UCLA
UCLA Electronic Theses and Dissertations

Title

Essays in Behavioral Economics

Permalink

Author

Fiorin, Stefano

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Behavioral Economics

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in Management

by

Stefano Fiorin

2017

ABSTRACT OF THE DISSERTATION

Essays in Behavioral Economics

by

Stefano Fiorin

Doctor of Philosophy in Management

University of California, Los Angeles, 2017

Professor Leonardo A. Bursztyn, Chair

This dissertation is comprised of three essays in behavioral economics. These essays share common empirical methodologies and intellectual themes. First, each essay uses randomized controlled trials, in the form of natural field experiments or online lab experiment. Second, each paper attempts to measure economically important but difficult-to-observe determinants of behavior – moral considerations in Chapter 1, social status concerns in Chapter 2, and social norms in Chapter 3.

In Chapter 1, coauthors Leonardo Bursztyn, Daniel Gottlieb, Martin Kanz and I study the role of morality in debt repayment using an experiment with credit card customers of a large Islamic bank in Indonesia. In our main treatment, clients receive a text message stating that "non-repayment of debts by someone who is able to repay is an injustice." This increases the share of customers meeting their minimum payment by 15%, which is more than the effect of substantial financial incentives. Additional treatments help understand

the underlying mechanisms and rule out competing explanations, such as reminder effects, priming religion, signaling the lender's commitment to debt collection, and provision of new information.

In Chapter 2, coauthors Leonardo Bursztyn, Bruno Ferman, Martink Kanz, Gautam Rao and I provide novel field-experimental evidence on status goods. We work with an Indonesian bank that markets platinum credit cards to high-income customers. In a first experiment, we show that demand for the platinum card greatly exceeds demand for a nondescript control product with identical benefits, suggesting demand for the pure status aspect of the card. Transaction data reveal that platinum cards are more likely to be used in social contexts, implying social image motivations. Combining price variation with information on the use of the card sheds light on the magnitude of the demand for social status. In a second experiment, we provide evidence of positional externalities from the consumption of these status goods. The final experiment shows that increasing self-esteem causally reduces demand for status goods. We infer that part of the demand for status is psychological in nature, and that social image is a substitute for self-image.

Social norms are typically thought to be persistent and long-lasting, sometimes surviving through growth, recessions, and regime changes. In some cases, however, they can quickly change. In Chapter 3, coauthors Leonardo Bursztyn, Georgy Egorov and I examine the unraveling of social norms in communication when new information becomes available, e.g., aggregated through elections. We build a model of strategic communication between citizens who can hold one of two mutually exclusive opinions. In our model, agents communicate their opinions to each other, and senders care about receivers' approval. As a result, senders are

more likely to express the more popular opinion, while receivers make less inference about senders who stated the popular view. We test these predictions using two experiments. In the main experiment, we identify the causal effect of Donald Trump's rise in political popularity on individuals' willingness to publicly express xenophobic views. Participants in the experiment are offered a bonus reward if they authorize researchers to make a donation to an anti-immigration organization on their behalf. Participants who expect their decision to be observed by the surveyor are significantly less likely to accept the offer than those expecting an anonymous choice. Increases in participants' perceptions of Trump's popularity (either through experimental variation or through the "natural experiment" of his victory) eliminate the wedge between private and public behavior. A second experiment uses dictator games to show that participants judge a person less negatively for publicly expressing (but not for privately holding) a political view they disagree with if that person's social environment is one where the majority of people holds that view.

The dissertation of Stefano Fiorin is approved.

Paola Giuliano

Adriana Lleras-Muney

Nico Voigtlander

Romain T. Wacziarg

Leonardo A. Bursztyn, Committee Chair

University of California, Los Angeles

2017

Contents

List of Figures

List of Tables

Acknowledgments

Review Board.

Education

Laurea Magistralis in Economics, *cum laude*   2010-2012
University of Bologna, Bologna, Italy
Laurea Triennale in Economia e Commercio, *cum laude*   2007-2010
University of Padova, Padua, Italy

Research Interests

Primary: Development Economics and Behavioral Economics
Secondary: Political Economy and Household Finance

Working Papers

"Moral Incentives in Credit Card Debt Repayment: Evidence from a Field Experiment",
with Leonardo Bursztyn, Daniel Gottlieb, and Martin Kanz.
"Status Goods: Experimental Evidence from Platinum Credit Cards", with Leonardo Bursztyn, Bruno Ferman, Martin Kanz, and Guatam Rao.
"From Extreme to Mainstream: How Social Norms Unravel", with Leonardo Bursztyn and
Georgy Egorov.
"Corruption and Legislature Size: Evidence from Bazil", with Diogo Britto.

