



working paper

eeeecon
[triple:e:con]

Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment

Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog

Working Papers in Economics and Statistics

2019-11

University of Innsbruck
<https://www.uibk.ac.at/eeeecon/>



University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform "Empirical and Experimental Economics"
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
<https://www.uibk.ac.at/eeecon/wopec/>

For a list of recent papers see the backpages of this paper.

Uncovering sophisticated discrimination with the help of credence goods markups – evidence from a natural field experiment¹

Jonathan Hall^a, Rudolf Kerschbamer^{2b}, Daniel Neururer^b and Eric Skoog^a

^a Uppsala University

^b University of Innsbruck

Abstract

We present the results of a pre-registered natural field experiment designed to uncover a sophisticated form of discrimination against an immigrant minority in a market for credence goods. For this purpose, we introduce two markups: (i) the credence goods markup defined as the difference between the price paid by the same person for an ordinary service and an otherwise equivalent credence goods service; and (ii) the discriminatory markup defined as the difference between the price paid by a member of an immigrant minority group and the price paid by a member of the majority group for the same kind of service. We document the existence of a large credence goods markup of about 40%, on average. Moreover, we find a sizeable discriminatory markup for the credence goods service but no discriminatory markup for the ordinary service. The results of an ex-post survey suggest that this sophisticated form of discrimination is mainly due to the prejudicial behavior of sellers belonging to an established local ethnic minority group towards buyers belonging to a low-status immigrant ethnic minority group.

Keywords: natural field experiment, credence goods, discrimination, expert services, credence goods markup, discriminatory markup

JEL classification: C93, D82, J15

June 14, 2019

¹ We benefitted a lot from comments by Marco Schwarz and our fieldwork coordinator Dogukan Cansis Karakus. Financial support from the Austrian Science Fund (FWF) through special research area grant SFB F63, as well as through grant numbers P26901 and P27912 is gratefully acknowledged.

² Corresponding Author. Address: University of Innsbruck, Department of Economics, Universitätsstraße 15, 6020 Innsbruck, Austria. E-mail: Rudolf.Kerschbamer@uibk.ac.at

1. Introduction

Credence goods markets are characterized by profound information asymmetries between customers lacking the expertise required to assess which product or service best fits their needs and expert sellers possessing the required expertise. Important examples are, among others, the market for health care services where the doctor is better informed than the patient about the appropriate treatment given the symptoms of the patient, the market for car repair services where the mechanic is better informed than the car owner about the appropriate repair, and the market for financial advice where the adviser can better assess which product fits the customer's needs.

The information asymmetries in markets for credence goods open the door to different forms of fraud and misbehavior on the experts' side.³ In this paper we address the question whether fraud interacts in a systematic way with discrimination in markets for credence goods. To address this question we conduct a natural field experiment in the Turkish market for cellphone repairs.⁴ Our main instrument is the manipulation of the seller's perception of whether the service under consideration is an ordinary good or a credence good service. This is done via the following treatment manipulation: In the ordinary goods environment the mystery shopper enters a repair store with a broken device and asks to change the specific part that causes the problem; in the credence goods context the customer enters a repair store with a broken device and asks for a repair, mentioning that he has no idea which kind of repair is needed to fix the problem. By comparing the average price across the two goods characteristics (keeping the customer type constant) we compute the additional sum the consumer has to pay when he is uninformed about the appropriate repair. We term this additional sum the "credence goods markup", because it represents the amount that is arguably exclusively due to the fact that the consumer has less information about the appropriate repair than the seller.

In a second treatment dimension, we employ mystery shoppers with two different ethnic backgrounds: half of our mystery shoppers come from the Turkish majority, the other half come from the Syrian immigrant minority. By comparing the average price across the

³ According to Darby and Karni (1973) and Dulleck and Kerschbamer (2006) the main fraud dimensions in markets for credence goods are (i) *overprovision* – the customer gets a more expensive treatment or service than he actually needs; (ii) *underprovision* – the customer receives a treatment or service that does not solve his problems; and (iii) *overcharging* – the customer is charged for a more expensive service than was actually provided.

⁴ The term 'natural field experiment' has been introduced by Harrison and List (2004). They distinguish between three variants of field experiments – an artefactual field experiment is the same as a conventional lab experiment but with a non-standard subject pool; a framed field experiment is the same as an artefactual field experiment but with field context added in some important parts of the instructions; and a natural field experiment is the same as a framed field experiment but where the subjects do not know that they are in an experiment.

two consumer groups (keeping the service type constant) we compute the additional sum a consumer has to pay when he is a member of the immigrant minority. We term this additional sum the “discriminatory markup”, because it represents the amount that is arguably exclusively due to the fact that the customer is a member of the immigrant minority and not a member of the local majority.

Our main hypothesis is that there is a non-trivial interaction between the credence goods markup and the discrimination markup. Specifically, our pre-registered ex ante hypothesis consists of three parts: (i) there is a positive credence goods markup, on average (over the two customer types); (ii) there is a positive discriminatory markup, on average (over the two service types); and (iii) there is a higher discriminatory markup for the credence goods service than for the ordinary service (or, equivalently, there is a higher credence goods markup for the minority than for the majority). The three parts of our ex ante hypothesis are based on the following considerations:

- In theory, repair services are credence goods for customers who are unable to self-diagnose their problem but ordinary goods for consumers who know what they need (see Dulleck and Kerschbamer 2006). If sellers exploit the asymmetric information, then prices are higher, on average, in the credence goods context than in the ordinary goods context resulting in a positive credence goods markup, on average.
- Based on the findings in the large and still growing literature on discrimination (we summarize parts of this literature below) we would expect that customers that are considered as “in-group” by a given seller are treated better than customers that are considered as “out-group”. If this is the case and if the discrimination takes place via the price, then it results in a positive discriminatory markup, on average.
- Following the logic of Zitzewitz (2012) discriminatory behavior is something that expert sellers would prefer to conceal rather than reveal. Since concealment is easier in the credence goods context where sellers can hide behind the preexisting information asymmetries than in the ordinary market context where customers can more easily assess which price is appropriate, sellers arguably have less concerns to discriminate in the credence goods context. If sellers act in accordance with these incentives then their behavior results in a higher discriminatory markup for the credence goods service than for the ordinary service – or, equivalently, a higher credence goods markup for the minority than for the majority.

Our data strongly support the existence of a large credence goods markup – it amounts to about 40% of the price of the ordinary good, on average. The size of the discriminatory

markup is much smaller – it amounts to about 20% of the price paid by a member of the majority. Turning to our main hypothesis – the interaction between credence goods markup and discriminatory markup – we uncover a sizeable discriminatory markup for the credence goods service but no discriminatory markup for the ordinary service. Together these findings point to the existence of a sophisticated form of discrimination in the market under consideration – a discrimination that takes place almost exclusively in the market segment where sellers can hide behind the preexisting information asymmetry. As a consequence, this form of discrimination is difficult to detect by the party affected and also difficult to uncover empirically. Our ex-post survey reveals that our discrimination result is mainly driven by the discriminatory behavior of sellers belonging to the Kurdish ethnic minority in Turkey towards buyers belonging to the Syrian immigrant minority.

According to Zitzewitz (2012) natural field experiments are a promising tool to uncover behavior that agents would prefer to conceal and by now, misbehavior in credence goods markets is well documented with the help of this instrument: The pioneering natural field study by Schneider (2012) investigates the impact of reputational concerns on fraudulent behavior in the US market for car repairs and reports that although the overall level of undertreatment and overtreatment is pervasive, reputation does not help much to improve efficiency. The studies by Balafoutas et al. (2013) and Balafoutas et al. (2017) are conducted in the Greek market for taxi rides. Balafoutas et al. (2013) investigate the impact of the perceived information asymmetry between the driver and the passenger on the size of the different fraud dimensions and show that passengers who are perceived as having inferior information about optimal routes are taken on longer detours (which corresponds to overtreatment) while lack of information on the local tariff system leads to manipulated bills (which corresponds to overcharging). Balafoutas et al. (2017) focus on the perceived incentive of the passenger to contain the expenses and find that passengers who explicitly state that their expenses will be reimbursed by their employer are significantly more likely to be the victim of overcharging.

The studies by Lu (2014) and Kerschbamer et al. (2016) point in a similar direction as Balafoutas et al. (2017) although the experiments are conducted in completely different markets. Lu (2014) finds that doctors act out of self-interest by prescribing unnecessary or excessively expensive drugs to insured patients. Similarly, Kerschbamer et al. (2016) show that the repair price of computers increases by 80 % if the customer reveals to the expert that an insurance company will cover the cost of the repair. The study of Anagol et al. (2017) examines the Indian market for life insurance and shows that life insurance agents react to

material incentives by overwhelmingly recommending unsuitable, strictly dominated products that provide high commissions to them. Finally, Gottschalk et al. (2018) study the market for dental care and report significantly less overtreatment recommendations for uninformed patients with higher socio-economic status.

