Ethnic Discrimination and Signals of Trustworthiness in an Online Market: Evidence from Two Field Experiments

Wojtek Przepiorka[1]

University of Oxford, Department of Sociology, Manor Road, Oxford OX1 3UQ, UK, and ETH Zürich, Professur für Soziologie, Clausiusstrasse 50, CLU D 4, 8092 Zürich, Schweiz
wojtek.przepiorka@sociology.ox.ac.uk

Summary: Results of two field experiments which were designed to identify possible ethnic discrimination on a German internet auction platform are discussed. A first set of results was produced by a secondary analysis of an earlier experiment. The second experiment then tested whether costly signals could help to overcome problems of trust between buyers and sellers in online markets. The evidence is rather mixed with respect to ethnic discrimination, and it does not support the signaling hypothesis.

Keywords: Discrimination; Costly Signaling; Trust; Online-Market; Experimental Sociology; Field Experiment.

Zusammenfassung: Die Ergebnisse von zwei Feldexperimenten zur Untersuchung einer möglichen ethnischen Diskriminierung auf einer deutschen Internet-Auktions-Plattform werden diskutiert. Hierfür wird zunächst ein früheres Experiment einer Sekundäranalyse unterzogen. Das zweite Experiment untersucht darüberhinaus, ob kostspielige Signale helfen können, Vertrauensprobleme zwischen Käufern und Verkäufern auf Online-Märkten zu lösen. Die Befunde sind hinsichtlich ethnischer Diskriminierung uneindeutig und stützen die „Signaling"-Hypothese nicht.

Schlagworte: Diskriminierung; Kostspielige Signale; Vertrauen; Online-Markt; Experimentelle Soziologie; Feldexperiment.

1. Introduction

The unequal treatment of people or groups based on their ethnicity or race is regarded as ethnic or racial discrimination. Economists have conceptualized discrimination as taste-based and information-based. Becker (1957) was the first to suggest that people might have a taste (i.e., preference) for interacting with one kind of people but not another. In his theoretical framework, Becker incorporates such a taste as a non-monetary part of a person's utility function which, if negative, is called discrimination and, if positive, is called nepotism (i.e., positive discrimination). Becker notes, however, that even if such a taste did not exist, we would observe unequal treatment of people belonging to different social groups because they either objectively differ with respect to a relevant characteristic or are erroneously believed to do so. Not before the early 1970s did economists come up with a formalization of the latter idea. Virtually at the same time, Arrow (1973) and Phelps (1972) developed a model of statistical discrimination showing that in the labor market, under incomplete information, differential judgments by employers can lead to wage differentials of employees from different social groups. For example, given employers' beliefs that qualified workers are less likely to be found among blacks than among whites and given an employer incurs a cost from determining a potential employee's actual qualification, skin color serves as "cheap" proxy for qualification. Thus, skin color becomes the characteristic upon which an employer decides whether or not to hire a particular person at a particular wage. Allport (1954) notes that if a person's belief is based on wrong information and would be revised in the light of new evidence, this person makes an error of prejudgment. A prejudgment becomes a prejudice, however, if a person does not revise his or her belief given new information. Arrow (1973) suggests cognitive dissonance as an explanation for people being prejudiced (i.e., having incorrect and sticky beliefs). According to the theory of cognitive dissonance (Festinger 1957), people tend to bring their beliefs into accord with their behavi-

[1] I would like to thank Ben Jann, Andreas Diekmann, Roger Berger, and Marc Höglinger for helpful comments and suggestions. I gratefully acknowledge the very helpful comments and suggestions of two anonymous reviewers. This work was partly supported by the Swiss National Science Foundation [grant number 100017_124877].

or. Hence, someone who has engaged in discriminatory behavior is reluctant to adjust his or her beliefs in light of new evidence as this would imply admission of wrongdoing. Theories of social identity (Tajfel 1981) and self-categorization (Turner et al. 1987), on the other hand, posit that the degree of one's identification with a social group determines whether a social interaction is perceived as interpersonal or inter-group. In the latter case, behavior is led by in-group norms rather than by a deliberate appraisal of the interaction at hand. Moreover, it has been suggested that out-group discrimination is a means to demonstrate one's loyalty and trustworthiness to in-group members (Posner 2000) and to gain intra-group social status (McAdams 1995). Thus, discrimination can be based on taste, wrong information, prejudice, and on social norms.