Research Grants

Center for Global Management (CGM) UCLA Anderson PhD Global Research Award 2016-
2017 ($4,000). Project: "Preference Based Home Bias: the Effect of Cultural Proximity on
Loan Repayment (Philippines)."
Center for Global Management (CGM) UCLA Anderson PhD Global Research Award 2015-
2016 ($4,000). Project: "Speaking the Same Language: the Effect of Cultural Proximity on
Loan Repayment (Kenya)."
Abdul Latif Jameel Poverty Action Lab (J-PAL) Governance Initiative Pilot Study Research
Grant ($49,050). Project: "Participatory Budgeting and Urban Services: Mohalla Sabhas in
Delhi."
Center for Global Management (CGM) UCLA Anderson PhD Global Research Award 2014-
2015 ($4,000). Project: "Speaking the Same Language: the Effect of Cultural Proximity on
Loan Repayment (Indonesia)."

Fellowships and Awards

Best Paper Award, KFUPM Islamic Banking and Finance Research Conference Riyadh 2016
Dissertation Year Fellowship, UCLA Graduate Division   2016-2017
Dean's Award for early Advancement to Candidacy, UCLA Anderson   2015-2016
Anderson Fellowship, UCLA Anderson   2012-2016
Laurea Magistrale in Economics Best Student Award, University of Bologna   2012

Seminars and Invited Presentations

Pacific Conference for Development Economics (Pac Dev 2017). UC Riverside.   2017
Bocconi Baffi Carefin Annual International Banking Conference. Bocconi University.   2016
Microeconometric and Public Policy Seminar. The University of Sydney.   2016
Warwick Economics PhD Conference 2016. The University of Warwick.   2016

Teaching Experience

TA for Managerial Economics (MBA students, Nico Voigtlaender). UCLA.   2013-2017
TA for Statistics (MA students, Paola Bortot). University of Bologna.   2011-2012
TA for Mathematics (BA students, Enrico Bernardi). University of Bologna.   2010-2011

Employment

Short Term Consultant, The World Bank.   2017
RA for Leonardo Bursztyn, UCLA Anderson.   2013-2017
Short Term Consultant, The World Bank.   2014-2015
RA for Rocco Macchiavello and Chris Woodruff, University of Warwick and IPA-Bangladesh.
2012
RA for Andrea Ichino, Margherita Fort and Giulio Zanella, University of Bologna. 2011-2012
RA for Erich Battistin, Sascha Becker, and Luca Nunziata, University of Padova.   2010

Chapter 1

Moral Incentives in Credit Card Debt Repayment: Evidence from a Field Experiment

1.1 Introduction

The ability to collect debts is one of the main pillars of any financial system. While economists have extensively examined the importance of screening, monitoring, and reputational considerations, little attention has been paid to the role of morality in establishing a norm of debt repayment. Still, from ancient philosophy to contemporary news media, there are countless references to the moral aspects of debt and debt repayment. In Plato's *Republic*, Socrates defines justice as "telling the truth and paying one's debts."[1] More recently, the burst of the real estate bubble left many observers puzzled by the fact that surprisingly few homeowners defaulted on mortgages whose value exceeded that of the property, while others suggested that moral considerations may have played an important role in these decisions (see Guiso, Sapienza, and Zingales, 2013). Similarly, a vocal debate over the morality of failing to repay one's student loans has been featured prominently in major newspapers.[2] Issues of morality have also played an important role in the context of sovereign debt, for example in the heated discussions on public debt and defaults in Argentina or Greece.[3]

---

[1]There are also numerous references to the morality of debt in religious texts. An example from the Bible is *Romans 13:7-8*: "Give to everyone what you owe them [...] and let no debt remain outstanding." An example from Islam is *Shahih al-Bukhari 3:575*: "[...] The best among you are those who repay their debts handsomely." Many languages, including German and Hebrew, share the same word for "debt" and "guilt." Nietzsche offers a detailed account of this association and its influence on the development of social norms in *The Genealogy of Morals* (1887).

[2]See, for example, Lee Siegel "Why I Defaulted on My Student Loans", *New York Times*, June 6, 2015. "Times Op-Ed Goes All In On Student Debt Silliness", *Forbes*, June 8, 2015.

[3]The prevalence of usury laws throughout history illustrates that moral issues regarding debt are not specific to the debtor's side. In the context of sovereign debt, philosophers have questioned not only the morality of default but also the morality of debt itself. The French philosopher Montesquieu, for example, argued that sovereign debt is fundamentally immoral because it "takes the true revenue of the state from those who have activity [...], to convey it to the indolent." Moral arguments have also played a prominent role in debates on debt forgiveness for highly indebted poor countries. See, for example, William Easterly "Debt Relief". *Foreign Policy*, December 2001.