All of the above studies have in common that they examine the effects of different treatment variations (i.e., suggesting the chance for repeated business, varying the information level of the customer, mentioning that the bill will be reimbursed by a 3rd party, ...) on the level of the different fraud dimensions within a credence goods market context. The experimental design in the present paper differs from previous studies in that our treatment manipulation changes the market context from an ordinary good to a credence good context. This enables us to compare experts' provision and charging behavior across these two market environments and to calculate the resulting credence goods markup. We consider the introduction of the credence goods markup as a major methodical innovation that can potentially be used to address a variety of research questions beyond the discrimination context studied here. As a measure of the amount a customer can save by investing time and effort to self-diagnose the problem, it could be used, for instance, in cross country studies gauging levels of trust and trustworthiness in society – acting as a substitute for the usual survey questions or experimental measures of trustworthiness (like the average back transfer in the investment game introduced by Berg et al. 1995).

Our idea of comparing the average repair price in a credence goods market to the average repair price in an ordinary goods market (keeping the type of service constant) bears some similarity to one of the treatment variations in Balafoutas et al. (2013). In their experiment on the cheating behavior of taxi drivers in Athens the authors let some passengers (the “locals”) only state the destination and others (the “nonlocal-natives”) state the destination and ask the driver whether he knows the destination, adding as an explanation for asking that they are not familiar with the city. In our interpretation the treatment with the “locals” is somehow in between an ordinary goods and a credence goods environment: Whether the ride is considered by the driver as a credence good or an ordinary good arguably depends on the beliefs of the taxi driver – it is plausible that the driver believes that a passenger without an accent is a local who knows the city very well and would detect a detour immediately (in this case the ride would be an ordinary good), but it is equally plausible that the taxi driver believes that the passenger has less than perfect knowledge about the optimal route and would not detect a short detour (in this case the ride would be more of a credence

good). Our treatment manipulation is much cleaner in this regard by leaving almost no room for speculation on the side of the expert service provider.

On the dimension of discrimination, there is a huge and still growing body of work – see the excellent review by Bertrand and Duflo (2017). Within this literature our paper is most closely related to field studies examining discrimination in product and service markets: The study by Ayres and Siegelmann (1995) is an early example of an audit study investigating race and gender discrimination in bargaining for a new car. The authors find that white males are quoted lower prices than white women and blacks (men or women), but their data do not allow one to discriminate clearly between the theories of taste-based and statistical discrimination.⁵ List (2004) examines the behavior of buyers and sellers in a sports cards market and documents discrimination against minorities (women and black males) relative to the majority (white males). By combining his field data with results from a dictator game, the author provides convincing evidence in support of statistical discrimination. Doleac and Stein (2013) examine racial discrimination in online marketplaces for iPods. The authors report that an iPod presented by a dark hand receives, on average, fewer and lower offers than the same product presented by a light-skinned hand. The fact that this effect is stronger in thin markets is in line with Becker’s hypothesis that discrimination can be competed away. In a similar fashion Nunley et al. (2013) study racial discrimination by simultaneously selling identical products on eBay under different racially identifying names. The authors detect a discriminatory pattern that is in line with in-group favoritism (i.e., white names receive higher bids for distinctively white products and black names receive higher bids for distinctively black products). They also find that the price differences disappear when sellers accumulate reviews and argue that this points in the direction of statistical discrimination.

Zussman (2013) conducts a field experiment with a follow-up survey to examine racial discrimination in the Israeli online market for used cars. The author finds discrimination against Arab buyers and sellers when it comes to different measures of response rates and argues that the discrimination is most likely motivated by statistical rather than taste considerations because the behavior of members of the majority towards members of the majority and minority are determined by perceptions regarding aggregate differences across groups in transaction-relevant characteristics. The study of Castillo et al. (2013) examines gender differences in bargaining outcomes in a market for taxi rides. The authors show that

⁵ These are the two leading explanations for discrimination against minorities in markets. According to the taste-based theory – due to Becker (1957) – discriminatory behavior is driven by personal prejudice against individuals of a given gender, race or ethnicity resulting in a willingness to pay for avoiding an interaction with such an individual. By contrast, the statistical theory – due to Arrow (1973) and Phelps (1972) – argues that discriminatory behavior is the result of profit- or utility-maximization as it is based on actual or perceived differences across groups in some transaction-relevant characteristics.

male passengers are treated as inferior by male drivers and they interpret this finding as evidence for statistical discrimination because it is driven by perceived higher valuations of male customers.

All of the above summarized studies have in common that they examine potential discrimination in ordinary product or service markets. Even in the study of Castillo et al. (2013) – investigating the market for taxi rides in Lima, Peru – the good can hardly be interpreted as a credence good because taxi rides are the most common means of transportation for households in Lima (implying that most customers know the shortest route to their destination quite well) and because bargaining over the fare for the ride takes place before the ride is actually taken (implying that the typical credence goods problem in the market for taxi rides – overtreatment in the form of taking detours and overcharging in the form of charging for items that are included in the fare – are not an issue in this market). Our study complements and extends previous work on discrimination by examining a credence goods context in addition to the ordinary goods context. As we have seen, this is important because it enables us to detect the sophisticated form of discrimination by expert providers uncovered in the present paper which would be hard – if not impossible – to detect otherwise.

The rest of the paper is organized as follows: Section 2 presents our experimental design and motivates the four treatments constituting a 2x2 factorial design. Our main results are presented and discussed in Section 3. Section 4 concludes the paper.

2. Experimental design and identification strategy

We conducted our natural field experiment in the Turkish city of Antalya from July 2018 until October 2018. Antalya is a major tourism hub located on Anatolia's southwestern coast with about one million inhabitants. At the latest council election in March 2019, the opposition party CHP was the winner in Antalya and this reflects the rather liberal attitudes of the people living in Antalya.

Four male undercover mystery shoppers – two Turks and two Syrians – were hired and trained by the fieldwork coordinator who was located in Antalya during the whole experiment.⁶ Choosing Syrians as minority shoppers allows us to assess discrimination towards a highly relevant and growing ethnic minority in Turkey. Globally, the largest stock of refugees are of Syrian origin (McAuliffe and Ruhs 2017). Turkey hosts more Syrian refugees, and indeed more refugees overall, than any other country in the world. Following

⁶ We check for the potential influence of personal characteristics of the mystery shoppers and find no significant difference in repair prices within the group of Turkish mystery shoppers and no significant difference within the group of Syrian mystery shoppers.

the recent influx of refugees, Turkish cities have witnessed an increase in intercommunal violence between Syrians and members of the majority community (International Crisis Group 2018). Hence, Syrians are a particularly salient ethnic minority in the Turkish context, and therefore arguably the target of negative sentiments and discrimination.⁷

The shops for our natural field experiment were selected as follows: We first compiled a list of all repair shops in the city of Antalya (without its surrounding suburbs) using information from exploratory tours in the streets of Antalya and available online (Google, city directory, etc.). Then we assigned to each shop a specific number and used a random number generator to match the shops to our four treatments (described below) and finally to the undercover shoppers.

Before the start of the experiment, we bought 24 identical, refurbished and perfectly working smartphones (Apple iPhone 5s) for an average price of 900 TRY.⁸ The smartphones were manipulated as follows (and the manipulation was repeated after every repair shop visit): First, we drained the battery of the phone until the phone switched off. Then we mechanically destroyed the charging port of the phone which is used to charge the battery. As a result, the mobile is off and it is not possible to switch it on anymore because the battery is empty and it can't be charged anymore. A defective charging port is not an unusual problem; rather, it happens relatively often (<https://tech.co/phones/iphone-troubleshooting-fix-common-problems#iphone-wont-charge>, accessed on the 5th of June 2019). The correct repair for this problem is to replace the charging port and charge the phone. With the help of internet tutorials and DIY repair kits, the problem could be easily self-diagnosed and self-repaired (<https://www.ifixit.com/Guide/iPhone+5s+Lightning+Connector+Replacement/20261>, accessed on the 5th of June 2019).

Because of the trivial nature of the problem we expected that almost all shops would be able to solve the problem. For our design it is essential that the problem can be identified and repaired easily by almost all mobile repair shops as we are interested in examining intentional misbehavior and not incompetence of the service provider. The fact that only about 2.5 % of the visited repair shops claim that they are not able to identify the problem confirms that we succeeded in this dimension.

The interaction between undercover customers and computer repair shops was done by the undercover shoppers themselves after a training session. The four treatments displayed in Table 1 constitute our 2x2 factorial design:

⁷ This is also confirmed in our ex-post survey – see Appendix B.

⁸ The exchange rate was 10 TRY = 1.55 EURO or 10 TRY = 1.75 USD on the 5th of June 2019.

