

THE VISIBLE HAND: RACE AND ONLINE MARKET OUTCOMES*

Jennifer L. Doleac and Luke C.D. Stein

We examine the effect of race on market outcomes by selling iPods through local online classified advertisements throughout the US. Each advertisement features a photograph including a dark or light-skinned hand, or one with a wrist tattoo. Black sellers receive fewer and lower offers than white sellers, and the correspondence with black sellers indicates lower levels of trust. Black sellers' outcomes are particularly poor in thin markets (suggesting that discrimination may not 'survive' competition among buyers) and those with the most racial isolation and property crime (consistent with channels through which statistical discrimination might operate).

Economic outcomes in the US are highly correlated with race but the causal mechanisms underlying these correlations are not well understood. In particular, it remains unclear how much of the correlation is due to discrimination and how much is due to other characteristics that are correlated with race, such as education.

There is an extensive literature on the effect of race on market outcomes, focusing on both labour and goods (including housing) markets. However, 'discrimination' is not a monolithic phenomenon, and it is important to investigate its extent and causes in a variety of different contexts (across which they may differ). Our study asks the following question: When the typical person engages in a consumer transaction (usually as a buyer) does he or she try to avoid dealing with minority sellers, and does she treat minority sellers differently?

This is an important question for at least two reasons. First, a large amount of commerce takes place through this kind of one-time consumer transaction. Second, discrimination by consumers may in fact underlie other forms of discrimination. For example, if white customers prefer not to deal with black sellers, retailers might avoid hiring black salespeople. Becker (1971) identifies discrimination by consumers, employers and fellow workers as the three potential sources of the racial disparity in labour market outcomes, and Nardinelli and Simon (1990) note that in a relatively competitive labour market like the US, consumer discrimination is the most likely cause of the persistent disparity.¹ However, the resulting labour market discrimination is difficult in practice to distinguish from lower ability because both affect observed productivity.

*Corresponding author: Jennifer L. Doleac, Frank Batten School of Leadership and Public Policy, University of Virginia, 235 McCormick Road, P.O. Box 400893, Charlottesville, VA 22904, USA. Email: jdoleac@stanford.edu.

We are grateful to B. Douglas Bernheim, Nicholas Bloom, Caroline Hoxby, Jörn-Steffen Pischke and several referees for useful advice and guidance, and have also benefited from conversations with participants in several Stanford seminars, the San Francisco Federal Reserve Bank's Applied Micro Summer Conference and the University of Chicago Experimental Economics Lunch. Brandon Wall made important contributions to our experimental design and piloting. We appreciate the generous support of the George P. Shultz Dissertation Support Fund.

¹ Altonji and Blank (1999) summarise the theory and evidence regarding race and the labour market in their *Handbook of Labor Economics* chapter, and document the persistent black–white gap in earnings, labour participation and education.

Our study contributes to the literature by examining the effect of consumer discrimination on sellers' market outcomes using a field experiment that rigorously isolates the effect of sellers' skin colour. This complements existing evidence of discrimination in labour and housing markets which – although striking – does not imply the existence of discrimination in common consumer transactions (although the latter could be one explanation of labour market discrimination). Indeed, we might expect that individuals are more likely to indulge in their taste for anti-black animus when renting a house or hiring an employee than when engaging in the limited interaction of a consumer transaction.

We posted classified advertisements offering an iPod Nano portable digital music player for sale on several hundred locally focused websites throughout the US, and analyse here the effect of the seller's skin colour on several outcomes of interest. Notably, we signal skin colour using photographs instead of one of the more problematic treatments used in previous discrimination studies.² By including a photograph of a dark-skinned (black) or light-skinned (white) hand holding the item, we were able to randomly vary the apparent race of the seller while fixing other advertisement and market characteristics.³ We also compare the effect of race with that of a social signal that can be communicated through the appearance of a seller's hand: a wrist tattoo. Tattooed sellers are likely statistically discriminated against for many of the same reasons as black sellers, so – in addition to providing general context for interpreting the magnitude of the black–white differences we observe – this third group of sellers can serve as a 'suspicious' white control group.

The market in which we run our experiment facilitates our investigation of consumer discrimination in multiple ways: buyers have no reason to make offers that they do not anticipate ending in a transaction. They anticipate having to meet a seller to complete the transaction – perhaps on the seller's terms – with the non-trivial possibility of deception or theft. Thus, trust plays a key role in the interactions we observe. These are characteristics of many (labour and goods) market transactions that may be less present in the decision to call back a job applicant, bid in an online auction or make a purchase guaranteed by a third party (such as from a store where the salesperson is merely an employee).

An additional attractive feature of our experimental setting is that the local focus of online classified advertisements allows us to analyse regional differences as well as variation by local economic and demographic characteristics. By examining whether black–white outcome differences vary systematically with market characteristics, we can shed some light on the type of discrimination we are observing. As these characteristics are not randomly assigned, causal inference is not entirely clear cut, but considering numerous local markets allows us to assess whether patterns of observed discrimination are consistent with those predicted by various theories.

² Ravina (2008) and Pope and Sydnor (2011) have also studied race through photographs in online markets, although in non-experimental settings where non-random samples and unobservable characteristics are a concern.

³ Skin colour is clearly highly correlated with race. We believe that discrimination based on skin colour is of primary interest when people discuss racism, but there are surely many other relevant components of race that our study ignores.

Given the challenges of adequately controlling for unobservable characteristics that may be correlated with race, we are, unsurprisingly, not the first to use an experimental approach. Riach and Rich (2002) offer a useful survey of field experiments designed to assess discrimination. Actor-based audit studies (Ayres and Siegelman, 1995; Neumark *et al.*, 1996) – in which actors apply for jobs, consider housing, or negotiate sales – attempt to match different race candidates on as many dimensions as possible but the match quality will never be perfect. In addition, these studies are typically not double blind and actors' awareness of the object of study and experimental design may bias the results.

One similar study of discrimination by buyers and sellers in a goods market is List (2004). As in more typical audit studies, List uses black and white agents who directly signal their race – in this case to dealers at a baseball card show. The agents are drawn from the general population of show attendees, so they inevitably differ along characteristics other than race. Despite these non-race differences, they find disparities between offers made to black and white agents consistent with what we find (average offers 3–30% worse *versus* our 11%); clearly there are many other differences between our market environments. List also offers evidence from complementary experiments that statistical discrimination is the main force explaining black–white gaps; given the large number of card dealers competing for business and their expertise (which could prepare them to effectively discriminate statistically), it is perhaps unsurprising that animus plays less of a role. The online classified advertising market we consider is perhaps more typical of interpersonal negotiation between non-experts where outside options (i.e. what a seller might do with an unsold iPod) are less clear. Although our experimental design does not provide similar opportunities to tease out mechanisms, we do find discrimination is heterogeneous across markets in patterns consistent with some expected channels through which statistical discrimination might operate.

A number of studies have avoided the issues (and costs) associated with hiring actors by signalling race through the use of racially distinctive names. Perhaps the best known application of this approach was conducted by Bertrand and Mullainathan (2004), who responded to job postings in Boston and Chicago using fictitious resumé, randomly assigned either a distinctively black or white name. The authors measured whether employers followed up with each application, and found that those with black names had a callback rate one third lower than those with white names. This difference was remarkably consistent across industries, and persisted for 'higher quality' (i.e. better educated and more experienced) applicants as well as for those randomly assigned mailing addresses in more affluent neighbourhoods. Similar studies have been conducted in several other countries considering, for example, Middle Eastern names in Sweden (Carlsson and Rooth, 2007) and Turkish names in Germany (Kass and Manger, 2012). A number of other authors have used racially distinctive names to experimentally investigate the impact of race in online markets including apartment rentals (Carpusor and Loges, 2006; Ahmed and Hammarstedt, 2008; Bosch *et al.*, 2010; Hanson and Hawley, 2011; Ewens *et al.*, forthcoming) and low-value auctions (Nunley *et al.*, 2011).