Empirical studies have repeatedly identified more or less subtle forms of ethnic and racial discrimination in job, housing, credit, and consumer markets (for a review, see Pager & Shepherd 2008). The most common approach has been to investigate outcome inequalities between social groups by means of statistical analysis. After controlling for all factors explaining between-group differences in, for instance, wages or loan conditions, the main effect of ethnicity or race on outcomes can be attributed to discrimination. However, objections that such evidence could be spurious because some unmeasured causes have not been accounted for in the model are hard to accommodate. Moreover, the fact that ethnicity and race are invariant within the same person limits the possibilities for causal inference. Ethnicity and race can be varied experimentally, though. In field experiments, researchers send pairs of applications on behalf of two bogus candidates seemingly applying for the same job or property. The two applications differ in one characteristic from which a naive employer or letting agent, respectively, could derive the candidate's ethnicity or race. Observed differences in response rates are ascribed to discrimination. Although limited in generalizability, the validity of results from such field experiments is particularly high. There is no self-selection of participants into experimental conditions; it is actual and not stated behavior that is being measured and participants are not aware of the fact that they are taking part in a scientific study about discrimination (for reviews, see Riach & Rich 2002; Pager 2007).[2] Klink & Wagner (1999) conducted 14 field experiments in which they varied the ethnicity of a confederate or fictitious person in diverse social interactions such as housing searches, help requests, restaurant reservations, or hitchhiking. Their results give evidence of ethnic discrimination against the Turkish minority in Germany. In a recent field experiment, Doleac & Stein (2010) sold iPods (i. e., portable multimedia players) through local online classified advertisements throughout the United States. They varied the skin color (black vs. white) of the hand holding the product on an advertisement photograph and find considerably lower sales for seemingly black sellers than for white sellers. Further evidence of buyers being reluctant to include their real names in the e-mail correspondence or to agree on mail delivery and long-distance payment when dealing with seemingly black sellers suggests that black sellers are perceived to be less trustworthy than white sellers.

In online markets, trust problems arise from the fact that buyers have to send the money before the seller ships the merchandise. Unlike the local online markets where Doleac & Stein (2010) conducted their experiment, many online markets implement a feedback forum which allows buyers and sellers to rate one another after a transaction (Kollock 1999). Sellers receive positive and negative ratings, which determine their reputation, and buyers take sellers' reputation into account when deciding to bid for an auctioned product. It can be shown theoretically that sellers entering the market have to invest in building their reputation by allowing a discount, and therefore, sellers' reputations must be correlated with selling prices (Shapiro 1983; Ockenfels 2003). The empirical evidence largely confirms that positive (negative) ratings positively (negatively) influence sales and selling prices (for reviews, see Bajari & Hortacsu 2004; Resnick et al. 2006). However, an interesting question is whether sellers entering the online market with an empty feedback record can signal their trustworthiness by other means. The theoretical argument is based on signaling theory, most prominently introduced into the social sciences by Spence (1973) and expanded on trust research by Voss (1998), Bacharach & Gambetta (2001), Raub (2004) and Diekmann &

[2] Field experiments investigating market discrimination are also known as audit studies. In general, field experiments, as opposed to laboratory experiments, denote experimental investigations in a natural environment. The term *nonreactive* is used, if, in addition, subjects do not know that they participate in a field experiment (e. g., Diekmann 2007: 630). For nonreactive studies on ethnic discrimination using non-experimental designs see, for instance, Kalter (1999) and Levitt (2004).

Przepiorka (2010). The formal argument derived from game theoretic models is that only sellers with a long-term interest to stay in the market have an incentive to behave cooperatively, and therefore afford to invest in costly signals. For example, producers of experience goods advertise their products even though this advertising cannot add valuable information prior to purchase. Nelson (1974) argues that consumers choosing products that have been advertised could be confident of the product's quality since only repeated sales would compensate the producer for the initial expenses on advertisement. Hence, advertisement can be a costly signal that allows buyers to distinguish between sellers of high and low quality products (Diekmann & Przepiorka 2010).

The present study investigates whether members of the Turkish minority in Germany are discriminated against in an online market. Given the previous evidence, I expect to find more sold items and higher prices attained by German sellers than by Turkish sellers. In online markets, social interactions are less likely to be observed by one's peers. Therefore, ethnic discrimination, if present, will less likely be induced by social norms. However, potential buyers could perceive Turkish sellers as less trustworthy than German sellers and discriminate against them based on these beliefs. Lower perceived trustworthiness clearly implies information-based discrimination. However, since buyers' actual beliefs about Turkish sellers' trustworthiness in online markets is unobserved, information-based and taste-based discrimination cannot be distinguished. The study also investigates whether market entrants can credibly signal their trustworthiness by an initial costly investment. According to the theoretical considerations outlined in the previous paragraph, I expect both German and Turkish sellers displaying a costly signal to be bought from with higher probability and to attain higher prices than sellers not displaying a costly signal. Moreover, if potential buyers discriminate against Turkish sellers because they believe Turkish sellers to be less trustworthy, Turkish sellers displaying the costly signal should be more successful and attain higher prices than Turkish sellers who do not display it. If, however, buyers discriminate against Turkish sellers based on taste, Turkish sellers should perform on a similar level, irrespective of whether or not they send a costly signal. These hypotheses are tested empirically in a field experiment described in the third section. Shohat & Musch (2003) conducted a similar field experiment to investigate discrimination of the Turkish minority in Germany. As I arrive at different conclusions from the secondary analysis of their data the next section reproduces their findings and the new evidence. The concluding section discusses the findings.