- **MAJORITY-ORDINARY:** In this treatment (abbreviated **MAJ-ORD**), mystery shoppers from Turkey stick to the following script when dropping off the cellphone at the repair shop: *“Hi! I can’t switch on my mobile anymore and I know that a defective charging port causes the problem. Could you please repair it?”* Here the customer reveals by the script that he knows the cause of the problem and as a result the transaction arguably involves an ordinary good and not a credence good. The ethnic origin of the mystery shopper is easily identified by the shop staff based on the accent.⁹ Since in this treatment the mystery shopper and the shop owner are both from Turkey, discrimination is expected to play a minor role in this treatment.
- **MAJORITY-CREDENCE:** In this treatment (abbreviated **MAJ-CRE**), mystery shoppers from Turkey stick to the following script when dropping off the cellphone at the repair shop: *“Hi! I can’t switch on my mobile anymore and I don’t know what the problem is. Could you please repair it?”* Here the mystery shopper gives the impression that he does not know the cause of the problem and as a result the transaction involves a credence good and not an ordinary good as in MAJ-ORD. As in MAJ-ORD, ethnic discrimination is expected to play a minor role in this treatment as the shop owner and the mystery shoppers are both from Turkey.
- **MINORITY-ORDINARY:** In this treatment (abbreviated **MIN-ORD**), the script is identical to that in MAJ-ORD but the mystery shopper is now from Syria and not from Turkey. As a result, ethnic discrimination could play a role in this treatment as the shop owners can infer the ethnic origin of the mystery shopper from his accent. On the other hand, the room for discrimination is arguably not large because discriminatory behavior is difficult to hide in an ordinary goods market.
- **MINORITY-CREDENCE:** In this treatment (abbreviated **MIN-CRE**), the script is identical to that in MAJ-CRE, but mystery shoppers are again from Syria – as in MIN-ORD. Compared to MAJ-CRE prices are expected to be higher, on average, because discrimination is predicted to play a role in MIN-CRE but not in MAJ-CRE. Compared to MIN-ORD the room for discrimination is expected to be larger because

⁹ This is also true the other way around and for the sake of experimental control we decided to stop the interaction of the mystery shopper and the repair shop in case the mystery shopper realized that the shop staff is not a Turkish citizen. This seemed necessary because otherwise the distinction between immigrant and native in the expert-customer relationship could be blurred or even reversed. The resulting cancellations led to a slight deviation from our preregistration plan with respect to the number of observations.

discriminatory behavior can easily be hidden behind the pre-existing information asymmetry regarding the appropriate repair.

Table 1: treatment design

customer type	service type	
	ORDINARY	CREDENCE
MAJORITY	<i>MAJ-ORD</i> majority customer ordinary goods market	<i>MAJ-CRE</i> majority customer credence goods market
MINORITY	<i>MIN-ORD</i> minority customer ordinary goods market	<i>MIN-CRE</i> minority customer credence goods market

Based on our four treatments we employ the following identification strategy:

- By comparing the average repair price across the two **ORD** treatments with the average repair price across the two **CRE** treatments we identify the average credence goods markup across the two customer types.
- By comparing the average repair price across the two **MAJ** treatments with the average repair price across the two **MIN** treatments we identify the average discriminatory markup across the two service types.
- By comparing the difference in the repair price between ***MIN-ORD*** and ***MAJ-ORD*** to the difference in the repair price between ***MIN-CRE*** and ***MAJ-CRE*** we address the question of whether there is a higher discriminatory markup for the credence good service than for the ordinary service.
- And by comparing the difference in the repair price between ***MAJ-CRE*** and ***MAJ-ORD*** to the difference in the repair price between ***MIN-CRE*** and ***MIN-ORD*** we address the question of whether there is a higher credence goods markup for the minority than for the majority.¹⁰

In addition to the treatments described above we also implemented an ex-post survey where we elicit characteristics and attitudes of the repair shops – see the results in subsection 3.3.

¹⁰ While in theory a higher discriminatory markup for the credence good than for the ordinary good implies a higher credence goods markup for the minority than for the majority (and vice versa), in the empirical analysis it could well be the case that one of the difference-in-differences is significantly positive while the other is not.

All the design features presented in this section have been preregistered prior to the data collection at <https://aspredicted.org> – see Appendix A for the time stamped preregistration document.

3. Results

In total, we visited 163 shops, 40 for **MAJ-ORD**, 43 for **MAJ-CRE**, 39 for **MIN-ORD** and 41 for **MIN-CRE**. Out of these shops, two shops in **MAJ-ORD** refused the repair with the excuse that they did not have enough free capacities to perform it, and four shops in **MAJ-CRE** were not able to identify the source of the problem. Since those six shops did not provide a price quote for a repair we exclude them from our database. This leaves us with 157 shops for the analysis.¹¹ In the next subsection we quantify the credence goods markup and the discriminatory markup. Then we examine the interaction between these two markups and in the last part of this section we analyze the results from the ex-post survey.

3.1 The credence goods markup and the discriminatory markup

Figure 1 and Table 2 present the average repair price for each treatment and Figure 2 displays the cumulative distribution function of the repair prices per treatment. The average repair price is 95 TRY in **MAJ-ORD**, 128 TRY in **MAJ-CRE**, 103 TRY in **MIN-ORD** and 158 TRY in **MIN-CRE**. To quantify the average credence goods markup across the two customer types we compare the average repair price of the two **ORD** treatments (99 TRY) to the average repair price of the two **CRE** treatments (143 TRY). This difference is economically large (44 TRY, or more than 40% of the average repair price in the **ORD** treatments) and statistically highly significant (2-sided T-test, $p=0.0000$). This finding is summarized in our first result:

Result 1 (credence goods markup): *The difference between the average repair price of the ordinary good and the average repair price of the credence good is economically large and statistically highly significant. The average credence goods markup across the two consumer types amounts to more than 40% of the price of the ordinary good.*

¹¹ Out of these 157 observations, 152 shops repaired the mobile successfully and as a consequence the repair price was paid to the shops. The remaining five shops (one in **MAJ-ORD**, two in **MIN-ORD** and two in **MIN-CRE**) are also included in our analysis because they still provided a diagnosis and stated a repair quote. The transactions with these shops were not completed because they did not repair the mobile within the agreed time period (three observations) or stated a price quote so high that it hurt the sentiments of our Syrian experimenters (two observations).

To quantify the average discriminatory markup across the two service types we compare the average repair price of the two **MAJ** treatments (112 TRY) to the average repair price of the two **MIN** treatments (131 TRY). This difference is (with 19 TRY, or about 20% of the average repair price in the **MAJ** treatments) much smaller than the difference between the two service types and it is statistically only mildly significant (2-sided T-test, $p=0.0670$). This finding is summarized in our second result:

Result 2 (discriminatory markup): *The difference between the average repair price members of the minority were charged and the average repair price members of the majority were charged is moderate and statistically only mildly significant. The average discriminatory markup across the two service categories amounts to about 20% of the price paid by a member of the majority.*

Figure 1: average repair price (in TRY)