Two primary criticisms of the Bertrand and Mullainathan design have been raised. The first results from the use of names as a proxy for race, rather than a more direct signal. Market actors may have viewed stereotypically black names as signals of

socioeconomic status or family background and responded in a way that they might not have to a more typical black candidate.⁴ This concern applies for most ‘correspondence’ studies, as names are typically the most appropriate way of signalling race to a potential buyer, seller or employer in an experimental setting. Second, the measured outcome – callbacks, for instance – is often not the ultimate outcome of interest. While the number of callbacks is interesting, it does not tell us how many of those applicants might have been offered a job or apartment rental, or what wage or price they might have received.

Our experimental design attempts to address these concerns in two ways. First, by signalling race through the inclusion of a photograph, we can vary race while holding constant all other signals sent about attributes of the seller.⁵ Second, the fact that online transactions are brought near completion without face-to-face contact makes it possible to consider outcome measures that are relatively ‘close’ to the true outcomes of economic interest. One attractive feature of the classified advertising market we consider relative to online auctions is that buyers expect that completing their transaction will involve face-to-face interaction with the black or white seller; as this is typical of many non-online transactions, we expect that our results will be informative about discrimination offline.

Ayres *et al.* (2011) conduct a similar, complementary experiment, selling baseball cards on eBay. Although they also use photographs to signal a seller’s race in a goods market, that environment differs from ours in key ways: sellers never expect to meet buyers in person; and purchases are insured by eBay (so are very low risk to the buyer). Despite this very different environment, that study finds similar evidence of racial discrimination.

One trade-off that we make by using this market is an inability to complete as many transactions as we might have on a site like eBay. We believe this trade-off was worth it, as we are able to measure how apparent discrimination varies with market characteristics. We also believe the context of this experiment – selling a well-known, popular good on a highly trafficked website – makes our experiment more representative of day-to-day interactions than studies conducted in more specialised markets full of experts (List, 2004).

This study proceeds as follows: We begin with an overview of our experimental procedure, including discussion of local classified advertising markets, the contents and timing of our advertisements and the manner in which we negotiated with respondents (Section 1). Next, we compare the response to white, black and tattooed sellers’ advertisements along a number of dimensions (Section 2). Black (and tattooed) sellers receive fewer, less trusting responses and fewer, less valuable offers. Section 3 considers whether the outcome differences vary systematically across markets. Black sellers seem to fare particularly poorly in less competitive markets and in markets with the highest degrees of racial isolation or property crime. Section 4 concludes.

⁴ Bertrand and Mullainathan (2004) are forthright in recognising this concern, suggesting even in the study’s title that their conclusions are fundamentally about applicants with names that are much more common among either blacks or whites (such as Lakisha, Jamal, Emily and Greg).

⁵ Of course, our treatment does signal that the seller has chosen to show her hand – and thus reveal her race. Approximately 16% of the iPod Nano advertisements on the websites we consider with personal photos include the seller’s hand, suggesting that the practice is not uncommon (although not typical). To the extent that buyers interpret this as a signal of the seller’s confidence or naïveté, for instance, this could be a confounding factor, although we do not believe it has a large effect on our results. We would also note that it is impossible to hide one’s race in most market transactions.

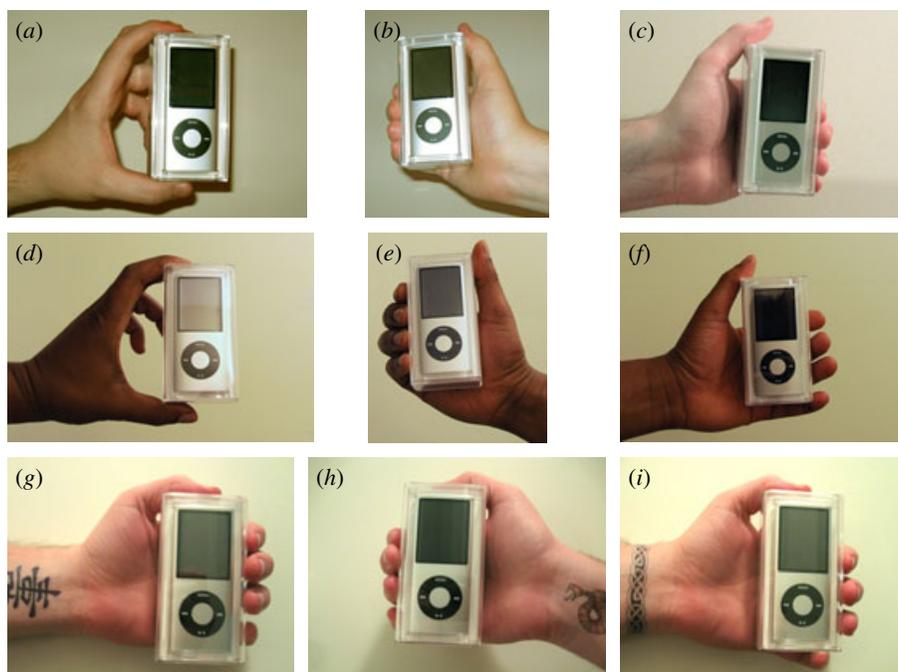


Fig. 1. *Advertisement Photographs*

Note. These photographs have been slightly scaled down from the size included in our advertisements.

1. Procedure

The goal of this study is to isolate the effect of race on market outcomes rigorously via a carefully constructed field experiment. We have developed a procedure that avoids several confounding factors present in other studies and which is replicable in a variety of settings. We posted online classified advertisements on locally focused websites throughout the US over the course of one year, with variation along three key dimensions: race or social group of the seller (as indicated by a photograph), asking price and the ‘quality’ of the advertisement text. This results in a $3 \times 3 \times 2$ design (three types of sellers, three prices and two qualities). The photographs used are shown in Figure 1. Online Appendix Table A2 tabulates these advertisement characteristics which – along with the markets in which the advertisements are posted and our posting and negotiation procedures – are discussed in greater detail below.

1.1. Overview: Online Classified Advertisements

We posted advertisements for new Apple iPod Nano 8 GB Silver portable digital music players⁶ on local classified advertising websites in approximately 300 geographical

⁶ Apple released an updated iPod Nano model in the midst of our experiment. Our advertisements offer the current model – the ‘fourth generation’ (model MB598LL/A) before 9 September 2009 and the ‘fifth generation’ (model MC027LL/A) after that date. The two models appear almost identical in their packaging.

markets. The sites together compose a network that is a major national source of online classified advertising. All sites are publicly accessible and fee-free for those looking to buy or sell items. We used all sites available in the network as of March 2009.

Potential buyers responded via advertisement-specific, anonymised e-mail addresses. We then followed typical practice in these markets, where a seller replies to individual buyers to negotiate a final price and – if they reach agreement – arrange to complete the transaction. An advertisement might receive zero responses or a dozen, depending on the market demand for a particular good, the contents of the advertisement and any number of idiosyncratic factors. The ensuing negotiations are, in general, *ad hoc*; that is, there is no formal bidding mechanism, and either party can cease communication at any time without facing any consequences.⁷

Among the experimental advantages of considering classified advertising in this setting are the local focus and the lack of information each potential buyer has about other buyers and their offers. Given the local focus of the sites on which we posted, buyers generally assume that sellers are local. In addition, the network of sites provides no facility for viewers to browse or search for advertisements across multiple markets, further encouraging local use. This is in contrast to online auctions like eBay, where it is normal to do national searches. The local focus allows us to analyse regional differences as well as variation by local economic and demographic characteristics; it also made it feasible to post multiple advertisements in a limited time frame while minimising the risk that our analysis of any given advertisement is contaminated by our other postings. Clearly, potential buyers' bids are affected not only by how much the iPod is worth to them but their assumptions about the seller and the other buyers who might be bidding.

1.2. Local Markets

Over the course of the experiment we posted at least three advertisements in each market available in the network,⁸ which collectively covers the geography of all 50 states and Washington, D.C. There are over 300 local sites, which include a wide variety of locations – from small towns in rural areas to the centres and suburbs of large cities. More information on the specific markets is included in online Appendix A. Within a single market, sellers choose a category in which to list their advertisement; we posted in the 'electronics for sale' category (as do the vast majority of other advertisers offering iPods for sale).

⁷ Indeed, our experience confirms anecdotal evidence that potential buyers regularly cease communicating in the midst of discussing a potential sale.