2. Experiment 1: Discrimination in an online market – a secondary analysis

Shohat & Musch (2003) conducted a field experiment in order to test whether members of the Turkish minority in Germany are subject to discrimination in the German online market eBay.de. In their experiment, one seller was given a Turkish alias (mehmet.orgum) and another seller was given a German alias (michael.ottersbach). In a two weeks period, both sellers sold the same sat of 30 different DVDs. The Turkish seller auctioned the first subset of 15 DVDs in the first week and the other 15 DVDs in the second week. The German seller did it the other way round. In other words, 30 different movie titles were sold each one by a seemingly Turkish and a German seller, but in different succession (cf. also Shohat 2001). Contrary to their hypothesis, Shohat & Musch (2003) found no statistically significant difference between German and Turkish sellers in the average number of page views, the average number of bidders, or the average selling price. From these findings the authors conclude that, unlike in other studies, they could find no direct evidence for ethnic discrimination in the online market under study.[3] In what follows, I first argue that the heterogeneity induced by the different movie titles and the fact that the two different subsets of movies were auctioned by the Turkish and the German seller in two different weeks have to be accounted for in the data analysis. Then I show that buyers do discriminate between German and Turkish sellers – but only if there is a lack of alternative offers by other sellers.

Shohat (2001: 100) provides the raw data from the experiment. These data allow replicating the findings reported in Shohat & Musch (2003). The results section in Shohat & Musch (2003) reports t-statistics from mean-comparison tests for the outcome variables mentioned in the previous para-

[3] Their exploratory data analysis, however, revealed that German sellers received their highest bid significantly earlier than Turkish sellers. This time difference would indicate that "buyers turn to auctions of Turkish sellers only when the same item cannot be bought from a German seller at a favorable price" (Shohat & Musch 2003). Unfortunately, the authors did not test this conjecture.

Table 1 Fixed effects regression models

	M1		M2		M3	
	Coef	SE	Coef	SE	Coef	SE
German seller	0.60	0.84	2.74*	1.03	0.93	0.72
Week 2	2.00*	0.84	2.34*	0.78	1.07+	0.54
Competition			–0.48	0.31	0.21	0.22
German 6 Compet.			–0.34*	0.15	–0.24*	0.10
Retail price					0.11*	0.03
Movie title dummies, F-statistic	58.78*		44.27*		2.47*	
N	60		60		60	
Adj. R^2	0.989		0.991		0.940	

Notes: Fixed effects regression models with selling price (in DM) as the dependent variable in the first two models and number of bidders in model three. All models account for movie title fixed effects (+ p<0.10, * p<0.05). The analysis is based on data from Shohat (2001: 100).

graph. In laboratory experiments in which subjects are randomized into experimental groups, differences in the means of the outcome variable can be fully attributed to the treatment (or stimulus). In the field experiment under discussion randomization of subjects (i. e., potential buyers in the eBay market) was not possible. Buyers always have the possibility to choose sellers contingent on their characteristics. Self-selection into experimental treatments, however, is not the only problem many field experiments must cope with. As time goes by, especially in such a dynamic environment like an online market, the context, and with it unobserved confounding factors, can change considerably. If information about subjects' characteristics or the context is not available, it is difficult to compensate for the lack of randomization and time-induced heterogeneity in a statistical model. The data provided in Shohat (2001: 100) though allow accounting for some potentially confounding factors.

Table 1 contains coefficient estimates from OLS regressions models each with 29 dummy variables accounting for movie title fixed effects (e.g., Gujarati 2003). The dependent variable in models *M1* and *M2* is the selling price in DM ($\bar{x}=29.57$, $\tilde{x}=31$, $sd=7.05$) and in model *M3* the number of bidders ($\bar{x}=7.68$, $\tilde{x}=7.5$, $sd=2.39$). The estimation of model *M1* shows that movie titles and week of experiment account for nearly the complete variance in the dependent variable ($adj.R^2=0.989$). Moreover, it confirms the finding by Shohat & Musch (2003) that buyers do not pay more or less depending on the sellers' alleged ethnicity. The latter, however, should be questioned because buyers can choose the seller they like. If, for instance, self-selection depends on nepotism such that German buyers prefer buying from German sellers and non-German buyers find offers by Turkish sellers more attractive, the actual treatment effect could be blurred by this unobserved heterogeneity. The data, however, do not allow controlling buyer characteristics in a statistical model. The estimates in model *M2*, on the other hand, support the conjecture that unobserved heterogeneity is an issue. The model contains two additional explanatory variables accounting for the number of alternative offers from other sellers, i. e., seller competition ($\bar{x}=4.88$, $\tilde{x}=3$, $sd=5.05$), and its interaction with the focal seller's ethnicity.[4] While seller competition does not have a significant main effect, the main effect of the focal seller's ethnicity shows that given alternative offers by other sellers are lacking, buyers pay 2.7 DM more if the seller is German than if the seller is Turkish. The significant interaction effect indicates that the ethnicity main effect decreases as the number of alternative offers increases. Model *M3* confirms that the more other sellers offer the same item, the less potential buyers are attracted by the German seller. Hence, the actual difference in what buyers pay for an offer from a German and a Turkish seller is more pronounced if alternative offers are lacking and the bidding competition is high.