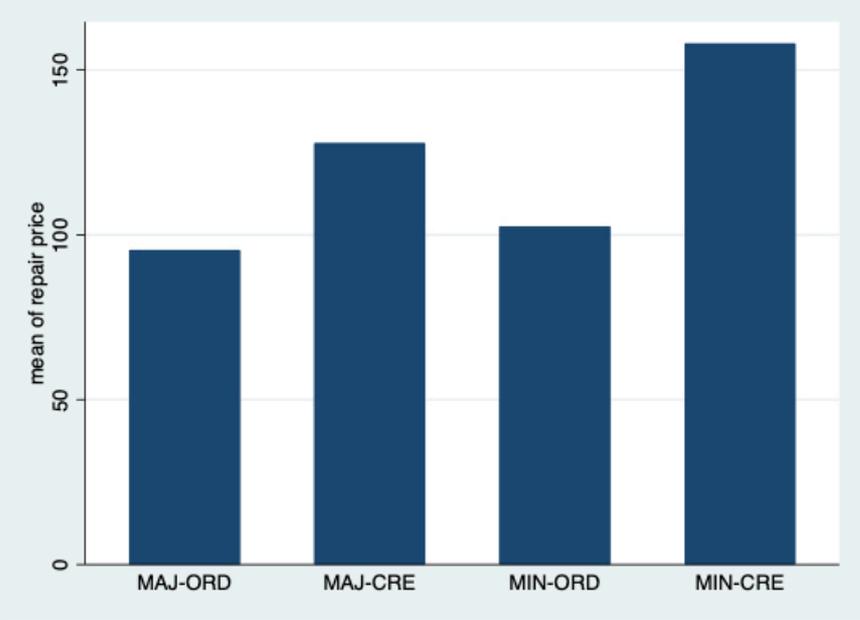
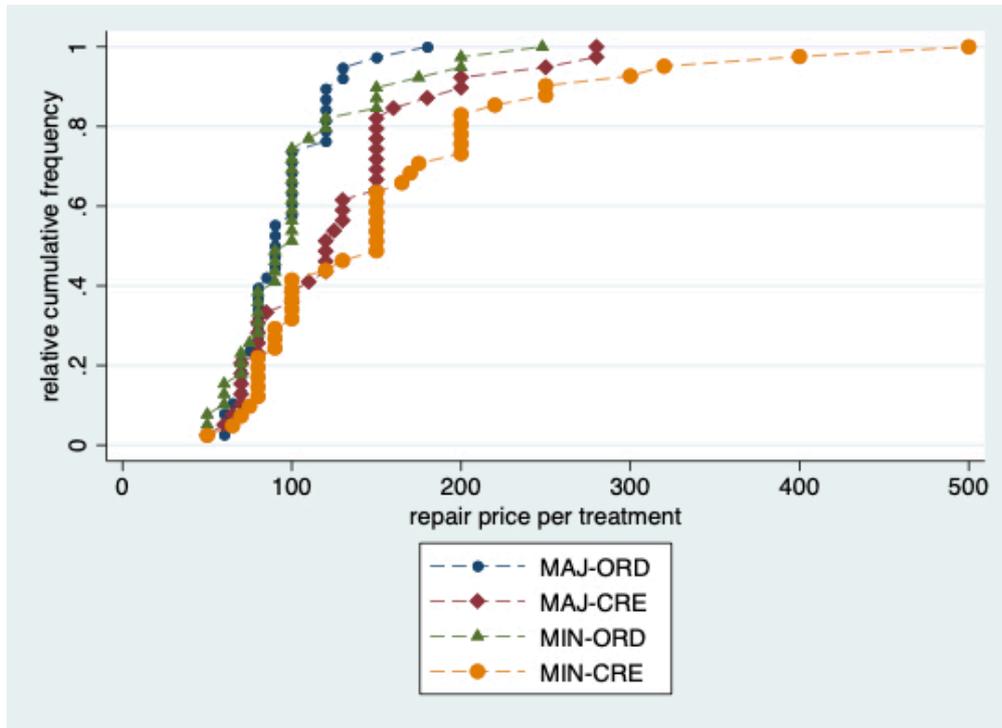


Table 2: average repair price (in TRY) and number of observations (in parentheses)

	ORDINARY	CREDENCE	both service types
MAJORITY	95 (38)	128 (39)	112 (77)
MINORITY	103 (39)	158 (41)	131 (80)
both customer types	99 (77)	143 (80)	122 (157)

Figure 2: Cumulative repair prices per treatment



3.2 Detecting sophisticated discrimination

Now, we examine our main hypothesis – the interaction effect between the credence goods markup and the discriminatory markup. For this purpose, we address two related questions: first, whether the discriminatory markup is higher for the credence goods service than for the ordinary service; and secondly, whether the credence goods markup is higher for the minority

than for the majority. Table 3 presents an OLS-regression with the repair price as the dependent variable and dummies for the treatments **MAJ-CRE**, **MIN-ORD** and **MIN-CRE** as explanatory variables – keeping **MAJ-ORD** as the (omitted) benchmark. In the first line of the table we see that **MAJ-CRE** increases repair prices significantly. This confirms and extends **Result 1** from the previous subsection; not only on average across the two consumer types, but also for the group of majority customers in isolation, the credence goods markup is (with about 30 TRY) large and statistically highly significant. From the fact that the dummy for **MIN-ORD** is not significant we conclude that there is no evidence in our data for a positive discriminatory markup in the ordinary service context. Finally, the dummy for **MIN-CRE** increases repair prices significantly. A pairwise comparison of the regression coefficients reveals that the effect of **MIN-CRE** (62.65 TRY) is almost twice as high as the effect of **MAJ-CRE** (32.43) and this difference is statistically significant ($p=0.029$). In sum, this means that we find evidence for a discriminatory markup in the credence goods context, but no evidence for a discriminatory markup in the ordinary service context; and we find a large credence goods markup for both consumer groups and an almost twice as large credence goods markup for members of the minority than for members of the majority. These findings are completely in line with our pre-registered ex ante hypothesis and they confirm our conjecture that discrimination is by far more important in markets where sellers can hide their discriminatory behavior behind the pre-existing information asymmetry. These findings are summarized in our third result:

Result 3 (interaction between discriminatory markup and credence goods markup):
Disentangling the discriminatory markup across the two service categories we find an economically large and statistically significant discriminatory markup for the credence goods service but no significant discriminatory markup for the ordinary goods service. Disentangling the credence goods markup across the two consumer types we find a large credence goods markup for both consumer groups and an almost twice as large credence goods markup for members of the minority than for members of the majority.

Table 3. Regression analysis of repair prices

	[1]
Dependent variable (OLS regressions)	<i>Repair price</i>
Independent variables	<i>(in TRY)</i>
MAJ-CRE treatment (1=yes)	32.43** (13.96)
MIN-ORD treatment (1=yes)	7.12 (13.96)
MIN-CRE treatment (1=yes)	62.65*** (13.79)
Constant	95.40*** (9.94)
# Observations	157

OLS-regressions with repair price (in TRY) as dependent variable, including, as explanatory variables, a dummy for **MAJ-CRE**, a dummy for **MIN-ORD** and a dummy for **MIN-CRE** — keeping the **MAJ-ORD** treatment as the (omitted) benchmark. ***, **, * denote significance at the 1%, 5%, 10% level, standard errors in parentheses.

3.3 Ex-post survey

After completing the data collection for our main experiment, we visited each repair shop in our sample a second time – this time asking the person in the shop to participate in a survey. The second visit was made solely by Turkish helpers and by a different person than the first visit. The fact that we had already visited the shop for our data collection was of course not mentioned. The ex-post survey was conducted to elicit general characteristics of the shop staff and their attitudes towards immigration. Most of the questions on our survey are from the migration module of the European Social Survey (<https://www.europeansocialsurvey.org>, accessed on the 5th of June 2019).

Unfortunately, only about half of the repair shops in our database agreed to participate in the survey as at the time of the survey the political situation in Turkey was rather tense and some shops were afraid that the survey had been commissioned by the government in an attempt to gather delicate information on Turkish citizens.¹² In total, 76 of the shops visited for repair participated in our survey. Table 4 replicates the contents of Table 2 for the shops

¹² In summer 2016 there was a failed coup attempt by the Turkish military and as a consequence the Turkish government declared a state of emergency until July 2018.

who participated in the survey – the numbers in the cells are again the average repair price and (in parentheses) the number of observations. A comparison of the contents of tables 2 and 4 suggests that there is no systematic pattern behind the decision to participate in the survey as the average repair price in each treatment for the subsample is rather similar to the corresponding entry for the whole sample.

Table 4: average repair price (in TRY) and number of observations (in parentheses) for the survey subsample

	ORDINARY	CREDENCE	both service types
MAJORITY	91 (17)	128 (21)	112 (38)
MINORITY	93 (22)	146 (16)	115 (38)
both customer types	92 (39)	136 (37)	113 (76)

Table 5 displays the mean response to some selected survey questions (see Appendix B for the complete survey and all the results) separately for two ethnicities of sellers – Turks and Kurds. The members of both ethnicities are Turkish citizens; Turks are the majority – according to various estimates only about 15% to 20% of the Turkish population consists of Kurds (see the report of the Home Office 2018). We decided to group the survey results by the ethnicity of the seller because it turned out that Kurdish and Turkish survey participants have significantly different attitudes towards immigration and this difference in attitudes translates into different behavior in the main experiment. The difference in attitudes can be seen in the last column of Table 5 reporting the p-values of the pairwise comparisons between Kurds and Turks for the corresponding question.