⁸ In fact, we attempt at least three postings, not all of which were successful. We randomly selected four markets each morning and evening, without replacement, until we exhausted the full list of 329 markets. Some markets are further divided geographically on their sites; for these we treated submarkets separately but would not simultaneously select multiple submarkets within the same market. Markets (and submarkets) were replaced after several weeks in cases where an advertisement was 'ghosted' – that is, algorithmic filters on the website prevented our advertisement from ever appearing on the site. Once we exhausted the full list of markets, we began again using the full pool of markets. We thus attempted to post at least three advertisements per market – more in markets with several submarkets – and those were, by design, spaced out over the course of the year.

Table 1
Market Characteristics—Summary Statistics

	Mean	Standard deviation	25%	50%	75%
Other ads in market (prior week)	15.7	33.2	1	3	11
20+ ads in market (prior week)	0.18				
Northeast	0.13				
Midwest	0.24				
South	0.36				
West	0.27				
% population white	77.0	16.1	67.1	81.5	90.1
% population black	12.8	14.6	2.4	7.2	16.9
% population Hispanic	13.5	16.6	3.2	6.9	16.7
% population Asian	3.3	4.1	1.3	2.0	3.6
Poverty rate	15.7	6.3	11.7	14.7	19.1
Median household income (\$K)	46.3	10.9	39.4	44.5	51.1
Property crime rate	357.6	125.0	275.7	337.9	411.5
Black isolation index	0.19	0.17	0.02	0.13	0.32
Observations	1,200				

Notes. All observations equally weighted. Local racial composition, poverty rate and household income are from the 2007 American Community Survey. 2008 property crimes are per ten thousand people (from United States Department of Justice and Federal Bureau of Investigations, 2009). Black isolation index is degree to which 'the average black resident lives in a census tract in which the black share of the population exceeds the overall metropolitan average' in 2000 (from Glaeser and Vigdor, 2001, or constructed using their methodology). Crime and isolation data are not available for all markets.

At the time of our listing, the average market had 15.7 other advertisements for iPod Nanos that had been listed in the previous week, and 18% of our advertisements were posted in markets with at least 20 other advertisements.⁹ Markets with more advertisements posted presumably get more traffic from potential buyers. Thus, advertisements in thicker markets may get more responses on average. On the other hand, in markets with more sellers our advertisements face greater competition for prospective buyers' attention and dollars.

Table 1 shows summary statistics for several market characteristics. In addition, online Appendix Table A1 shows average values for these characteristics broken down by advertisement type.

1.3. *Advertisement Contents*

The contents of our advertisements varied along three dimensions: photograph (including skin colour), advertisement text and asking price. Photographs were randomly assigned to each advertisement, but skin colours were not replaced within

⁹ This count is based on a search for other advertisements in the same market that include the phrase 'iPod Nano' (regardless of capitalisation) in their title. (This count therefore includes both new and used items, and some non-iPod items, such as accessories.) Note that we have data on the stock of advertisements listed on the site when we post, but not on any flow of advertisements posted. Sellers can remove their listings, so the number of listings of vintage less than one week gives only a lower bound on the number of sellers to which a potential buyer may have been exposed during that week.

markets – that is, if a market’s first advertisement included a white hand, the second advertisement posted there only randomised over the photos with black or tattooed hands. Advertisement texts and asking prices were randomly assigned with replacement.

1.3.1 *Photograph*

Each advertisement included a photograph of a new, unopened iPod held in a black hand, a white hand or a white hand with a wrist tattoo. (The tattoo treatment was introduced somewhat after the start of the experiment.) All the photographs were of a man’s hand so, strictly speaking, our experimental design will only allow us to assess the discrimination faced by black men. We use multiple photographs of each type to limit the chance that a buyer might see the same photo twice, and therefore to help make the advertisements independent observations. Three styles of photographs were used for the black and white hands; the need to display the tattoo prevented us from reproducing all three of those hand positions perfectly in the last series of photos. Nevertheless, the pictures (reproduced in Figure 1) are very similar in all ways other than the apparent race or social group of the seller.

Photographs are very common in online classified advertisements, and are included in approximately 60% of the other iPod Nano advertisements we observe. Typically, these are either stock/marketing images or personal photographs of the item for sale; our photos are similar in style to the personal photos many others use.

1.3.2 *Advertisement text*

Our advertisements (and the ensuing e-mail correspondence discussed in subsection 1.5) randomised over six different texts: three types, each with a ‘high-quality’ and a ‘low-quality’ variant. We used multiple text types to create within-market variety that minimises the apparent suspiciousness associated with repeatedly posting advertisements in the same market. (We were concerned here both with the websites’ users and with spam filters present on the sites themselves.) All six texts are included in online Appendix B.

The three high-quality texts use proper capitalisation, punctuation and grammar, and were generally well-written. Our low-quality advertisements had the same content, but with less sophisticated wording and incorrect spelling, grammar and capitalisation. Our aim was to provide a signal of the seller’s socioeconomic status, proxied by his education level and writing ability.

1.3.3 *Asking price*

Each advertisement also included an asking price (both in a searchable price field and in the text of the listing) of either \$90, \$110 or \$130.¹⁰ The iPod we advertised were popular and widely available through electronics retailers, mass market stores, online vendors and Apple Stores. It had a list price of \$149.99 (plus local sales tax) and was available for sale prices of approximately \$135 throughout our experimental

¹⁰ We limited asking prices to \$90 and \$110 beginning in December, 2009, due to \$130 advertisements’ very low response rates.

period, so all three asking prices were below the amount buyers would have paid in a store. This asking price represented the ‘first offer’ in the sale negotiation, and we expect to see buyers’ responses depending on it. In addition to producing an anchoring effect (as in Tversky and Kahneman, 1974), the specific asking price also sent prospective buyers a signal about market conditions, the seller and the quality of the product.

1.4. *Timing*

Our experimental period covered 16 March 2009 to 15 March 2010, excluding the periods around major holidays (and various other times at which we suffered technical difficulties). Advertisements were posted in the morning and evening (at approximately 9:30 A.M. and 9:30 P.M., Pacific Time), with no more than four online at any given time. A tabulation of advertisement timing by advertisement type is provided in online Appendix Table A3.

We removed our advertisements approximately 12 hours after they were posted; no new potential buyers would view or respond to an advertisement after that point, although ongoing e-mail exchanges could and did continue well beyond the 12-hour mark. During a pilot of the experiment in which we posted advertisements for longer durations, we found that the vast majority of responses were received within 12 hours, and it was common practice to complete transactions within a day or two after posting. Thus, our 12-hour window gives us sufficient time to receive responses from most likely buyers.

We added the non-race social signal dimension of this experiment after we had already begun posting advertisements with black and white photos. Thus, a larger share of the later advertisements include tattoos. The results reported below are robust to the inclusion of a quadratic time trend to control for this correlation between advertisement type and timing, as well as several alternate strategies for controlling for advertisement timing such as including in regression specifications the order in which advertisements were posted within each market.

The weeks around two particular gift-giving holidays, Christmas and Valentine’s Day, saw a large increase in responses to our advertisements and the offers received. Our analyses therefore include controls for these two periods.¹¹

1.5. *Negotiation with Respondents*

Beginning approximately two hours after each advertisement is posted, we sent a response via e-mail to each respondent saying that we had received many e-mails and asking for her best offer (or to confirm that an offer made in an initial e-mail was indeed her best). The text of all interactions was scripted and is included in online Appendix C, together with additional details about our negotiation procedure.

¹¹ We define the Christmas period as the Monday after Thanksgiving (30 November) to 21 December. We did not post advertisements from 22 December to 5 January. The Valentine’s Day period runs from two weeks before the holiday to one week after (31 January–21 February). We include the days after the holiday because some buyers reported looking for gifts to reciprocate gifts they had unexpectedly received.