[4] The number of alternative offers from other sellers is a variable provided in Shohat (2001: 100). The variable measures the number of alternative offers of the same movie title one day before the auction of the DVD offered in the experiment ended.

3. Experiment 2: Discrimination and costly signaling

Similar to Shohat & Musch (2003), I implemented a field experiment on eBay.de which tried to meet some of the shortcomings of the previous study. I established 44 online seller accounts which varied in two characteristics. Firstly, half of the accounts were opened with aliases comprising a Turkish first name (e.g., Ali, Ercan, Murat, Ozkan, etc.) and the other half with aliases comprising a German first name (e.g., Axel, Christian, Helmut, Steffen, etc.).[5] Secondly, half of the sellers had a verified identity. Obtaining a verified identity is costly in terms of time and money. A seller can verify his or her identity by filling out a form that includes the applicant's real name, address, and passport or identity card number, and bringing the form to the municipal administration for notarization at a cost of 15 Euro. About ten days after eBay receives the form, the verified identity icon appears next to the seller's alias on the seller's profile page as well as on every item page. I auctioned 84 identical 1 GB digital memory cards which are widely used in small digital devices like cameras, mobile phones, and handheld computers. The criteria for the choice of the market for SD memory cards were the homogeneity, the low complexity, and the convenience of handling of the merchandise. The retail price of such a memory card was, at that time, about 20 Euro. The cards were posted at an initial price of 1.99 plus shipping and handling charges of around 9.90 Euro. The title, description, and picture of the item were identical across the 84 posts. However, to increase the appearance of different sellers, the order of words in the title and sentences in the description was sometimes altered. Figure 1 shows an offer that was posted on the market in the field experiment. The seller has a German alias (christian0935), zero ratings, and the verified identity icon appears next to this information (see „Angaben zum Verkäufer" on the right hand side of Figure 1).

I implemented a randomized block design for the experiment. One block consisted of a pair of items which were put online at about the same time. The two items of a pair differed in at least one characteristic – either in the alias (Turkish vs. German), in the identity (verified vs. not verified), or in both. Table 2 summarizes the experimental conditions. The randomized block design is especially suited for experiments in which one expects more variability between than within homogeneous subgroups. In order to account for the between-pair variability, I estimated fixed effects regression models. Another advantage of this approach is that within-pair constant unobservables cancel out in the estimation process. In addition, as the matched pairs are randomized over time model misspecification due to unobserved heterogeneity becomes less of an issue.

After a short pretest, the experiment started on October 30 and lasted until December 21, 2006. One pair of items was put online every day. In order to eliminate potential sources of reactivity, I started the item pairs at different times in the evening. Unfortunately, the marketplace restricted new sellers, allowing offers to be posted only at quarters of an hour. Since I wanted to avoid the two items of a pair to start at exactly the same instant, one item started 15 minutes after its pair. Another restriction was that an offer had to last for at least three days. Because the time restriction made it impossible to offer all item pairs consecutively during the desired time frame, three pairs of items were offered during overlapping periods. Figure 2 illustrates the described experimental design and the scheduling. The four letters A, B, C, and D denote the four treatment conditions (see Table 2).

Despite the efforts at camouflaging the auctions, eBay closed down some of the seller accounts eleven days into the experiment, suspecting some sort of fraud. The problem was resolved through direct communication with eBay and the experiment could be continued with a delay of three days (see Figure 2).[6] I had to cope with a further restriction

[5] At the time the experiment was conducted, eBay.de allowed every person to open up to six user accounts with different aliases. In order to open a user account on eBay.de, a valid credit card or direct debit account was required. For this reason, I asked several friends and colleagues to provide me with their credentials. The alias is one's user name on eBay. It can be freely chosen, but must not offend common decency and has to be unique in the whole market platform. People usually choose a fancy name or their first name followed by a number as their alias. Aliases as such are not legally binding, but the person registered under a certain alias is responsible for the activities of his or her online alter ego. I chose aliases comprising either a German or Turkish first name because I wanted potential buyers to believe that they are trading with a German or Turkish seller, respectively.