Table 5: attitudes of repair shops grouped by ethnicity

N°	selected question from survey	Turkish	Kurdish	p-value
1	<i>What is the highest level of education that you have completed? (1) I have no formal education – (6) I have completed more than 12 years of schooling.</i>	4.64	4.26	0.0192 (MW-Test)
2	<i>Imagine Turkish society as arranged on a scale where the worst off socially and economically are on the left (0) and the best off are on the right (10). Please select the place where you feel you stand today.</i>	4.67	4	0.3787 (T-test)
3	<i>Is Turkey made a worse (10) or a better place (0) to live by people coming to live here from other countries?</i>	4.87	5.29	0.6246 (T-test)
4	<i>To what extent do you think Turkey should allow people of the same race or ethnic group as most Turkish people to come and live here. (1) allow many – (4) allow none.</i>	2.33	1.90	0.1022 (MW-test)
5	<i>To what extent do you think Turkey should allow people of different race or ethnic group than most Turkish people to come and live here. (1) allow many – (4) allow none.</i>	2.49	2.07	0.1189 (MW-test)
6	<i>Some people come to this country and apply for refugee status on the grounds that they fear persecution in their own country. Please say how much you agree (5) or disagree (1) that most applicants for refugee status aren't in real fear of persecution in their own countries.</i>	3.53	4.29	0.0066 (MW-test)
7	<i>Think of people who have come to live in Turkey from another country who are of a different race or ethnic group from most Turkish people. Please tell me how much you would mind (10) or not (0) if someone like this marries a close relative of yours.</i>	6.4	4.23	0.0059 (T-test)
8	<i>Would you say that people who come to live here generally take jobs away (10) from workers in Turkey or generally help to create new jobs (0)?</i>	7.22	5.19	0.0062 (T-test)
9	<i>Most people who come to live here work and pay taxes. They also use health and welfare services. On balance, do you think people who come here take out more from society than they put in (10) or put in more to society than they take out (0)?</i>	6.62	5.39	0.09672 (T-test)
10	<i>Would you say that Turkey's cultural life is generally undermined (10) or enriched (0) by people coming to live here from other countries?</i>	5.71	5.13	0.4031 (T-test)
11	<i>How often do you have any contact with people who are of a different race or ethnic group from most Turkish people when you are not at home? This could be on public transport, in the streets, in shops or in the neighborhood (any contact should be included whether verbal or non-verbal)? (1) never – (7) every day.</i>	4.71	5.26	0.1370 (MW-test)
12	<i>Thinking about this contact, in general how bad (10) or good (0) is it?</i>	3.8	3.91	0.8877 (T-test)
13	<i>Do you think some races or ethnic groups are born less intelligent than others? (1) yes – (0) no.</i>	0.42	0.74	0.0062 (MW-test)
14	<i>Do you think some races or ethnic groups are born harder working than others? (1) yes – (0) no.</i>	0.6	0.84	0.0260 (MW-test)
15	<i>Compared to people like yourself, would you say that those who have come to live here from other countries are better or worse off financially? (5) much better off – (1) much worse off.</i>	3.49	3.32	0.3550 (MW-test)
16	<i>How religious would you say you are? (0) not at all religious – (10) very religious.</i>	4.2	4.52	0.6522 (T-test)
17	<i>How emotionally attached do you feel to Turkey? (0) not at all emotionally attached – (10) very emotionally attached.</i>	6.91	4.13	0.0009 (T-test)
# Observations		45	31	

We switched between two-sided T-tests and Mann-Whitney tests according to the measurement scale of the variable of interest.

In the first row we see that Turks attended school significantly longer than Kurds but this doesn't translate into differences in socio-economic status across the two groups according to the answer to question 2. Regarding attitudes towards immigration in general, row 3 reveals that Kurds and Turks think to the same degree that Turkey is neither a better nor a worse place because of immigration. However, when we look at the attitudes towards immigration from an economic and cultural perspective we find significant differences between the two groups: In rows 8 and 9 we see that Turks think significantly more often that immigrants take away jobs from the local society and that they generally take out more than they put in. Row 10 reveals that Kurds and Turks believe to the same degree that their culture is neither enriched nor undermined by immigration. Interestingly, Turks would mind significantly more that a non-local with a different ethnicity marries a close relative (see row 7).

The partially diverging views towards the consequences of immigration do not translate into differences when it comes to the question to what extent Turkey should allow people of the same (different) ethnic group to come and live here. We conclude this from the relatively positive answers of Kurds and Turks to questions 3 and 4. Very interesting in this context is that Kurds think significantly more often that people who apply for refugee status in Turkey are not in real fear of prosecution in their home countries (see question 6). When it comes to estimate the financial wealth of immigrants (question 15), Kurds and Turks believe to the same extent that migrants are financially slightly better off than themselves.

One of the most interesting differences between the two groups is displayed in rows 13 and 14: Kurds exhibit more racist attitudes as they think significantly more often that some ethnic groups are born more intelligent and harder working than others. None of these differences in attitudes towards immigration between Kurds and Turks can be explained by the contact hypothesis (see Allport 1954 and Amir 1969) because both groups report a relatively high level of contact quantity (question 11) and quality (question 12). Finally, question 16 reveals that both groups are only moderately religious and that, not surprisingly, Turks feel significantly more attached to Turkey than Kurds (question 17).

Taken together the results displayed in Table 5 suggest that Turks believe more often that migration has negative consequences from an economic perspective and that Kurds exhibit more racist attitudes in general.¹³ It seems important to note that the survey did not specifically ask for attitudes towards Syrian immigrants but rather about attitudes towards immigration in general. It is rather plausible, though, that most survey participants had Syrian

¹³ The stronger manifestation of general prejudices on the side of Kurds could be partly driven by the fact that Kurds had less years of schooling than Turks (see Wagner and Zick 1995 for further reading).

immigrants in mind when answering the questions. Evidence supporting this conjecture comes from an additional question where we asked the survey participants about the main countries they believe people immigrate from: out of our 76 survey participants, 73 think that people who come to live in Turkey are mainly from Syria (see question b32 in Appendix B).

Table 6 provides a first answer to the question of how the different attitudes of Kurds and Turks translate into differences in behavior in the main experiment. It presents the average price per treatment contingent on the ethnicity of the service provider and reveals that the average repair price is almost identical across the two ethnicities in the **ORD**-treatments but different in the **CRE**-treatments. In the latter Kurds charge higher prices than Turks, on average. In light of our main results more important is the fact that Kurds add a noticeable discriminatory markup in the **CRE**-treatments (143 vs. 192 TRY), while Turks do not (121 vs. 125 TRY). This already suggests that the higher discriminatory markup for the credence goods service is mainly driven by the behavior of ethnic minority Turkish Kurds.¹⁴

Table 6: average repair price (in TRY) and number of observations (in parentheses) for the survey subsample contingent on ethnicity of shop staff

	ORDINARY		CREDENCE	
	TURKISH	KURDISH	TURKISH	KURDISH
MAJORITY	92 (13)	90 (4)	121 (14)	143 (7)
MINORITY	93 (7)	93 (15)	125 (11)	192 (5)

The hypothesis that the interaction of the credence goods markup and the discriminatory markup is mainly driven by the behavior of ethnic minority Turkish Kurds is confirmed by the OLS-regression in Table 7. This regression includes the repair price as the dependent variable and dummies for different combinations of our four treatments and the ethnicity of the seller. Taking the **MAJ-ORD** treatment with **TURKISH** expert providers as the (omitted) benchmark, **MIN-CRE** & **KURDISH** providers has the strongest and biggest price increasing effect (100.46 TRY). The effect of **MAJ-CRE** & **KURDISH** providers is also statistically significant and economically impressive (51.32 TRY) but at the same time smaller than the effect of **MIN-CRE** & **KURDISH** providers (a pairwise comparison of the regression

¹⁴ The differences in animus towards other ethnicities are also manifested within the subset of experts in the **MIN-CRE** treatment: In the **MIN-CRE** treatment, 100 % of the Kurdish experts answer question 14 with “yes” (versus 64 % of the Turkish experts) and 80 % answer question 13 with “yes” (versus 34 % of the Turkish experts).

coefficients between **MAJ-CRE** & KURDISH providers vs. **MIN-CRE** & KURDISH providers results in a p-value of 0.072) and this means that the discriminatory markup is mainly driven by the behavior of Kurdish providers. The effects of **MAJ-CRE** & TURKISH providers and **MIN-CRE** & TURKISH providers are significant at the 10% level, suggesting that there is also evidence for a credence goods markup when we focus solely on Turkish expert providers. None of the remaining explanatory variables has a significant effect on the repair price.

Table 7. regression analysis of repair prices contingent on treatment and ethnicity

Dependent variable (OLS regressions)	[1] <i>Repair price</i>
Independent variables	<i>(in TRY)</i>
MAJ-ORD & KURDISH providers	-1.54 (26.29)
MAJ-CRE & TURKISH providers	29.53* (17.71)
MAJ-CRE & KURDISH providers	51.32** (21.26)
MIN-ORD & TURKISH providers	1.32 (21.55)
MIN-ORD & KURDISH providers	1.13 (17.43)
MIN-CRE & TURKISH providers	33.46* (18.84)
MIN-CRE & KURDISH providers	100.46*** (24.20)
Constant	91.54*** (12.75)
# Observations	76

OLS-regression with repair price (in TRY) as dependent variable, including, as explanatory variables, the treatment dummies grouped by the ethnicity of the expert provider (the treatment **MAJ-ORD** with TURKISH expert providers serves as the benchmark treatment). ***, **, * denote significance at the 1%, 5%, 10% level, standard errors in parentheses.

We summarize our findings in this subsection as follows:

Result 4 (drivers of the sophisticated form of discrimination): *The results of our ex post survey suggest that the sophisticated form of discrimination uncovered in our analysis is mainly due to the behavior of ethnic minority sellers and that the nature of the observed discriminatory behavior is mainly taste-based, i.e., driven by racial bias as opposed to realistic conflict over economic resources.*

4. Conclusion

The present study has made two main contributions: First, by introducing a novel toolkit for the investigation of discrimination against minorities in the marketplace it has paved the way for future work in this discipline. Second, by applying the toolkit to a specific credence goods market it has shown that discriminatory behavior is by far more important in markets where sellers can hide behind a preexisting information asymmetry than in ordinary goods markets where they cannot.

