FABO, LUDMILA FADEJEVA, VALENTIN JOUVANCEAU, JAN-OLIVER MENZ, TERESA MESSNER, PAVLOS PETROULAS, PAU ROLDAN-BLANCO, FABIO RUMLER, SERGIO SANTORO, ELISABETH WIELAND and HÉLÈNE ZIMMER. New facts on consumer price rigidity in the euro area.

2226 MARIO BAJO and EMILIO RODRÍGUEZ: Integrating the carbon footprint into the construction of corporate bond portfolios.

2227 FEDERICO CARRIL-CACCIA, JORDI PANIAGUA and MARTA SUÁREZ-VARELA: Forced migration and food crises.

2228 CARLOS MORENO PÉREZ and MARCO MINOZZO: Natural Language Processing and Financial Markets: Semi-supervised Modelling of Coronavirus and Economic News.

2229 CARLOS MORENO PÉREZ and MARCO MINOZZO: Monetary Policy Uncertainty in Mexico: An Unsupervised Approach.

2230 ADRIAN CARRO: Could Spain be less different? Exploring the effects of macroprudential policy on the house price cycle.

2231 DANIEL SANTABÁRBARA and MARTA SUÁREZ-VARELA: Carbon pricing and inflation volatility.

2232 MARINA DIAKONOVA, LUIS MOLINA, HANNES MUELLER, JAVIER J. PÉREZ and CRISTOPHER RAUH: The information content of conflict, social unrest and policy uncertainty measures for macroeconomic forecasting.

2233 JULIAN DI GIOVANNI, MANUEL GARCÍA-SANTANA, PRIIT JEENAS, ENRIQUE MORAL-BENITO and JOSEP PIJOAN-MAS: Buy Big or Buy Small? Procurement Policies, Firms' Financing and the Macroeconomy*.

2234 PETER PAZ: Bank capitalization heterogeneity and monetary policy.

2235 ERIK ANDRES-ESCAYOLA, CORINNA GHIRELLI, LUIS MOLINA, JAVIER J. PÉREZ and ELENA VIDAL: Using newspapers for textual indicators: which and how many?

2236 MARÍA ALEJANDRA AMADO: Macroprudential FX regulations: sacrificing small firms for stability?

2237 LUIS GUIROLA and GONZALO RIVERO: Polarization contaminates the link with partisan and independent institutions: evidence from 138 cabinet shifts.

2238 MIGUEL DURO, GERMÁN LÓPEZ-ESPINOSA, SERGIO MAYORDOMO, GAIZKA ORMAZABAL and MARÍA RODRÍGUEZ-MORENO: Enforcing mandatory reporting on private firms: the role of banks.

2239 LUIS J. ÁLVAREZ and FLORENS ODENDAHL: Data outliers and Bayesian VARs in the Euro Area.

2240 CARLOS MORENO PÉREZ and MARCO MINOZZO: "Making text talk": The minutes of the Central Bank of Brazil and the real economy.

2241 JULIO GÁLVEZ and GONZALO PAZ-PARDO: Richer earnings dynamics, consumption and portfolio choice over the life cycle.

2242 MARINA DIAKONOVA, CORINNA GHIRELLI, LUIS MOLINA and JAVIER J. PÉREZ: The economic impact of conflict-related and policy uncertainty shocks: the case of Russia.

2243 CARMEN BROTO, LUIS FERNÁNDEZ LAFUERZA and MARIYA MELNYCHUK: Do buffer requirements for European systemically important banks make them less systemic?

2244 GERGELY GANICS and MARÍA RODRÍGUEZ-MORENO: A house price-at-risk model to monitor the downside risk for the Spanish housing market.