In the course of our correspondence with potential buyers, we received a large number of ‘scam’ offers (both as initial responses to our advertisements, and following our first e-mails). These scams generally comprised offers to pay high prices to have the item shipped overseas; several samples are included in online Appendix D.¹² We coded all requests for shipping or non-cash payments (and other similar responses) as scams and ceased correspondence with these respondents.¹³

Approximately 48 hours after removing each advertisement, we offered to sell the iPod, by postal mail, to the respondents who made and confirmed the highest offer and (when available) the second-highest offer. We apologised for being out of town, and told the respondent we were willing to mail her the iPod in exchange for payment via PayPal, an electronic payment system widely used for online person-to-person transactions. The time delay was intended to make our shipment proposal less suspicious; buyers might think we were local but had to leave town after posting the advertisement. We sent the iPods to those who agree to this, and replied to all other confirmed bidders that the iPod was no longer available.

The reasons we chose to offer shipment rather than in-person delivery were principally logistical, but we also sought to avoid introducing unobservable (and uncontrollable) variation. Given the local nature of our advertisements and the sites we posted them on, most high bidders are understandably wary of a long-distance transactions; those who agreed to trade this way are unlikely to compose a representative sample of potential buyers. Nonetheless, we completed as many transactions as possible in the spirit of honestly following through on our advertised offer to sell.

2. Results: Average Effects

We consider six types of outcome measures: whether our advertisements were prematurely removed by website users (described below); the number of responses received; qualitative characteristics of the responses’ contents; the dollar amounts offered; high bidders’ reactions to our stated inability to deliver the iPod in person; and the probability that an advertisement resulted in a successful sale. Average values for these measures by advertisement type are reported in Table 2.

The following subsections present findings on the effects of race on these outcomes, controlling for a variety of advertisement, timing and market characteristics. Given that seller race was experimentally varied independent of these other characteristics, the race effects are consistently estimated by the difference or ratio of means reported in

¹² The associated fraud appears to operate in at least two ways. First, the ‘buyer’s’ payment – whether by online payment service, cheque or money order – is counterfeit, allowing her to acquire the item at no cost. The second technique is more insidious. The seller receives an e-mail purporting to be from her bank or an online payment service, confirming that a payment has been received. The web links in this e-mail lead to sites controlled by the scammer, who hopes that the seller will enter her bank account or online payments password.

¹³ After several months of reading and responding to potential buyers’ e-mails, it became increasingly obvious which e-mails were attempted scams. As not all of these e-mails result in follow-ups that would confirm our suspicions, we coded such responses as ‘probable scams’ to distinguish them from genuine offers. We code responses as probable scams if the text of the e-mail or e-mail address is identical to those from a confirmed scam e-mail we received earlier. Our results are robust to this alternative coding procedure.

Table 2
Key Outcome Averages by Advertisement Type

	White	Black	Tattoo	Total
Prematurely removed	0.028	0.056	0.041	0.041
<i>Number of responses</i>				
Number of non-scams	2.46	2.06	2.07	2.21
Number of offers	1.70	1.36	1.44	1.50
Received ≥ 1 offer	0.624	0.559	0.586	0.590
<i>Indicators of trust in responses (given ≥ 1 non-scam response)</i>				
Includes name	0.391	0.301	0.315	0.339
Uses polite language	0.415	0.370	0.354	0.383
Includes personal story	0.038	0.046	0.048	0.044
<i>Offer amount</i>				
Mean offer	53.51	46.84	48.93	49.86
Best offer	58.51	50.36	52.93	54.05
<i>Offer amount (given ≥ 1 offer)</i>				
Mean offer	85.76	83.78	83.45	84.46
Best offer	93.79	90.07	90.25	91.56
<i>Reaction to delivery proposal (given delivery proposed)</i>				
Scam/payment concern	0.075	0.107	0.084	0.088
No response	0.376	0.424	0.398	0.398
Other	0.191	0.139	0.199	0.176
Prefer to wait	0.303	0.297	0.260	0.289
Willing to ship	0.056	0.033	0.059	0.049
iPod shipped	0.037	0.017	0.031	0.028

Notes. Mean values are reported. Observations are weighted by state population/number of advertisements posted in each state.

Table 2; including additional controls merely allows us to increase the estimates' precision.

Online Appendix E reports these results for specifications without controls, excluding prematurely removed advertisements, using linear models, replacing the market controls with market fixed effects and with alternate weighting.

2.1. *Premature Advertisement Removal*

The sites on which we posted provide tools for users to mark advertisements as inappropriate or unwelcome. If enough users protest a particular advertisement in this way, it is removed from the website: 4.1% of our advertisements were removed in this manner. In addition to legitimate use of this feature, other sellers may disingenuously mark competing advertisements as inappropriate to reduce competition.¹⁴

¹⁴ In addition, the websites implement filters (based on unknown algorithms that appear to change frequently) to identify unwelcome advertisements. On several occasions, all of the advertisements we posted on a given morning or evening were immediately removed from the site. This universal, simultaneous premature removal suggested that our advertisements were caught in the websites' filters. Similarly, some advertisements did not show up in search results despite appearing to have posted successfully; this is also due to the websites' screening for unwelcome advertisements. All of these advertisements are entirely excluded from our analyses. Stratifying the advertisements we posted simultaneously by advertisement quality and market size greatly reduced this automatic removal.

Table 3
Key Outcome Regressions

	(1) Prem. rem. Probit	(2) Nonscams Neg. Bin.	(3) Offers Neg. Bin.	(4) Any offer Probit	(5) Mean offer OLS	(6) Best offer OLS	(7) Shipped Probit
Black	0.0270* (0.0149)	0.869** (0.0532)	0.8292** (0.0643)	-0.0646* (0.0370)	-5.720** (2.760)	-7.069** (2.978)	-0.0127 (0.00776)
Tattoo	0.0136 (0.0141)	0.826*** (0.0523)	0.836*** (0.0589)	-0.0560 (0.0387)	2.533* (2.939)	-6.597** (3.072)	-0.00381 (0.00731)
High quality	-0.0158 (0.0103)	0.988 (0.0535)	1.022 (0.0614)	0.0508 (0.0324)	4.384* (2.521)	4.750* (2.663)	-0.00560 (0.00672)
Price \$110	0.00607 (0.0116)	0.4292*** (0.0266)	0.388*** (0.0281)	-0.287*** (0.0369)	-11.55*** (2.849)	-16.38*** (3.012)	-0.0119* (0.00613)
Price \$130	-0.0242*** (0.0109)	0.224*** (0.0180)	0.183*** (0.0190)	-0.494*** (0.0347)	-27.84*** (3.057)	-34.35*** (3.237)	-0.0289*** (0.00717)
Christmas	0.0209 (0.0307)	1.873*** (0.181)	2.027*** (0.239)	0.120 (0.0752)	12.27** (5.681)	17.47*** (6.520)	0.0224 (0.0246)
Valentine's Day	-0.00724 (0.0195)	1.312*** (0.110)	1.361*** (0.120)	0.195*** (0.0520)	14.28*** (4.746)	15.00*** (4.803)	0.00448 (0.0136)
Night	-0.00836 (0.00972)	0.698*** (0.0432)	0.663*** (0.0426)	-0.133*** (0.0317)	-10.68*** (2.549)	-12.41*** (2.670)	0.00152 (0.00642)
20+ weekly advertisements	-0.00482 (0.0187)	2.002*** (0.179)	2.103*** (0.203)	0.210*** (0.0420)	16.08*** (3.548)	20.69*** (3.999)	-0.00754 (0.00837)
Median household income (log)	-0.0615 (0.0447)	2.665*** (0.833)	3.249*** (1.121)	0.357** (0.178)	34.43*** (13.02)	34.68*** (14.26)	0.0698** (0.0353)
Poverty rate	-0.00198 (0.00155)	1.026** (0.0122)	1.038*** (0.0137)	0.0102* (0.00615)	0.924** (0.455)	0.948* (0.498)	0.00238* (0.00133)
% population white	0.000108 (0.000455)	0.995* (0.00244)	0.997 (0.00261)	-0.00206 (0.00139)	-0.143 (0.0967)	-0.179* (0.106)	-0.000154 (0.000282)
Northeast	-0.0111 (0.0144)	0.952 (0.0926)	1.061 (0.109)	0.0330 (0.0565)	0.154 (4.211)	1.153 (4.532)	-0.00669 (0.0119)
Midwest	-0.0107 (0.0136)	0.981 (0.0829)	0.977 (0.0929)	-0.0152 (0.0511)	-1.894 (3.977)	-1.361 (4.240)	0.00295 (0.0106)
South	-0.0154 (0.0152)	0.887 (0.0828)	0.937 (0.0996)	-0.0280 (0.0498)	-2.086 (3.764)	-3.086 (4.049)	0.00527 (0.00988)
Observations	1,200	1,200	1,200	1,200	1,200	1,200	1,200
White mean	0.0278	2.459	1.691	0.624	53.51	58.51	0.0367

Notes: Marginal effects from probit estimation are reported in (1), (4) and (7); (2)–(3) report incidence rate ratios from negative binomial estimation; (5)–(6) report OLS coefficients. Standard errors clustered by market are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Observations are weighted by state population/number of advertisements posted in each state.