[6] eBay closed down six accounts which had all been opened with the credentials of one person. There was no apparent reason for this, and I have never learned why eBay closed exactly these accounts and all other accounts remained unaffected. Note that during this study I never engaged in illegal practices, deceived eBay users or violated other rules established by eBay. The experiment was conducted within the constraints determined by the platform

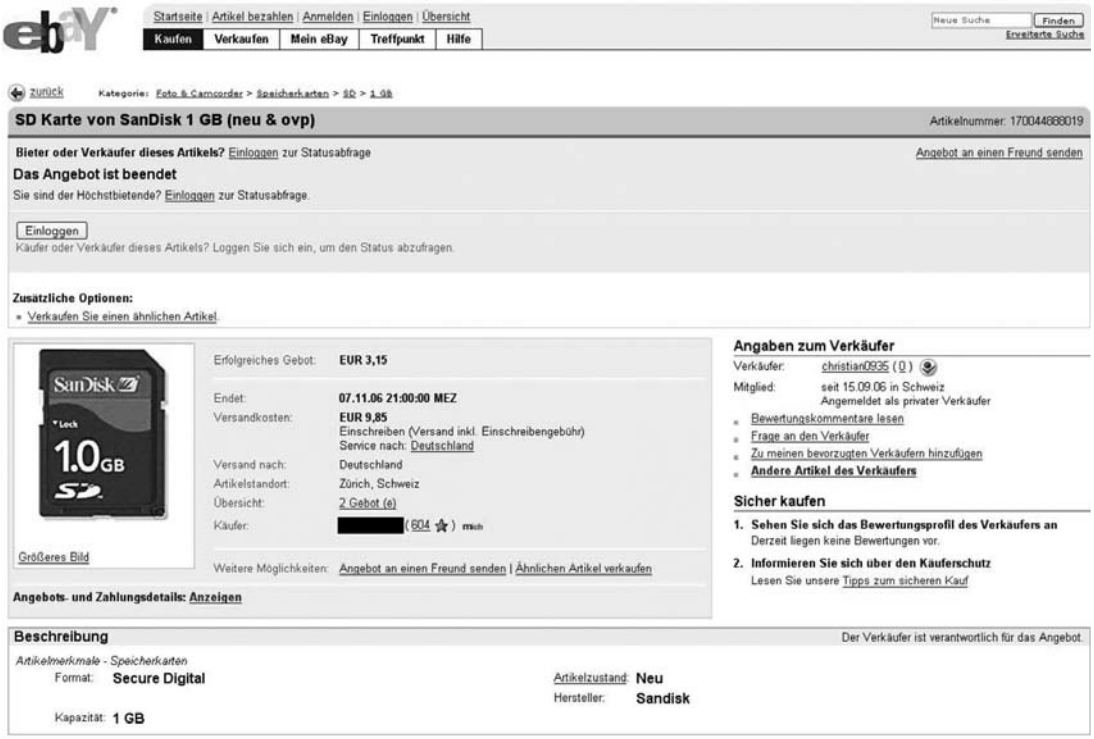

Fig. 1 Item offered in the field experiment

Table 2 Number of items posted by treatment

	Seller's Alias		
Verified ID	German	Turkish	Total
yes	A 21	B 21	42
no	C 21	D 21	42
Total	42	42	84

which resulted from the fact that buyers could leave feedback after a completed transaction. It took about ten days from the moment an offer ended until the buyer received the item and could be expected to leave feedback. A seller account with an entry in the feedback history would clearly violate the experimental conditions. Therefore, it was possible to use the same seller account at most three times in a row. While an item was active in the market, I often received e-mails from potential buyers. The e-mails I received were either inquiries about shipping conditions or suggestions to make the deal outside of the market place. I did not reply to e-mails until an offer had ended. Once an offer had ended and had received at least one bid, the highest bidder received my bank details and was asked to transfer the money within the next few days. Not until this point potential buyers had the possibility to learn who the real person was with whom they were trading. As soon as the money arrived on my bank account, I shipped the merchandize to the address I had been given by the buyer. Neither the bidders nor the buyers ever learned that they had participated in a scientific experiment.

Table 3 contains descriptive statistics of the number of sales and mean selling prices (incl. shipping costs). 77 of the 84 items were sold at an average price of 13.86 Euro. The differences between mean prices across experimental conditions are rather small and statistically insignificant (Wilcoxon signed-rank test for within pair differences in selling price: ID verified vs. not verified, $n=23$, $z=1.521$, $p=0.128$; Alias German vs. Turkish, $n=22$, $z=$

provider and all buyers received the product they had paid for.