2245 JOSÉ E. GUTIÉRREZ and LUIS FERNÁNDEZ LAFUERZA: Credit line runs and bank risk management: evidence from the disclosure of stress test results.
2301 MARÍA BRU MUÑOZ: The forgotten lender: the role of multilateral lenders in sovereign debt and default.
2302 SILVIA ALBRIZIO, BEATRIZ GONZÁLEZ and DMITRY KHAMETSHIN: A tale of two margins: monetary policy and capital misallocation.
2303 JUAN EQUIZA, RICARDO GIMENO, ANTONIO MORENO and CARLOS THOMAS: Evaluating central bank asset purchases in a term structure model with a forward-looking supply factor.
2304 PABLO BURRIEL, IVÁN KATARYNIUK, CARLOS MORENO PÉREZ and FRANCESCA VIANI: New supply bottlenecks index based on newspaper data.
2305 ALEJANDRO FERNÁNDEZ-CEREZO, ENRIQUE MORAL-BENITO and JAVIER QUINTANA: A production network model for the Spanish economy with an application to the impact of NGEU funds.
2306 MONICA MARTINEZ-BRAVO and CARLOS SANZ: Trust and accountability in times of pandemic.
2307 NATALIA FABRA, EDUARDO GUTIÉRREZ, AITOR LACUESTA and ROBERTO RAMOS: Do renewable energies create local jobs?
2308 ISABEL ARGIMÓN and IRENE ROIBÁS: Debt overhang, credit demand and financial conditions.
2309 JOSÉ-ELÍAS GALLEGOS: Inflation persistence, noisy information and the Phillips curve.
2310 ANDRÉS ALONSO-ROBISCO, JOSÉ MANUEL CARBÓ and JOSÉ MANUEL MARQUÉS: Machine Learning methods in climate finance: a systematic review.
2311 ALESSANDRO PERI, OMAR RACHEDI and IACOPO VAROTTO: The public investment multiplier in a production network.
2312 JUAN S. MORA-SANGUINETTI, JAVIER QUINTANA, ISABEL SOLER and ROK SPRUK: Sector-level economic effects of regulatory complexity: evidence from Spain.
2313 CORINNA GHIRELLI, ENKELEJDA HAVARI, ELENA MERONI and STEFANO VERZILLO: The long-term causal effects of winning an ERC grant.
2314 ALFREDO GARCÍA-HIERNAUX, MARÍA T. GONZÁLEZ-PÉREZ and DAVID E. GUERRERO: How to measure inflation volatility. A note.
2315 NICOLÁS ABBATE, INÉS BERNIELL, JOAQUÍN COLEFF, LUIS LAGUINGE, MARGARITA MACHELETT, MARIANA MARCHIONNI, JULIÁN PEDRAZZI and MARÍA FLORENCIA PINTO: Discrimination against gay and transgender people in Latin America: a correspondence study in the rental housing market.
2316 SALOMÓN GARCÍA: The amplification effects of adverse selection in mortgage credit suply.
2317 METTE EJRNÆS, ESTEBAN GARCÍA-MIRALLES, METTE GØRTZ and PETTER LUNDBORG: When death was postponed: the effect of HIV medication on work, savings and marriage.
2318 GABRIEL JIMÉNEZ, LUC LAEVEN, DAVID MARTÍNEZ-MIERA and JOSÉ-LUIS PEYDRÓ: Public guarantees and private banks' incentives: evidence from the COVID-19 crisis.
2319 HERVÉ LE BIHAN, DANILO LEIVA-LEÓN and MATÍAS PACCE: Underlying inflation and asymmetric risks.
2320 JUAN S. MORA-SANGUINETTI, LAURA HOSPIDO and ANDRÉS ATIENZA-MAESO: The numbers of equality regulation. Quantifying regulatory activity on non-discrimination and its relationship with gender gaps in the labour market.
2321 ANDRES ALONSO-ROBISCO and JOSÉ MANUEL CARBÓ: Analysis of CBDC Narrative of Central Banks using Large Language Models.
2322 STEFANIA ALBANESI, ANTÓNIO DIAS DA SILVA, JUAN F. JIMENO, ANA LAMO and ALENA WABITSCH: New technologies and jobs in Europe.
2323 JOSÉ E. GUTIÉRREZ: Optimal regulation of credit lines.
2324 MERCEDES DE LUIS, EMILIO RODRÍGUEZ and DIEGO TORRES: Machine learning applied to active fixed-income portfolio management: a Lasso logit approach.
2325 SELVA BAHAR BAZIKI, MARÍA J. NIETO and RIMA TURK-ARISS: Sovereign portfolio composition and bank risk: the case of European banks.
2326 ANGEL-IVAN MORENO and TERESA CAMINERO: Assessing the data challenges of climate-related disclosures in european banks. A text mining study.
2327 JULIO GÁLVEZ: Household portfolio choices under (non-)linear income risk: an empirical framework.
2328 NATASCHA HINTERLANG: Effects of Carbon Pricing in Germany and Spain: An Assessment with EMuSe.
2329 RODOLFO CAMPOS, SAMUEL PIENKNAGURA and JACOPO TIMINI: How far has globalization gone? A tale of two regions.
2330 NICOLÁS FORTEZA and SANDRA GARCÍA-URIBE: A Score Function to Prioritize Editing in Household Survey Data: A Machine Learning Approach.
2331 PATRICK MACNAMARA, MYROSLAV PIDKUYKO and RAFFAELE ROSSI: Taxing consumption in unequal economies.
2332 ESTHER CÁCERES and MATÍAS LAMAS: Dividend Restrictions and Search for Income.
2333 MARGARITA MACHELETT: Gender price gaps and competition: Evidence from a correspondence study.