Column 1 of Table 3 provides estimated marginal effects associated with probit estimation assessing which advertisements are most likely to be prematurely removed. The regression controls for a variety of advertisement, market and timing characteristics that explain a substantial amount of the variation in our dependent measures.

On average, black sellers' advertisements are 2.7 percentage points more likely to be prematurely removed than white sellers' advertisements; the likelihood is thus almost twice as high that a black seller's advertisement will be removed.

Clearly if an advertisement is removed, it limits the seller's opportunity to receive responses and bids from potential buyers. Although this is an economically relevant outcome for a seller, he can also attempt to repost his advertisement. We include prematurely removed advertisements in the analyses that follow, but report results excluding these advertisements in online Appendix Table E4.

2.2. Number of Responses

80% of our advertisements received some response, and on average they received 2.7 responses. We identify a number of our responses as disingenuous 'scams', and partition the remainder based on whether or not they result in a specific dollar offer. Table 4 provides summary statistics on the number of responses received broken down by response type.

Given that the number of responses (of each type) received are count variables, we estimate the impact of race and other covariates using models of the form

$$\text{responses}_i \sim \text{Poisson}[v_i \exp(\mathbf{x}_i \boldsymbol{\beta})] \quad (1)$$

$$v_i \sim \text{Gamma}(1/\alpha, \alpha), \quad (2)$$

where i indexes advertisements and \mathbf{x}_i is the i th row of the data matrix \mathbf{X} , containing the covariates for advertisement i . This yields a negative binomial distribution for the outcome of interest (conditional on covariates).¹⁵ Note that this negative binomial distribution has $\mathbb{E}(\text{responses}_i) = \exp(\mathbf{x}_i \boldsymbol{\beta})$; thus, the reported exponentiated coefficient estimates (corresponding to $\exp(\beta_j)$ in (1)) should be interpreted as incidence

Table 4
Number of Responses—Summary Statistics

	Mean	Standard deviation	25%	50%	75%	95%	Max.	Frac. > 0
Responses	2.65	2.76	1	2	4	8	17	0.80
Scams	0.44	0.78	0	0	1	2	10	0.32
Non-scams	2.21	2.73	0	1	3	8	17	0.70
Offers	1.50	2.05	0	1	2	6	15	0.59
Observations	1,200							

Note. Observations are weighted by state population/number of advertisements posted in each state.

¹⁵ In these negative binomial models, $\alpha \leq 0$ parameterisation over/under-dispersion relative to the Poisson distribution, since $\text{responses}_i \rightarrow \text{Poisson}[\exp(\mathbf{x}_i \boldsymbol{\beta})]$ as $\alpha \rightarrow 0$.

rate ratios. A covariate has a positive effect on the outcome measure precisely when its corresponding exponentiated coefficient is greater than one; to determine the combined effect of several covariates, multiply the exponentiated coefficients together.

Responding to an advertisement requires no commitment and limited time, so it is cheap but it is not free. There is no incentive for anyone to respond to an advertisement in which he is completely uninterested. Also, the number of responses received is unaffected by our subsequent e-mail correspondence, which may send additional signals about the seller and the local market. In particular, our first scripted e-mail response suggests that there is a lot of interest in our iPod (i.e. that the market is competitive) and that the seller is fairly savvy and organised in his approach to selling the item. Thus, the number of responses may best reflect local buyers' prior assumptions about black and white sellers, as well as the demand to purchase from them. To the extent that our correspondence provided additional information that contradicts these assumptions, some buyers might have ceased communication because they were no longer interested in purchasing from us, not because they were not serious to begin with.

Column 2 of Table 3 reports the results of a maximum likelihood estimation of (1–2) for the number of non-scam responses received. While our average advertisement received 2.2 non-scam responses, black sellers received 13% fewer responses than white sellers. Tattooed sellers appear to suffer even more discrimination than blacks along this margin, receiving 17% fewer responses than white sellers.

Several other covariates seem to have the expected effects: high asking prices severely depress response and advertisements posted at night or in markets with few other advertisements fare poorly. Perhaps surprisingly, advertisement quality appears to have no effect on the number of responses received. Based on the degree to which many responses to our advertisements were poorly written, it is possible that our high and low-quality advertisements were simply insufficiently differentiated.

The number of dollar-valued offers¹⁶ received may be a more reliable measure of serious interest, especially if we think that some buyers were searching for a good deal by indiscriminately responding to many sellers' advertisements. We record the dollar amount of an offer whether it comes in the initial inquiry or in response to our reply. Approximately two thirds of non-scam responses resulted in an offer and the average advertisement received 1.5 offers.

We report negative binomial regression results for the number of offers in column 3 of Table 3. Black sellers receive 18% fewer offers than white sellers, whereas tattooed sellers receive 16% fewer.

2.3. *Response Characteristics*

The manner in which buyers respond to advertisements may indicate their underlying level of respect or trust. We analyse the text of the first e-mail each buyer sends, identifying whether:

¹⁶ Throughout the study, we refer as 'offers' only to cash offers. Approximately 4% of non-scam respondents offered to trade various goods and services – from live snakes to auto detailing – for our iPod. Several examples are included in online Appendix D.

- (i) The buyer included or signed their name (34% of responses);
- (ii) The buyer was polite, including the words ‘please’, ‘thank you’, or variations such as ‘pls’, ‘thx’, or ‘thanks’ anywhere in the e-mail text (38%); and/or
- (iii) The buyer included a personal story, presumably to appeal to the seller’s sentiments and get a lower price (4%).

Note these characteristics are neither mutually exclusive nor collectively exhaustive; examples of responses exhibiting each are included in online Appendix D.

Table 5 reports probit regression results for these three attributes of buyers’ responses. Obviously, this analysis is restricted to advertisements which received at least one non-scam response, which may introduce some selection effect. Overall, buyers are more likely to act respectfully when communicating with white sellers. Approximately 7% fewer buyers sign their names when responding to black rather than white sellers; thus, the average response received by a black seller is 19% less likely to include the buyer’s name. This is similar to the effect observed for tattooed sellers.

Buyers are slightly less likely to use polite language when responding to black or tattooed sellers’ advertisements, although these results do not rise to the level of statistical significance.

2.4. Offer Amount

The ultimate reason that the number of responses is economically important to a seller is that it increases the probability of receiving a good offer and of completing a sale. We thus look at both the mean and maximal offers made in response to each advertisement. To the extent that a seller is able to successfully complete a sale with the highest bidder at that bidder’s offered price, the ‘best offer’ received is the outcome of primary economic importance to the seller. In our main specifications, we treat the failure to receive any offer as a \$0 offer.

Table 5
Probit Regression of Response Characteristics

	(1) Name	(2) Polite	(3) Personal
Black	-0.0746*** (0.0286)	-0.0252 (0.0295)	0.0113 (0.0121)
Tattoo	-0.0782** (0.0307)	-0.0398 (0.0319)	0.00694 (0.0138)
Standard controls	✓	✓	✓
Observations	2,547	2,547	2,547
White mean	0.391	0.415	0.038

Notes. Observations are weighted by reciprocal of number of responses per advertisement. ‘Standard controls’ are: high advertisement quality, asking price (\$130 and \$110 dummies; \$90 excluded), holidays (Christmas and Valentine’s day dummies), night, 20+ iPod Nano advertisements in market over previous week, median household income, poverty rate, non-Hispanic white fraction of local population and region (Northeast, Midwest, and South dummies; West excluded). Probit marginal effects are reported. Standard errors clustered by advertisement are reported in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

Note that there are some challenges in interpreting the effects of advertisement type (or any other covariate) on the mean offer. The mean offer observed is an average across the subset of potential respondents who chose to make an offer. We have already observed that black sellers receive fewer offers; to the degree that the buyers who adjust on the extensive margin would have been particularly low or high bidders, selection could drive the mean offer up or down.