The toolkit introduced in the present paper consists of four experimental treatments complemented by an ex post survey. The four treatments are organized in a 2x2 factorial design varying in one dimension the characteristic of the good – an ordinary good in two treatments and a credence good in the other two treatments – and in the second dimension the characteristic of the customer – a native of Turkey in two of the treatments and a member of an immigrant minority in the other two treatments. Our factorial design allows the derivation of two markups: (i) the credence goods markup defined as the difference between the price paid by a customer for an ordinary service and the price paid by the customer for an (otherwise equivalent) credence goods service; and (ii) the discriminatory markup defined as the difference between the price paid by a member of the minority and the price paid by a member of the majority for the same kind of service. The complementary survey is conducted after the main experiment and exposes the expert providers visited during the main experiment to a series of questions that can help to identify the drivers of discrimination. The toolkit can easily be adapted to different potential objects of discrimination (race, gender, age, sexual orientation, ...) and to different credence goods markets.

Uncovering discrimination in credence goods markets seems especially important because the size of those markets is huge and the potential for discrimination therefore

large.¹⁵ At the same time uncovering discrimination in such markets is tricky as expert providers can hide behind the pre-existing information asymmetry and are willing to do so since discrimination is a socially sanctioned behavior. As a consequence, the chances for the victims of discrimination to detect this kind of misbehavior on the experts' side on their own are slim. Academic research on this issue seems therefore important – not only to pin down the extent and the drivers of discrimination but also to find policy instruments to fight it.

In the present paper, we have applied the proposed toolkit to the Turkish market for cellphone repairs. Our data suggest that discrimination is by far more important in the credence goods market where the discriminatory markup is economically large and statistically highly significant than in the ordinary goods market where it is negligible. This result suggests that – by focusing exclusively on ordinary goods markets – the previous literature has underestimated the adverse effects of discrimination.

Gneezy and List (2013) point out the importance of examining the mechanism behind discriminatory behavior because without understanding the drivers of discrimination it is hard to protect the victims from disparate treatment. The two most prominent explanations for discrimination in the market place are the taste-based theory by Becker (1957) and the statistical theory by Arrow (1973) and Phelps (1972). The taste-based theory explains discriminatory behavior with animus or racism towards a specific group. Looking back to the 50s, at the time where this theory was established, discriminatory behavior based on animus manifested itself in extreme ways – e.g., US banks refused loans to black home buyers only because of their race, shop owners barred ethnic minorities from access to their shops, etc. In these examples taste-based discrimination resulted in a monetary loss for the discriminator – which is often considered as a defining property of taste-based discrimination. By contrast, statistical discrimination is per definition consistent with the notion of profit maximization: in his pursuit of higher profits, the discriminator uses observable characteristics to make statistical inferences about the willingness to pay or the outside option of the victim; this information is then used to (third-degree) price-discriminate among customers - see Charles and Guryan (2013) for a deeper discussion. The latest psychological research blurs the sharp line between statistical and taste-based discrimination that economists tend to establish (see Bertrand and Duflo 2017 for a deeper discussion). In a similar vein, Gneezy and List (2013) argue that animus and racism can combine with economic discrimination and that the consequences are terrible in these cases.

¹⁵ For instance, health care expenditures alone account for about 10% of GDP in the OECD-countries (www.oecd-library.org). The finance sector represents 9% of worldwide GDP (see *The Economist*, 2014: <http://www.economist.com/news/finance-and-economics/21604574-new-paper-shows-industrys-take-has-been-rising-counting-cost-finance>), and repair services generate more than 100 billion Euro per annum in Europe alone (ec.europa.eu/eurostat). Links accessed on 10 January 2019.

Our study provides support for the conjecture by Gneezy and List (2013): The results of our ex-post survey suggest that our discrimination result is largely driven by the behavior of ethnic minority expert providers who, compared to ethnic majority expert providers, are less likely to perceive immigrants as an economic threat while exhibiting more racial prejudice. This suggests that the discriminatory behavior uncovered in our analysis is mainly motivated by animus – and that the discrimination is therefore taste-based. However, the monetary consequence of the discriminatory behavior for the discriminator is not to decrease, but rather to increase the profit. In other words, the unfavorable position of established minority groups vis-à-vis the majority seems to pave the way for prejudice towards immigrants and these prejudices seem to translate into money maximizing on the experts' side.¹⁶

But can animus-based discrimination really be profit-maximizing behavior? At first glance, this combination seems to lead to a contradiction: If the behavior towards the members of the minority is profit-maximizing behavior, why isn't the same behavior also applied to the members of the majority? Framed differently, why do certain sellers charge high prices only from the members of the minority and not also from members of the majority? The answer seems to lie in moral costs. Exploiting their information advantage seems to impose a moral cost on the sellers. And this moral cost seems to be lower for minority sellers if the trading partner belongs to a low-status "outgroup". In this sense, animus-based discrimination still comes at a cost, but the cost is a moral and not a material one.

The animus-based explanation of our results is in line with the findings by Gneezy et al. (2012). In a study investigating the nature and extent of discrimination against several distinct groups across several different markets, the authors find that the nature of discrimination (taste-based vs. statistical) is less driven by the particulars of the market or the minority considered, but rather by whether the object of discrimination is chosen by the individual or uncontrollable: When the object of discrimination is chosen by the individual, the evidence suggests that the nature of discrimination is animus; by contrast, when the object of discrimination is perceived to be out of the control of the individual the underlying form of discrimination tends to be statistical. Given that 100 percent of the Kurdish experts who serve immigrant minority customers in a credence goods setting believe that most applicants for refugee status aren't in real fear of prosecution in their home countries, the decision to

¹⁶ The specific reasons for the observed prejudices towards the out-group could be manifold. The civil war in Syria and the ongoing ambitions of Kurds for an autonomous Kurdistan could be two channels that influence the social identity and the intergroup relationship between Kurds and Syrians.

migrate is arguably interpreted by them as a free choice. In this sense our results are completely in line with those in Gneezy et al. (2012).

The results of our ex post survey are in line with claims derived from social identity theory in psychology – see Tajfel and Turner (1996), for instance: minority group members – in our case Kurds – often deal with their own unfavorable position in society and with threats to their self-esteem by making downward social comparisons to other minority groups – i.e., ‘we are more intelligent and hardworking than other ethnicities’, etc. In our case the self-esteem enhancing downward social comparisons seem to have resulted in hard-to-eliminate taste-based discrimination, thereby imposing costs on third parties, and the economy as a whole.

Uncovering sophisticated discrimination in the marketplace and identifying the drivers of it is one thing; fighting against this form of misbehavior is another. Does our research suggest any specific policy recommendations or advice for customers that are potentially helpful in this context? Given that we identified taste-based discrimination as the main driver for our results, it is unlikely that there exists a quick fix for this problem and this is bad news.¹⁷ The good news is that our data suggest some advice for customers that is potentially helpful in containing the extent of exploitation. The simplest advice would be to invest time and effort to get a reliable self-diagnosis of the problem: This transforms the credence good transaction into an ordinary goods transaction and in the market under consideration there is no evidence for discrimination against minorities in the ordinary goods arm of the experiment.¹⁸ An alternative approach would be for customers in credence goods situations to search for an expert seller with whom they have a shared identity. However, as greater contact across groups is a reliable predictor of prejudice reduction such an approach may only serve to reinforce prejudice, reducing prospects for social change through day-to-day intergroup contact over the long-term.

¹⁷ Prejudices between Kurds and Turks are the product of a long and complex joint history between these two ethnicities, and the possibility that Kurds may offset their unfavorable social position by discriminating against low-status outgroups seems to be rooted in basic social psychological processes. It is unlikely that an intervention from a 3rd party will change the minds of the involved people in the short term.

¹⁸ Of course, the advice of getting informed might also be valuable for members of the majority – if they reveal that they know the source of the problem they save, on average, 33 TRY or about 35% of the price for the ordinary good. However, the cost-saving effect of becoming informed is much larger for the members of the minority than for members of the majority because the former can save not only the credence goods markup but also the discrimination markup while the latter can save only the credence goods markup.