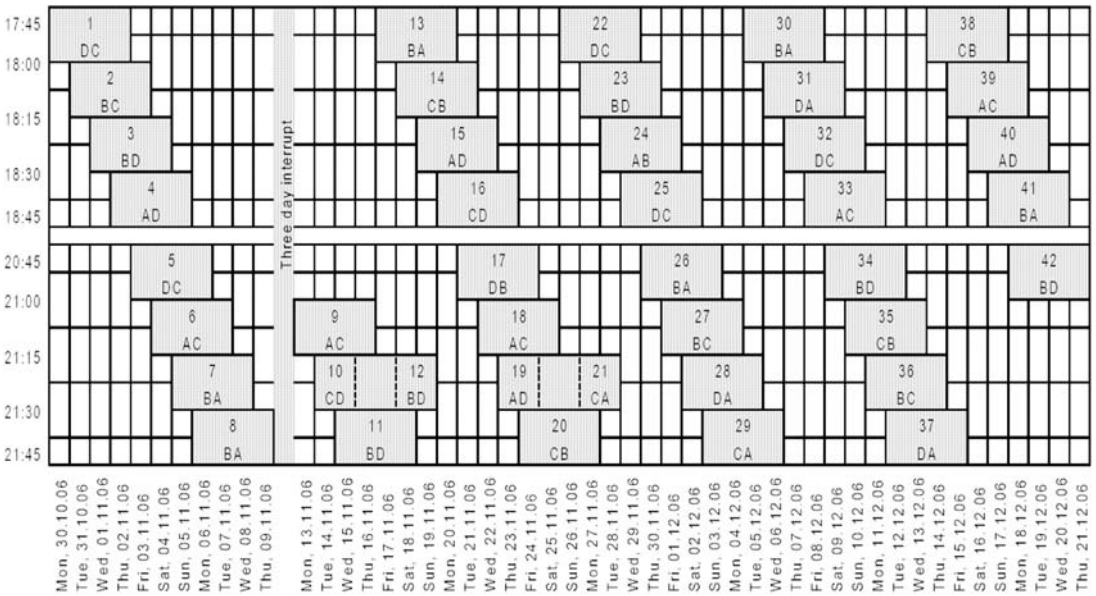

Fig. 2 Design and time schedule of field experiment

Table 3 Summary statistics by treatment

	Seller's Alias		
Verified ID	German	Turkish	Total
	20 (1)	20 (1)	40 (2)
yes	13.71	14.14	13.92
	(0.36)	(0.66)	(0.37)
	18 (3)	19 (2)	37 (5)
no	14.33	13.30	13.80
	(0.53)	(0.37)	(0.33)
	38 (4)	39 (3)	77 (7)
Total	14.00	13.74	13.86
	(0.31)	(0.39)	(0.25)

Notes: Number of items sold (not sold), mean selling prices (incl. shipping costs in Euro), and standard deviations of the mean in parentheses.

0.86, $p = 0.39$). The same is true for the number of sales (Fisher's exact test: ID verified vs. not verified, $p = 0.352$; Alias German vs. Turkish, $p = 0.669$). In both cases H_0 ($\delta = 0$) was tested against the two-sided alternative.

Table 4 shows regression estimates of selling price on the main treatment and control variables. Each model contains 41 dummy variables absorbing the between-pair variability. Model 1 estimates treatment main effects and a two-way interaction. The coefficient estimate for the verified identity has the expected direction and is statistically significant at the 10 % level. It turns out, however, that it makes a difference whether an item was posted first or second in a pair. Model 2 includes a dummy variable accounting for this fact and the coefficient estimate is positive and significantly different from zero. The inclusion of this factor changes the magnitude of the other coefficient estimates suggesting that treatments were not perfectly randomized across the sequence within item pairs.

Out of the 84 items auctioned in the experiment, 7 items were not sold. Thus, potential buyers' willingness to pay for these items is censored at a starting bid of approximately 12 Euro and the actual values remain unobserved. However, using a truncated sample (i. e., sold items only) or the censored values in the data analysis yields inconsistent coefficient estimates (Long 1997: 201–203). This problem can be tackled by the use of other estimation techniques. The estimation of Model 3 in Table 4 is based on a censored normal regression (Greene 2002: 764–768). The coefficient estimates do not deviate much from the estimates in Model 2, but are statistically significant. The results can be interpreted as follows. Having a verified identity or a German alias increases selling prices by 1.09 or 0.87 Euro on average, respectively. The negative interaction effect suggests, however, that having both

Tabelle 4 Fixed effects regression models of selling price

	M1		M2		M3	
	Coef	SE	Coef	SE	Coef	SE
Verified ID	1.25+	0.64	0.94	0.61	1.09*	0.41
German seller	0.66	0.64	0.57	0.59	0.87*	0.40
Verified × German	–1.08	0.87	–0.86	0.81	–1.02+	0.56
First in pair			0.87*	0.35	0.77*	0.24
Const.	14.72*	1.67	14.81*	1.54	11.22*	0.81
Pair dummies, F(41,38) statistic	2.79*		3.31*		6.67*	
N	77		77		84	
Adj./pseudo R^2	0.50		0.57		0.35	

Notes: OLS and censored normal regression models with selling price (incl. shipping cost, in Euro) as the dependent variable. All models account for item pair fixed effects. Models 1 and 2 are estimated on the truncated sample. Model 3 is a censored normal regression ($^+$ $p < 0.10$, * < 0.05).

does not further increase a seller's trustworthiness. In other words, identity verification pays off for a Turkish seller, but being German, a verified ID does not further increase a seller's profits. The interaction effect is only significant at the 10 % level, though.