On average, advertisements received a mean offer of \$49.86 and a maximal offer of \$54.05. We present ordinary least squares results assessing the effect of advertisement type and other covariates on these outcomes in columns 5 and 6 of Table 3. Compared with white sellers, black sellers receive average offers \$5.72 (11%) lower and tattooed sellers \$5.53 (10%) lower. Black and tattooed sellers' best offers are also lower than whites', by \$7.07 (12%) and \$6.60 (11%), respectively.

Online Appendix Table E6 shows the effect of race on mean and best offers conditional on receiving at least one offer. The offers received from potential buyers are lower for black and tattooed sellers (although not all differences are statistically significant), despite the selected – and presumably less biased – pool of respondents.

2.5. *Reactions to Delivery Proposal*

After we took an advertisement down, we contacted the highest bidder to say that we would mail the iPod to her if she would pay us using PayPal. Because the websites include warnings about the risks of non-local transactions, we did not expect many buyers to accept this offer. However, the manner in which they declined can tell us something about their inclination to trust the seller. Buyers' initial responses to our delivery proposal fall into one of five mutually exclusive categories, listed here in order of most to least positive:

- (5) suggesting an openness to receiving the iPod by mail (5% of proposed deliveries);
- (4) offering to wait and meet when we get back into town (29%);
- (3) declining for some other reason (18%);
- (2) no response (40%), which we interpret as a signal of some distrust; or
- (1) explicitly accusing us of trying to scam them, or saying they do not want to use PayPal, which we interpret as a concern about being scammed (9%).

Examples of each type of reaction are included in online Appendix D.

In Table 6, we report the results of ordered probit regressions of buyers' reactions to our delivery proposal on advertisement type. These regression specifications allow us to test whether each seller type received 'more positive' reactions as measured by the ordinal ranking above. The statistically significant negative coefficients on blacks support the hypothesis that black sellers receive worse reactions to their delivery offers, suggesting an underlying distrust of black sellers. We are no longer relying on truly experimental variation but note that black and tattooed sellers suffer even though the sample consists only of the (presumably less biased) potential buyers who not only chose to respond to those sellers' advertisements but made the highest offers.

Table 6
Ordered Probit Regression of 'Positivity' of Reaction to Delivery Proposal

	(1) Reaction to delivery proposal	(2) Reaction to delivery proposal
Black	-0.168* (0.101)	-0.172* (0.102)
Tattoo	-0.135 (0.108)	-0.136 (0.108)
Standard controls		✓
Observations	622	622

Notes. Ordered probit coefficients are reported. Standard errors clustered by advertisement are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Observations are weighted by reciprocal of number of delivery proposals per advertisement. Outcomes ranked from most to least positive are: willing to ship, prefer to wait, other, no response and scam/payment concern. 'Standard controls' are: high advertisement quality, asking price (\$130 and \$110 dummies; \$90 excluded), holidays (Christmas and Valentine's day dummies), night, 20+ iPod Nano advertisements in market over previous week, median household income, poverty rate, non-Hispanic white fraction of local population and region (Northeast, Midwest and South dummies; West excluded).

Online Appendix Table E7 reports probit estimates of the frequency of receiving each individual reaction type. Although no results rise to conventional levels of statistical significance, black sellers are somewhat less likely to face the three most positive reactions and more likely to face the two most negative ones.

2.6. Shipment

After offering to ship the iPod to the highest bidder, our procedure becomes more *ad hoc* out of necessity (we must respond to questions, and work out the logistics of shipment and payment) but remains blind to the seller's type. Column 7 of Table 3 reports the effect of seller type on the probability that advertisement results in a successful transaction. The number of successes is small (as delivery by mail is not typical in this market), so the estimates are imprecise and statistically insignificant. On average, advertisements posted by black sellers ultimately result in sales almost 35% less often than advertisements posted by white sellers.

3. Understanding Observed Discrimination

In the previous Section, we analysed the differences in a number of outcomes faced by white, black and tattooed sellers. We now investigate whether these differences vary systematically across markets, focusing on three key outcome measures: the number of offers received, the mean offer and the best offer.

We examine several possible drivers of systematic variation that may speak to four broad hypotheses about discrimination in this market: competition limits discrimination; racial disparities are driven by statistical discrimination; racial disparities are driven by animus; and racial disparities are the result of attracting different pools of buyers.

Competition among buyers – who have different preferences for discrimination – should improve outcomes for black sellers. Meanwhile, the presence of larger

black–white outcome gaps in some settings than others could suggest the presence of statistical discrimination as distinct from animus. The former generally refers to discrimination where race is used as a proxy for other characteristics that buyers cannot observe directly but wish to avoid (e.g. low socioeconomic status). Animus, or taste-based discrimination, is a negative reaction to race itself, independent of other characteristics.

We expect that buyers might statistically discriminate in this market to avoid one or more of the following: buying fake or stolen goods;¹⁷ sellers they would need to meet in an inconvenient/dangerous neighbourhood and unreliable sellers who would not complete the transaction.¹⁸

It will of course not be possible for us to disentangle these types of discrimination completely. In practice, animus can be a by-product of statistical discrimination, and *vice versa*; the presence of one type does not preclude the presence of the other. Readers should interpret our results as merely suggestive of the underlying mechanisms at work in this market.

Except for advertisement text, which we varied experimentally, all of the relevant market characteristics (degree of competition, racial isolation, property crime rates and racial composition) are at least weakly correlated. We therefore test for all of the heterogeneous effects in a single regression. This reduces our statistical power but lessens concerns about omitted variable bias.

3.1. Market Competition Reduces Discrimination

In theory (Becker, 1971), discrimination against black sellers – and perhaps especially taste-based discrimination – should be less present in markets with more competition among buyers. Buyers for whom it is more costly to interact with a black seller will be outbid by buyers who do not discriminate between black and white sellers. This should improve outcomes for black sellers.¹⁹

We do not observe the number of potential buyers in each market but do know the number of offers our advertisements receive. We use an indicator of the average number of offers received by all our advertisements posted in each market as a proxy for the degree of competition among buyers. Specifically, we create an indicator variable for each market equal to one if our advertisements receive two or fewer offers, on average, and zero otherwise. (Our advertisements average two or fewer offers in 80% of markets.) We then test the hypothesis that the impact of sellers' race varies significantly with the degree of competition.

Table 7 suggests that black sellers indeed face more discrimination in less competitive markets. (That is, the coefficient $\hat{\beta}_{\text{Mkt avg} \leq 2 \times \text{Black}}$ is less than one in the negative binomial

¹⁷ Indeed, the iPod Nano we sell is probably more likely to be fake or stolen than many other goods, so it is a particularly good test for this form of statistical discrimination.

¹⁸ The best test for statistical discrimination is whether gradually revealing additional information to the buyer – a receipt proving that the iPod was not stolen, a mailing address in a good neighbourhood etc. – changes the buyer's response. Animus will not be affected by new information. We were unable to do this here, but it would be a useful extension of this and similar studies.

¹⁹ Two recent studies related to ours assess the impact of competition on racial discrimination, with findings broadly consistent with ours: List and Livingstone (2010) and Nunley *et al.* (2011).