References

- Allport, G.W. (1954). The nature of prejudice. *Addison-Wesley Publishing Company*, Cambridge, Mass.
- Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin* 71, 319-342.
- Anagol, S., Cole, S., Sarkar, S. (2017). Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance market. *Review of Economics and Statistics* 99, 1-15.
- Arrow, K.J. (1973). The Theory of Discrimination. *Discrimination in Labor Markets* 3, 3-33.
- Ayres, I., Siegelman, P., (1995). Discrimination in bargaining for a new car. *The American Economic Review* 85, 304-321.
- Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M. (2013). What drives taxi drivers? A field experiment on fraud in a market for credence goods. *Review of Economic Studies* 80, 876-891.
- Balafoutas, L., Kerschbamer, R., Sutter, M. (2017). Second-degree moral hazard in a real-world credence goods market. *Economic Journal* 127, 1-18.
- Becker, G.S. (1957). The Economics of Discrimination. *University of Chicago Press*.
- Bertrand, M., Duflo, E. (2017). Chapter 8 – Field experiments on discrimination. *Handbook of Economic Field Experiments* 1, 309-393.
- Castillo, M., Petrie, R., Torero, M., Vesterlund, L. (2013). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics* 99, 35-48.
- Darby, M., Karni, E. (1973). Free competition and the optimal amount of fraud. *Journal of Law and Economics* 16, 67-88.
- Doleac, J.L., Stein L.C.D. (2013). The visible hand: Race and online market outcomes. *The Economic Journal* 123, 469-492.
- Dulleck, U., Kerschbamer, R. (2006). On doctors, mechanics and computer specialists – The economics of credence goods. *Journal of Economic Literature* 44, 5-42.
- Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L. (2001). Measuring trust. *The Quarterly Journal of Economics* 115, 811-846.
- Gneezy, U., List, J.A., Price, M.K. (2012). Towards an understanding of why people discriminate: Evidence from a series of natural field experiments. NBER working paper series 17855, <http://www.nber.org/papers/w17855>.
- Gneezy, U., List J.A. (2013). The why axis -hidden motives and the undiscovered economics of everyday life. *Public Affairs*, New York.

Gottschalk, F., Mimra, W., Waibel, C. (2018). Health services as credence goods: A field experiment. Available at SSRN: <https://ssrn.com/abstract=3036573>.

Guryan, J., Charles, K.K. (2013): Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal* 123, 417-432.

Harrison, G.W., List, J.A. (2004). Field experiments. *Journal of Economic Literature* 42, 1009-1055.

International Crisis Group (2018). *Turkey's Syrian Refugees: Defusing Metropolitan Tensions*. Report No. 241.

Kerschbamer, R., Neururer, D., Sutter, M. (2016). Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proceedings of the National Academy of Sciences* 113, 7454-7458.

Kerschbamer, R., Sutter, M. (2017). The economics of credence goods – A survey of recent lab and field experiments. *CESifo Economic Studies* 63, 1-23.

Knack, S., Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics* 112, 1251-1288.

List, J.A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics* 119, 49-89.

Lu, F. (2014). Insurance coverage and agency problems in doctor prescriptions: Evidence from a field experiment in China. *Journal of Development Economics* 106, 156-167.

McAuliffe, M., Ruhs, M. (2017). *World Migration Report 2018*. Geneva: International Organization for Migration.

Nunley, J.O., Mark, O., Howard, S. (2011). The effects of information and competition on racial discrimination: Evidence from a field experiment. *Journal of Economic Behavior & Organization* 80, 670-679.

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

Schneider, H.S. (2012). Agency problems and reputation in expert services: Evidence from auto repair. *Journal of Industrial Economics* 60, 406-433.

Tajfel, H., Turner, J. C. (1986). "The Social Identity Theory of Inter-group Behavior". In S. Worchel & L. W. Austin (Eds.), *Psychology of Intergroup Relations*. Chicago: Nelson-Hall.

The Home Office (2018). "Country Policy and Information Note, Turkey: Kurds", Version 2.0.

Wagner, U., Zick, A. (1995). The relation of formal education to ethnic prejudice: Its reliability, validity and explanation. *European Journal of Social Psychology* 25, 41-56.

Zitzewitz, E. (2012). Forensic economics. *Journal of Economic Literature* 50, 731-769.

Zussman, A. (2013). Ethnic discrimination: Lessons from the Israeli online market for used cars. *The Economic Journal* 123, 433-468.

Appendix A



Faces of Discrimination (#12363)

Created: 07/02/2018 02:43 AM (PT)

Shared: 03/24/2019 02:40 AM (PT)

This pre-registration is not yet public. This anonymized copy (without author names) was created by the author(s) to use during peer-review. A non-anonymized version (containing author names) will become publicly available only if an author makes it public. Until that happens the contents of this pre-registration are confidential.

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

We plan to conduct a natural field experiment in the Turkish market for cell phone repairs to address the following research questions: (i) To what extent are non-locals discriminated against in an ordinary service market? (ii) To what extent are non-locals discriminated against in a credence goods market? (iii) How large is the 'credence markup' (that is, the price difference between a service without credence goods character and the same service with credence goods character) in general? (iv) Is there a difference in the credence markup for non-locals and local customers?

3) Describe the key dependent variable(s) specifying how they will be measured.

The main dependent variable is the average repair price in each treatment group.

4) How many and which conditions will participants be assigned to?

We will use a mystery shopper approach, where research assistants enter cell phone repair shops with a non-functioning phone and ask for a repair. The mystery shoppers are either locals (Turkish) or non-local (Syrian), and either know (ordinary) or do not know (credence) the problem with the phone. The participants (i.e. the cell phone repair shops) are thus assigned to four treatments:

- ordinary_local: ordinary service market and local customer
- credence_local: credence goods market and local customer
- ordinary_non-local: ordinary service market and non-local (Syrian) customer
- credence_non-local: credence goods market and non-local (Syrian) customer

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

- By comparing the average repair price in the treatment ordinary_local to the average repair price in the treatment ordinary_non-local we address research question (i).
- By comparing the average repair price in the treatment credence_local to the average repair price in the treatment credence_non-local we address research question (ii).
- By comparing the average repair price in the treatment ordinary_local to the average repair price in the treatment credence_local we address research question (iii).
- By comparing the difference of average repair prices between ordinary_local and credence_local to the difference of average repair prices between ordinary_non-local and credence_non-local we address research question (iv).

In addition to the binary treatment comparisons with non-parametric tests sketched above, the collected data would also allow us to perform a classical difference-in-difference estimation using standard econometric techniques.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude observations if the participants (shop owners) realize that it is an experiment. If a shop declines to repair the phone, we will not be able to include it in the price comparisons. This is not a real exclusion though, and if this happens more often, we will control whether there is an effect of the treatments on refusal to repair the phone.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

In total, we plan to collect 200 observations (50 observations per treatment). We will strive to select the sample in such a way that the person behind the counter is Turkish (native). However, we will make sure to document this and control for it if it happens that several shop owners are non-native.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

In addition to the data on repair prices we also plan to collect data on characteristics and attitudes of repair shops via a questionnaire to be filled out by our mystery shoppers and a survey to be completed by the shop owners.

Verify authenticity: <http://aspredicted.org/blind.php?x=vu97v7>

Appendix B

In the following we present all questions stated in the ex-post survey and report the differences between Turkish and Kurdish participants.

Table B1: ex-post survey grouped by ethnicity

N°	selected question from survey	Turkish	Kurdish	p-value
b1	What is your gender? (0) male – (1) female.	0.16	0.13	0.7470 (Chi2-test)
b2	What year were you born? (1) 2018 – (101) 1918.	32.6	36	0.1182 (T-test)
b3	Were you born in Turkey? (1) Yes – (0) no.	1	1	---
b5	What is your ethnicity? (1) Turkish, (2) Kurdish and (3) other.	1	2	---
b6	What is your religion? (1) Islam Sunni, (2) Islam Shia, (3) Christian and (4) other.	1.44	1.68	0.005 (Chi2-test)
b7	What is the highest level of education that you have completed? (1) I have no formal education – (6) I have completed more than 12 years of schooling.	4.64	4.26	0.0192 (MW-test)
b8	Imagine Turkish society as arranged on a scale where the worst off socially and economically are on the left (0) and the best off are on the right (10). Please select the place where you feel you stand today.	4.67	4	0.3787 (T-test)
b9	In general, people often face risks when making financial, career, or other life decisions. Overall, do you feel comfortable, uncomfortable, or neither comfortable nor uncomfortable taking risks? (1) Extremely comfortable taking risks (1) – (7) extremely uncomfortable taking risks.	2.58	2.48	0.8951 (MW-test)
b10	Is Turkey made a worse (10) or a better place (0) to live by people coming to live here from other countries?	4.87	5.29	0.6246 (T-test)
b11	To what extent do you think Turkey should allow people of the same race or ethnic group as most Turkish people to come and live here. (1) allow many – (4) allow none.	2.33	1.90	0.1022 (MW-test)
b12	To what extent do you think Turkey should allow people of different race or ethnic group than most Turkish people to come and live here. (1) allow many – (4) allow none.	2.49	2.07	0.1189 (MW-test)
b13	To what extent do you think Turkey should allow people from the poorer countries in the region to come and live here? (1) Allow many to come and live here – (4) allow none.	2.47	2.10	0.1421 (MW-test)
b14	Some people come to this country and apply for refugee status on the grounds that they fear persecution in their own country. Please say how much you agree (5) or disagree (1) that most applicants for refugee status aren't in real fear of persecution in their own countries.	3.53	4.29	0.0066 (MW-test)
b15	Think of people who have come to live in Turkey from another country who are of a different race or ethnic group from most Turkish people. Please tell me how much you would mind or not if someone like this was appointed as your boss? (0) Not mind at all – (10) mind a lot.	5.62	4.42	0.1263 (MW-test)
b16	Think of people who have come to live in Turkey from another country who are of a different race or ethnic group from most Turkish people. Please tell me how much you would mind (10) or not (0) if someone like this marries a close relative of yours.	6.4	4.23	0.0059 (T-test)
b17	Would you say that people who come to live here generally take jobs away (10) from workers in Turkey or generally help to create new jobs (0)?	7.22	5.19	0.0062 (T-test)