There are several reasons for why, in this case, the estimates from the censored normal regression model (*M3*) may be misleading. Firstly, unlike in OLS, estimates of the censored normal regression are inconsistent if the assumptions of homoscedasticity and normally distributed errors do not hold (Greene 2002: 768–773). Moreover, the model estimation is based on assumptions which are only known to hold for large samples. Given the sample size of 84 cases, these assumptions are likely to be violated. Secondly, while the mere numbers in Table 3 tend to suggest lower sales if sellers' identity is unverified, they seem not to correspond to the ethnicity main effect and the interaction effect obtained in Model 3. Finally, the number of censored cases is relatively small. Usually, a small number of censored cases will decrease the bias of the OLS coefficients estimated on the truncated sample. Since only 8.3 % of the cases are censored, the bias in Model 2 will be rather small. Therefore, I consider the coefficients estimated in Model 2 as the more reliable ones and will base my conclusions on the results obtained in Model 2 only.

4. Discussion and conclusions

The paper reports results from two field experiments designed to identify possible ethnic discrimination on a German internet auction platform. The second experiment also tests whether costly signals can help to overcome trust problems between buyers and sellers in such markets. The first experiment was conducted by Shohat & Musch (2003); the secondary analysis of their data reveals that their conclusion of no evidence for ethnic discrimination may be incomplete. Controlling for possibly confounding factors in a multiple linear regression model shows that ethnicity indeed has a statistically significant and substantial effect on selling price. In particular, if the number of similar offers by other sellers is small, seemingly German sellers attain higher selling prices than Turkish sellers. This indicates that ethnic discrimination becomes apparent when stakes rise due to a lack of competition. However, the data provided in Shohat (2001: 100) do not allow controlling for every potential source of bias. That is to say, the secondary analysis presented here could likewise suffer from omitted-variable bias. It shows, however, that if a randomization of subjects in experimental treatments is not possible, one should either be able to control for potentially confounding factors in a statistical model, or devise a research design that better copes with possible confounds. The second field experiment replicated the first, albeit with a more elaborate design. Forty-two pairs of items which differed in at least one seller characteristic (either ethnicity, verified identity, or both) were put online at different occasions over a period of 50 days. Thus, experimental treatments

were randomized over time and unobservable confounding factors were held constant within item-pairs. Results of this second experiment provided no conclusive evidence either for ethnic discrimination or for the signaling hypothesis.

While it is good news, of course, to find no evidence for ethnic discrimination on this online market, there are several reasons to believe that under slightly different conditions we might have observed both discrimination and a signaling effect. Firstly, in the second experiment, potential buyers also had the possibility to choose a seller based on his observable characteristics. Hence, discrimination against German sellers and favoritism of Turkish sellers could have canceled out discrimination against Turkish sellers. However, the buyers' surnames are predominantly German (not reported) and therefore, such a compensating effect appears rather unlikely. Secondly, for technical reasons the merchandise had to be shipped from Switzerland. The information about the item location is displayed on every offer page (see „Artikelstandort" in Figure 1). The results of the second experiment could be explained, if one assumed that potential buyers were predominantly concerned with the fact that the seller shipped from Switzerland, completely neglecting the other seller characteristics. However, it is difficult to think of a reason why this should be the case. Finally, at the time of the second experiment approximately 200,000 offers of similar memory cards were posted on eBay.de by about 3,500 different sellers. The fact that only 49 % of these offers resulted in a sale suggests an oversupply of the merchandize and consequently low selling prices (cf. Przepiorka 2010). In fact, the average selling price was 13.86 Euro, significantly below the retail price of about 20 Euro. Hence, if potential losses are not very high, buyers might consider information provided about the seller less carefully. I would expect a replication of this study with a more expensive product to give evidence for both ethnic discrimination and a signaling effect of the verified identity. In this case indeed, it would be interesting to see whether trust-enhancing costly signals are necessary and sufficient to compensate the negative effect of ethnic discrimination. If so, this would indicate that ethnic discrimination in online markets is information-based rather than taste-based.

References

Allport, G.W., 1954: The Nature of Prejudice. New York: Addison-Wesley.

Arrow, K., 1973: The Theory of Discrimination. Pp. 3-33 in: O. Ashenfelter & A. Rees (eds.), Discrimination in Labor Markets. Princeton: Princeton University Press.

Bacharach, M. & D. Gambetta, 2001: Trust in Signs. Pp. 148–184 in: K.S. Cook (ed.), Trust in Society. New York: Sage Foundation.

Bajari, P. & A. Hortacsu, 2004: Economic Insights from Internet Auctions. Journal of Economic Literature 42: 457–486.

Becker, G.S., 1957: The Economics of Discrimination. Chicago: University of Chicago Press.

Diekmann, A., 2007: Empirische Sozialforschung. Reinbek bei Hamburg: Rowohlt.

Diekmann, A. & W. Przepiorka, 2010: Soziale Normen als Signale. Der Beitrag der Signaling-Theorie. Pp. 220–237 in: G. Albert & S. Sigmund (eds.), Soziologische Theorie kontrovers. Sonderheft der Kölner Zeitschrift für Soziologie und Sozialpsychologie. Wiesbaden: VS.