Table 7
Heterogeneous Effects by Market Characteristics and Advertisement Quality

	(1) Number of offers	(2) Mean offer	(3) Best offer
Black	1.467 (0.568)	13.33 (16.93)	12.49 (17.85)
Tattoo	1.298 (0.501)	15.10 (18.44)	11.84 (19.49)
Mkt avg ≤ 2 offers	0.489*** (0.0501)	-10.07** (5.053)	-14.26*** (5.447)
Mkt avg $\leq 2 \times$ black	0.771* (0.114)	-12.45** (6.169)	-12.96* (6.695)
Mkt avg $\leq 2 \times$ tattoo	0.865 (0.134)	-13.54** (6.489)	-14.48** (6.780)
High property crime rate	0.966 (0.108)	0.0458 (4.965)	-0.753 (5.359)
High crime \times black	0.903 (0.165)	-15.73** (6.643)	-16.30** (7.196)
High crime \times tattoo	0.957 (0.158)	-3.113 (7.082)	-2.557 (7.403)
High isolation	1.414*** (0.158)	9.878* (5.922)	12.40* (6.366)
High isolation \times black	0.578*** (0.114)	-21.45*** (8.064)	-26.71*** (8.808)
High isolation \times tattoo	0.730* (0.119)	0.0637 (8.261)	-4.619 (8.741)
High quality advertisement	0.981 (0.0846)	4.468 (4.266)	4.904 (4.526)
HQ \times black	1.091 (0.148)	-4.138 (5.908)	-4.457 (6.182)
HQ \times tattoo	1.042 (0.159)	0.808 (6.484)	1.441 (6.801)
Racist Google search index	1.005 (0.00467)	0.185 (0.254)	0.172 (0.269)
Racist Google \times black	0.993 (0.00686)	-0.0933 (0.300)	-0.0846 (0.316)
Racist Google \times tattoo	0.993 (0.00704)	-0.181 (0.343)	-0.136 (0.364)
% population black	0.988** (0.00452)	-0.335* (0.189)	-0.390* (0.208)
% Black \times black	1.009* (0.00562)	0.558*** (0.207)	0.610** (0.236)
% Black \times tattoo	1.015*** (0.00535)	0.0798 (0.261)	0.179 (0.275)
Regions \times black/tattoo	✓	✓	✓
Standard controls	✓	✓	✓
Observations	1,042	1,042	1,042
All black = 0 p.	0.0375	0.000731	0.000483
All black = tattoo p.	0.880	0.0881	0.0691

Notes. Observations are weighted by state population/number of advertisements posted in each state. 'High isolation' markets are top 25% as measured by degree to which 'the average black resident lives in a census tract in which the black share of the population exceeds the overall metropolitan average' in 2000 (from Glaeser and Vigdor (2001) or constructed using their methodology) 'High crime' markets are top 25% as measured by 2008 property crimes per capita (from United States Department of Justice and Federal Bureau of Investigations, 2009). 'Racist Google search index' is relative frequency of Google searches containing 'nigger' at state level from (Stephens–Davidowitz, 2013). 'Standard controls' are: high advertisement quality, asking price (\$130 and \$110 dummies; \$90 excluded), holidays (Christmas and Valentine's day dummies), night, 20+ iPod Nano advertisements in market over previous week, median household income, poverty rate, non-Hispanic white fraction of local population and region (Northeast, Midwest and South dummies; West excluded). Incidence rate ratios from negative binomial estimation are reported in (1); OLS coefficients are reported in (2)–(3). Standard errors clustered by market are reported in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

and less than zero in the ordinary least squares regressions.) Relative to white sellers, black sellers receive 23% fewer offers in markets with a low degree of competition among buyers. Similarly, in less competitive markets, black sellers' mean and best offers are \$12–13 further behind white sellers'. Results are similar for tattooed sellers.²⁰

3.2. *Heterogeneous Effects Consistent with Statistical Discrimination*

3.2.1 *Property crime rate*

Buyers might discriminate against black sellers statistically if they think those sellers are more likely to sell stolen goods or that it is more dangerous to meet those sellers in person (because the sellers living in high-crime markets are criminals themselves). Using data from the Uniform Crime Reports (United States Department of Justice and Federal Bureau of Investigations, 2009), which map 88% of our markets, we designate markets with 2008 property crime rates in the top quartile of our sample as 'high crime' areas. We then test the hypothesis that buyers are more likely to discriminate against black sellers in areas with high property crime rates than they are in areas with less crime, as might be expected if transacting with a black seller is perceived as posing a disproportionate risk of exposure to criminality.

Indeed, we do find that black sellers face worse outcomes in highest crime markets: they receive mean and best offers roughly \$16 lower (relative to white sellers) than in lower crime areas. (A statistically insignificant coefficient estimate suggests they also receive 10% fewer offers.) The effect is directionally similar, albeit much smaller and statistically insignificant, for tattooed sellers.

3.2.2 *Racial isolation*

Buyers might also discriminate against black sellers statistically if they assume it would be inconvenient to travel to meet those sellers. (If the seller is the one travelling, buyers might assume he is less reliable because of the inconvenience.) This is more likely when local black and white populations are more geographically isolated from one another. Glaeser and Vigdor (2001) created an 'isolation' index to measure segregation in metropolitan areas across the country. Their data map approximately 80% of our markets, and we use census data to construct the measure for additional markets. The index increases from zero to one with greater isolation and indicates the degree to which 'the average black resident lives in a census tract in which the black share of the population exceeds the overall metropolitan average'. That is, it measures how geographically segregated the local black population is from the local white population.

We denote markets in the top quartile of racial isolation scores as exhibiting 'high isolation' and consider the differential effect of race in those markets. If statistical

²⁰ An alternate interpretation of these results is that black sellers face less discrimination in cities, where thicker online markets are more likely to be found. Residents of cities tend to be more racially diverse and younger (according to the 2000 Census), and may be more accustomed to interacting with people of other races and ethnicities. Because these market characteristics (thickness and urbanity) are highly correlated, we are unable to distinguish whether market competition has 'crowded out' discrimination, or whether buyers inclined against discrimination are merely more likely to live in thick, competitive markets.

discrimination is operative in this market due to buyers assuming that buying from a black seller is likely to involve greater inconvenience, black sellers should have worse outcomes in high-isolation markets. This would result in coefficients on the interactions between black and high isolation that are less than one in the negative binomial and less than zero in the OLS regressions.

Indeed, this is what we find. For example, black sellers receive 42% fewer offers (relative to white sellers) in high than in low isolation markets. The best offers black sellers receive are nearly \$27 lower in high-isolation markets than in markets with less racial isolation. Interestingly, tattooed sellers also receive significantly fewer offers in high-isolation markets, although the size of their offers is largely unaffected. We interpret this as further support for the hypothesis that buyers are using race (and the presence of a tattoo) as a proxy for living in a different, perhaps lower socioeconomic status, part of town, rather than indulging in taste-based discrimination.

3.2.3 *Advertisement text*

Each of our advertisements was randomly assigned either a high or low-quality text to provide a signal of the seller's socioeconomic status, proxied by his education level and writing ability. If low socioeconomic status is highly correlated with the characteristics that buyers are trying to avoid, discrimination should decrease in the presence of a high-quality advertisement. That is, if statistical discrimination against black sellers is operative, it should be smaller when advertisements are high quality; it might therefore manifest itself as coefficients on the interaction between black and high advertisement quality being greater than one in the negative binomial and greater than zero in the ordinary least squares regressions reported in Table 7.

This is fairly typical of statistical discrimination tests in experiments like this one, but it is not as successful here. The impact of our high-quality text treatment does not vary with race, and sometimes even has the wrong sign. Given the limited importance of our quality measure even in our average effects analysis, it seems likely that our low and high-quality advertisements are simply insufficiently different to affect response. Buyers might also have interpreted the low-quality text as signalling youth or 'hipness' instead of low socioeconomic status. To the extent that different buyers had opposing reactions to advertisement quality, the effects might have cancelled out. In any case, the results of this test are inconclusive.

3.3. *Testing for the Impact of Animus*

Stephens-Davidowitz (2013) creates a local measure of racial animus using the relative frequency of Google searches containing a racial slur. This measure had significant power in explaining voting behaviour in the 2008 presidential election between Barack Obama and John McCain. We consider whether black–white outcome gaps vary systematically with this measure as a test of whether the racial disparities we find are driven by animus and fail to find such evidence. The coefficients on the Google search variable – alone and interacted with black and tattooed – are approximately zero and statistically insignificant. We consider this further evidence that statistical discrimination is driving the racial disparities found in this market.