b18	<i>Most people who come to live here work and pay taxes. They also use health and welfare services. On balance, do you think people who come here take out more from society than they put in (10) or put in more to society than they take out (0)?</i>	6.62	5.39	0.09672 (T-test)
b19	<i>Would you say it is generally bad or good for the Turkish economy that people come to live here from other countries? (0) Good for economy – (10) bad for economy.</i>	6.53	6.48	0.8217 (MW-test)
b20	<i>Are Turkey's crime problems made worse or better by people coming to live here from other countries? (0) Crime problems made better – (10) crime problems made worse.</i>	7.09	6.26	0.2564 (MW-test)
b21	<i>How much do you agree or disagree with the following statement: If you want to make money, you can't always act honestly. (5) Strongly agree – (1) strongly disagree.</i>	3.84	3.39	0.2697 (MW-test)
b22	<i>Would you say that Turkey's cultural life is generally undermined (10) or enriched (0) by people coming to live here from other countries?</i>	5.71	5.13	0.4031 (T-test)
b23	<i>Do you think the religious beliefs and practices in Turkey are generally undermined or enriched by people coming to live here from other countries? (0) Religious beliefs and practices enriched – (10) religious beliefs and practices undermined.</i>	5.33	5.58	0.7613 (MW-test)
b24	<i>How would you describe the area where you currently live? (1) An area where almost nobody is of a different race or ethnic group from most Turkish people, (2) some people are of a different race or ethnic group from most Turkish people and (3) many people are of a different race or ethnic group.</i>	1.82	1.71	0.3165 (MW-test)
b25	<i>How often do you have any contact with people who are of a different race or ethnic group from most Turkish people when you are not at home? This could be on public transport, in the streets, in shops or in the neighborhood (any contact should be included whether verbal or non-verbal)? (1) never – (7) every day.</i>	4.71	5.26	0.1370 (MW-test)
b26	<i>Thinking about this contact, in general how bad (10) or good (0) is it?</i>	3.8	3.91	0.8877 (T-test)
b27	<i>Do you have any close friends who are of a different race or ethnic group from most Turkish people. (1) Yes, several, (2) yes, a few and (3) no, none at all.</i>	1.87	1.42	0.0066 (MW-test)
b28	<i>Do you think some races or ethnic groups are born less intelligent than others? (1) yes – (0) no.</i>	0.42	0.74	0.0062 (MW-test)
b29	<i>Do you think some races or ethnic groups are born harder working than others? (1) yes – (0) no.</i>	0.6	0.84	0.0260 (MW-test)
b30	<i>Out of every 100 people living in Turkey, how many do you think were born outside Turkey?</i>	27.29	23.29	0.4317 (T-test)
b31	<i>Compared to people like yourself, would you say that those who have come to live here from other countries are better or worse off financially? (5) much better off – (1) much worse off.</i>	3.49	3.32	0.3550 (MW-test)
b32	<i>Thinking about people who come to live in Turkey from other countries, which is the main country you think they came from? (3) Syria, (4) Iraq, (5) Afghanistan and (6) Iran.</i>	2.98	3.07	0.5320 (Chi2-test)
b33	<i>How religious would you say you are? (0) not at all religious – (10) very religious.</i>	4.2	4.52	0.6522 (T-test)
b34	<i>Apart from special occasions such as weddings and funerals, about how often do you attend religious services nowadays? (7) Every day – (1) never.</i>	3.84	4.23	0.3997 (MW-test)
b35	<i>Apart from when you are at religious services, how often, if at all, do you pray? (7) Every day – (1) never.</i>	5.51	5.61	0.6234 (MW-test)

<i>b36</i>	<i>How emotionally attached do you feel to Turkey? (0) not at all emotionally attached – (10) very emotionally attached.</i>	6.91	4.13	0.0009 (T-test)
# Observations		45	31	

We switched between two-sided T-tests, Chi2 tests and Mann-Whitney tests according to the measurement scale of the variable of interest.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

<https://www.uibk.ac.at/eeecon/wopec/>

- 2019-11 **Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog:** Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment
- 2019-10 **Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Collective intertemporal decisions and heterogeneity in groups
- 2019-09 **Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller, Jean-Robert Tyran:** Distributional Preferences Explain Individual Behavior Across Games and Time
- 2019-08 **Daniel Müller, Sander Renes:** Fairness Views and Political Preferences - Evidence from a representative sample
- 2019-07 **Florian Lindner, Michael Kirchler, Stephanie Rosenkranz, Utze Weitzel:** Social Status and Risk-Taking in Investment Decisions
- 2019-06 **Christoph Huber, Julia Rose:** Individual attitudes and market dynamics towards imprecision
- 2019-05 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, Stefan Zeisberger:** What Drives Risk Perception? A Global Survey with Financial Professionals and Lay People
- 2019-04 **David M. McEvoy, Tobias Haller, Esther Blanco:** The Role of Non-Binding Pledges in Social Dilemmas with Mitigation and Adaptation
- 2019-03 **Katharina Momsen, Markus Ohndorf:** When do people exploit moral wiggle room? An experimental analysis in a market setup
- 2019-02 **Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** Credence goods markets and the informational value of new media: A natural field experiment
- 2019-01 **Martin Geiger, Eric Mayer, Johann Scharler:** Inequality and the Business Cycle: Evidence from U.S. survey data
- 2018-18 **Matthias Sutter, Jürgen Huber, Michael Kirchler, Matthias Stefan, Markus Walzl:** Where to look for the morals in markets?
- 2018-17 **Rene Schwaiger, Michael Kirchler, Florian Lindner, Utz Weitzel:** Determinants of investor expectations and satisfaction. A study with financial professionals

- 2018-16 **Andreas Groll, Julien Hambuckers, Thomas Kneib, Nikolaus Umlauf:** LASSO-Type Penalization in the Framework of Generalized Additive Models for Location, Scale and Shape
- 2018-15 **Christoph Huber, Jürgen Huber:** Scale matters: Risk perception, return expectations, and investment propensity under different scalings
- 2018-14 **Thorsten Simon, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Lightning prediction using model output statistics
- 2018-13 **Martin Geiger, Johann Scharler:** How do consumers interpret the macroeconomic effects of oil price fluctuations? Evidence from U.S. survey data
- 2018-12 **Martin Geiger, Johann Scharler:** How do people interpret macroeconomic shocks? Evidence from U.S. survey data
- 2018-11 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Low visibility forecasts for different flight planning horizons using tree-based boosting models
- 2018-10 **Michael Pfaffermayr:** Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach
- 2018-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model
- 2018-08 **Lisa Schlosser, Torsten Hothorn, Reto Stauffer, Achim Zeileis:** Distributional regression forests for probabilistic precipitation forecasting in complex terrain
- 2018-07 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Delegated decision making and social competition in the finance industry
- 2018-06 **Manuel Gebetsberger, Reto Stauffer, Georg J. Mayr, Achim Zeileis:** Skewed logistic distribution for statistical temperature post-processing in mountainous areas
- 2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach
- 2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals
- 2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items
- 2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies
- 2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding

University of Innsbruck

Working Papers in Economics and Statistics

2019-11

Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog

Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment

Abstract

We present the results of a pre-registered natural field experiment designed to uncover a sophisticated form of discrimination against an immigrant minority in a market for credence goods. For this purpose, we introduce two markups: (i) the credence goods markup defined as the difference between the price paid by the same person for an ordinary service and an otherwise equivalent credence goods service; and (ii) the discriminatory markup defined as the difference between the price paid by a member of an immigrant minority group and the price paid by a member of the majority group for the same kind of service. We document the existence of a large credence goods markup of about 40%, on average. Moreover, we find a sizeable discriminatory markup for the credence goods service but no discriminatory markup for the ordinary service. The results of an ex-post survey suggest that this sophisticated form of discrimination is mainly due to the prejudicial behavior of sellers belonging to an established local ethnic minority group towards buyers belonging to a low-status immigrant ethnic minority group.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)