Doleac, J.L. & L.C.D. Stein, 2010: The Visible Hand: Race and Online Market Outcomes. SIEPR Discussion Paper 09–015, Stanford University.

Festinger, L., 1957: A Theory of Cognitive Dissonance. Stanford: Stanford University Press.

Greene, W.H., 2002: Econometric Analysis. New Jersey: Prentice Hall.

Gujarati, D.N., 2003: Basic Econometrics. New York: McGraw-Hill.

Kalter, F., 1999: Ethnische Kundenpräferenzen im professionellen Sport? Der Fall der Fußballbundesliga. Zeitschrift für Soziologie 28: 219–234.

Klink, A. & U. Wagner, 1999: Discrimination against Ethnic Minorities in Germany: Going Back to the Field. Journal of Applied Social Psychology 29: 402–423.

Kollock, P., 1999: The Production of Trust in Online Markets. Advances in Group Processes 16: 99–123.

Levitt, S.D., 2004: Testing Theories of Discrimination: Evidence from Weakest Link. Journal of Law and Economics 47: 431–453.

Long, J.S., 1997: Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks: Sage.

McAdams, R.H., 1995: Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination. Harvard Law Review 108: 1003–1084.

Nelson, P., 1974: Advertising as Information. Journal of Political Economy 82: 729–754.

Ockenfels, A., 2003: Reputationsmechanismen auf Internet-Marktplattformen. Zeitschrift für Betriebswirtschaft 73: 295–315.

Pager, D., 2007: The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future. ANNALS of the American Academy of Political and Social Science 609: 104–133.

Pager, D. & H. Shepherd, 2008: The Sociology of Discrimination: Racial Discrimination in Employment, Hou-

sing, Credit, and Consumer Markets. Annual Review of Sociology 34: 181–209.
Phelps, E.S., 1972: The Statistical Theory of Racism and Sexism. American Economic Review 62: 659–661.
Posner, E.A., 2000: Law and Social Norms. Cambridge (MA): Harvard University Press.
Przepiorka, W., 2010: Buyers Pay for and Sellers Invest in Reputation. More Evidence from eBay. Working Paper, ETH Zurich Sociology.
Raub, W., 2004: Hostage Posting as a Mechanism of Trust: Binding, Compensation, and Signaling. Rationality and Society 16: 319–365.
Resnick, P., R. Zeckhauser, J. Swanson & K. Lockwood, 2006: The Value of Reputation on Ebay: A Controlled Experiment. Experimental Economics 9: 79–101.
Riach, P.A. & J. Rich, 2002: Field Experiments of Discrimination in the Market Place. Economic Journal 112: F480–F518.
Shapiro, C., 1983: Premiums for High Quality Products as Return to Reputation. The Quarterly Journal of Economics 98: 659–680.
Shohat, M., 2001: Online-Auktionen als Forschungstool: Ein Feldexperiment zur ethnischen Diskriminierung. Diplomarbeit, Rheinische Friedrich-Wilhelms-Universität, Psychologisches Institut.
Shohat, M. & J. Musch, 2003: Online Auctions as a Research Tool: A Field Experiment on Ethnic Discrimination. Swiss Journal of Psychology 62: 139–145.
Spence, M.A., 1973: Job Market Signaling. Quarterly Journal of Economics 87: 355–374.
Tajfel, H., 1981: Human Groups and Social Categories: Studies in Social Psychology. Cambridge: Cambridge University Press.
Turner, J.C., M.A. Hogg, P.J. Oaks, S.D. Reicher & M.S. Wetherell, 1987: Rediscovering the Social Group: A Self-Categorization Theory. Oxford: Blackwell.
Voss, T., 1998: Vertrauen in Modernen Gesellschaften. Eine Spieltheoretische Analyse. Pp. 91–129 in: R. Metze, K. Mühler & K.-D. Opp (eds.), Der Transformationsprozess: Analysen und Befunde aus dem Leipziger Institut für Soziologie. Leipzig: Leipziger Universitätsverlag.

Autorenvorstellung

Wojtek Przepiorka, geb. 1977 in Pulawy (Polen). Studium der Soziologie, Informatik und VWL in Bern. Promotion 2009 in Zürich. Seit 2010 John Fell Fund Research Fellow in Experimental Social Sciences am Nuffield College, University of Oxford.
Forschungsschwerpunkte: Entstehung sozialer Kooperation, Verhaltensspieltheorie, experimentelle Methoden in den Sozialwissenschaften.
Wichtigste Publikationen: Der Beitrag der Signaling-Theorie zur Erklärung sozialer Normen (mit A. Diekmann), in: G. Albert & S. Steffen, (Hrsg.), Soziologische Theorie kontrovers. Sonderheft 50 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, Wiesbaden 2010; zuletzt in dieser Zeitschrift: Die Präventivwirkung des Nichtwissens im Experiment (mit A. Diekmann und H. Rauhut), ZfS 40, 2011: 74–84.