3.4. *Different Pools of Potential Buyers*

There are two possible ways to model the pool of buyers who respond to each seller type and their apparent discrimination between sellers: First, all buyers may be part of the same pool and their offers drawn from a single distribution, which has a lower mean valuation for iPods from black sellers than from white sellers. In this case, the buyers are actively discriminating between sellers by race. Alternatively, it is possible that buyers are self-segregated into separate pools that are more likely to respond to certain types of advertisements. Black sellers might then receive fewer or lower offers because the pools are different sizes, or have different valuation distributions (perhaps due to underlying characteristics like age, income or race). The offers to each seller type would then be drawn from these different distributions, perhaps producing worse outcomes for black sellers even though their own buyers are not discriminating against them. (Note that this still implies discrimination by the white sellers' buyers, who are choosing not to respond to black sellers' advertisements.)

So far we have implicitly assumed that the market functions as in the first case, with offers drawn from a single distribution. In this subsection we try to test the hypothesis that there are separate pools of buyers – in particular, that buyers show a preference for own-race sellers.

If buyers show a preference for own-race sellers, the local racial composition of a market will affect the size of the pool of potential buyers for black and white sellers. We test the hypothesis that black sellers' outcomes improve with the share of the local population that is black. If they do, the coefficient on % black \times black should be greater than one in the negative binomial regression and greater than zero in the OLS regression. Indeed, the regression results presented in Table 7 provide some evidence that black sellers do slightly better in markets where a larger share of the population is black. A 1 percentage point increase in the local black population increases the number of offers received by black sellers by 0.9% and the best offer received by 61 cents; this effect is statistically significant. (Less expectedly, it also increases the number of offers received by tattooed sellers by 1.5%.) We interpret this as consistent with the idea that part of the disparity found in our main results could be driven by buyers' preference for own-race sellers. Of course, it may also be that less discriminatory buyers live in communities with a larger black population (either by choice, or proximity makes them less discriminatory).

4. Conclusions

In this study, we present strong evidence that black sellers suffer worse market outcomes than their white counterparts in the environment we consider. In particular, their advertisements receive 13% fewer responses and 18% fewer offers. These effects are similar in magnitude to those associated with a seller's display of a wrist tattoo. A black seller's average offer is approximately \$5.72 lower than a white seller's, with an even greater difference in the highest offers: the best offer received by a black seller is typically \$7.07 lower. These represent gaps of 11% and 12%, respectively, below white sellers' offers.

Respondents to advertisements with black photographs also exhibit lower trust. Compared with correspondents with white sellers, they are 17% less likely to include

their name in their initial e-mail to the seller. Furthermore, the high bidders on black sellers' advertisements – presumably among the least biased of potential buyers – are 44% less likely to accept delivery by mail and are 56% more likely to express concern about making a long-distance payment.

By considering the degree to which black–white outcome disparities vary systematically with market and advertisement characteristics, we hope to shed some light on the various explanations of this observed discrimination. This exercise suffers both from limited statistical precision and an inability to sharply test theories of discrimination. The disadvantage faced by black sellers is greatly reduced in more competitive markets; this provides evidence in favour of Becker's hypothesis that discrimination can be competed away. Discrimination is greater in markets in which black and white residents are geographically isolated from one another and in markets with high property crime rates. This is consistent with statistical discrimination used to avoid a fraudulent, inconvenient or dangerous sale, although it is also possible that animus against black sellers is higher in high-crime or high-isolation markets. We also find evidence that black sellers do better in markets with larger black populations, suggesting that the disparities may be driven, in part, by buyers' preference for own-race sellers.

University of Virginia
Arizona State University

Additional Supporting Information may be found in the online version of this article:

Appendix A. Market, Advertisement and Timing Characteristics

Appendix B. Advertisement Text

Appendix C. E-mail Correspondence with Potential Buyers

Appendix D. Sample Responses from Buyers

Appendix E. Additional Results

References

- Ahmed, A.M. and Hammarstedt, M. (2008). 'Discrimination in the rental housing market: a field experiment on the internet', *Journal of Urban Economics*, vol. 64(2), pp. 362–72.
- Altonji, J.G. and Blank, R.M. (1999). 'Race and gender in the labor market', in (O.C. Ashenfelter and D.C. Card, eds.), *Handbook of Labor Economics*, vol. 3, pp. 3143–259, Amsterdam: Elsevier.
- Ayres, I., Banaji, M. and Jolls, C. (2011). 'Race effects on eBay', available at: <http://ssrn.com/abstract=1934432> (last accessed: 29 September 2011).
- Ayres, I. and Siegelman, P. (1995). 'Race and gender discrimination in bargaining for a new car', *American Economic Review*, vol. 85(3), pp. 304–21.
- Becker, G.S. (1971). *The Economics of Discrimination*, Chicago, IL: University of Chicago Press.
- Bertrand, M. and Mullainathan, S. (2004). 'Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination', *American Economic Review*, vol. 94(4), pp. 991–1013.
- Bosch, M., Carnero, M.A. and Farré, L. (2010). 'Information and discrimination in the rental housing market: evidence from a field experiment', *Regional Science and Urban Economics*, vol. 40(1), pp. 11–9.
- Carlsson, M. and Rooth, D.O. (2007). 'Evidence of ethnic discrimination in the Swedish labor market using experimental data', *Labour Economics*, vol. 14(4), pp. 716–29.
- Carpusor, A.G. and Loges, W.E. (2006). 'Rental discrimination and ethnicity in names', *Journal of Applied Social Psychology*, vol. 36(4), pp. 934–52.
- Ewens, M., Tomlin, B. and Wang, C. (forthcoming). 'Statistical discrimination or prejudice? A large sample field experiment', *Review of Economics and Statistics*.

- Glaeser, E.L. and Vigdor, J.L. (2001). 'Racial segregation in the 2000 census: promising news', Center on Urban and Metropolitan Policy, The Brookings Institution.
- Hanson, A. and Hawley, Z. (2011). 'Do landlords discriminate in the rental housing market? Evidence from an internet field experiment in US cities', *Journal of Urban Economics*, vol. 70(2–3), pp. 99–114.
- Kass, L. and Manger, C. (2012). 'Ethnic discrimination in Germany's labour market: a field experiment', *German Economic Review*, vol. 13(1), pp. 1–20.
- List, J.A. (2004). 'The nature and extent of discrimination in the marketplace: evidence from the field', *Quarterly Journal of Economics*, vol. 119(1), pp. 49–89.
- List, J.A. and Livingston, J.A. (2010). 'When the weak become weaker: the effect of market power on third degree price discrimination', available at: http://works.bepress.com/jeffrey_livingston/11/ (last accessed: 31 January 2010).
- Nardinelli, C. and Simon, C. (1990). 'Customer racial discrimination in the market for memorabilia: the case of baseball', *Quarterly Journal of Economics*, vol. 105(3), pp. 575–95.
- Neumark, D., Bank, R.J. and Van Nort, K.D. (1996). 'Sex discrimination in restaurant hiring: an audit study', *Quarterly Journal of Economics*, vol. 111(3), pp. 915–41.
- Nunley, J.M., Owens, M.F. and Howard, R.S. (2011). 'The effects of information and competition on racial discrimination: evidence from a field experiment', *Journal of Economic Behavior and Organization*, vol. 40(6), pp. 670–9.
- Pope, D.G. and Sydnor, J.R. (2011). 'What's in a picture? Evidence of discrimination from prosper.com', *Journal of Human Resources*, vol. 46(1), pp. 53–92.
- Ravina, E. (2008). 'Love & loans: the effect of beauty and personal characteristics in credit markets', *American Law & Economics Association Annual Meeting Papers*, vol. 2008(67), pp. 1–74.
- Riach, P.A. and Rich, J. (2002). 'Field experiments of discrimination in the market place', *ECONOMIC JOURNAL*, vol. 112(483), pp. F480–518.
- Stephens-Davidowitz, S. (2013). 'The effects of racial animus on a black presidential candidate: using Google search data to find what surveys miss', available at: <http://ssrn.com/abstract=2050673> (last accessed: 11 June 2012).
- Tversky, A. and Kahneman, D. (1974). 'Judgment under uncertainty: heuristics and biases', *Science*, vol. 185(4157), pp. 1124–31.
- United States Department of Justice and Federal Bureau of Investigations (2009). 'Crime in the United States, 2008', available at: <http://www2.fbi.gov/ucr/cius2008/> (last accessed: 14 September 2009).