In this paper, we study the role of morality in debt repayment, one of the most important financial decisions faced by the household. Over the last decades, the ratio of household debt to GDP has grown dramatically worldwide, and high interest consumer credit, including credit card debt, accounts for a large share of this increase both in the U.S. and in emerging markets (see Mian and Sufi, 2014, 2015 and Zinman, 2015). Many studies have documented patterns of inefficient borrowing and debt repayment among households in various settings (see, for example, Agarwal et al., 2009a,b and Stango and Zinman, 2015).[4] However, much less is known about the factors that determine individual debt repayment decisions, and the extent to which non-monetary considerations, including issues of morality, may affect repayment behavior.

To study these issues, we use a field experiment with the universe of late-paying customers of the most popular Islamic credit card in Indonesia, the world's largest Muslim country. Islamic banking is a large and rapidly growing industry in Indonesia and around the world, with more than 300 banks in over 75 countries and approximately US$ 1.5 trillion in assets (World Bank, 2014). Islamic banks offer a range of standard financial products that comply with the principles of Islamic law and often emphasize the ethical dimension of their business model.[5]

The credit card in our experiment is issued by one of Indonesia's leading Islamic banks,

---

[4]For example, households often hold liquid assets and high interest rate debt simultaneously. Existing evidence also suggests that households often fail to choose the lowest cost credit contracts and to prioritize repayment accordingly.

[5]References to moral values are also used in other areas of finance. Many investment management firms offer socially responsible investment (SRI) products that do not invest in "sin stocks," including purveyors of alcohol, tobacco, and gambling, or firms linked to unethical practices. Examples include the *HSBC Ethical Global Equity Fund* or the *iShares Human Rights Fund*. SRIs account for approximately US$ 5 trillion in assets worldwide.

which is part of a large non-religious conglomerate that targets a relatively secular customer segment. The card has all features of a standard credit card and is functionally equivalent to credit cards issued by non-Islamic banks.[6] Prior to our study, the bank had independently introduced a mobile phone text messaging system that automatically sends reminders to customers who have not made the required minimum payment one day after the due date. Between February 2015 and April 2016, we worked with the bank to develop a second set of messages that included basic reminders as well as moral appeals. These messages were randomly assigned at the individual customer level and sent to late-paying customers two days before the end of a ten-day grace period. A control group received only the first neutral reminder.

In the main treatment condition of our experiment, late-paying customers received a text message containing a moral appeal, which highlights that not repaying a debt when one is able to repay violates a moral norm. The text refers to the Islamic doctrine on non-repayment of debts using a quote from the *Shahih-al-Bukhari*, one of the main religious texts of Sunni Islam, which serves as a source of Islamic law and is widely known among Indonesian Muslims:[7,8]

---

[6]Not all clients of Islamic banks are driven by religious motivations. In fact, 10 percent of credit card clients at our partner bank are non-Muslims. This is roughly the same as the share of non-Muslims in the Indonesian population. Many non-Muslim customers seem to be attracted by the zero overdraft fees the bank charges. While the card has no explicit interest rates, it charges fees proportional to the balance so that the pricing is similar to credit cards outside Islamic finance. We discuss the institutional details of our setting in Section 1.2 below.

[7]The *Shahih-al-Bukhari* is one of the six major *hadith* collections of Sunni Islam (*Kuttub al-Sittah*). It reports on the sayings, deeds, and teachings of the Prophet and was transcribed by the Persian scholar Muhammad al-Bukhari after being transmitted orally for centuries. It is considered one of the most important texts in Sunni Islam and is widely used in the application and interpretation of Islamic law.

[8]This quote was suggested to us by our partner bank. The bank had previously used it on a smaller scale in debt collection calls to delinquent customers and ensured us that customers found it appropriate. The available evidence, described in more detail below, indicates that customers indeed had no objections to the

The Prophet (Peace and blessings be upon Him) says: "non-repayment of debts by someone who is able to repay is an injustice." (Imam al-Bukhari) Please repay your credit card balance at your earliest convenience. Call [customer service number].

The design of our experiment has several important features that help us identify the effect of moral appeals on debt repayment decisions. First, debt repayment is a common financial decision with high stakes. We are able to study this decision directly, by designing a real-stakes field experiment that is integrated into the credit card repayment cycle of a large bank. Second, we implement the experiment using text messages sent through the bank's automated system. This enables us to address the moral appeal to delinquent customers directly. Third, the bank routinely uses text messages to communicate with its customers, and religious or moral content, such as that used in our experiment, is not uncommon in these messages. Therefore, both the channel of communication and the content of the messages used in the experiment are credible and natural in this setting.[9] Finally, many moral appeals used in practice rely on a reference to a moral authority (such as religion, family values, or the law). Our main treatment consists of two separate components: an appeal to a moral authority (a religious text quoting the Prophet) and a moral statement ("non-repayment of

content of the message: 80% of recipients stated that they would like to receive a similar message again in the future. We also find no reduction of card usage or transaction volumes in the weeks after customers first receive the message.

[9]Messages with moral suasion content that seek to incentivize repayment have also been frequently used by banks in other settings. In India, for example, banks have aired television and radio commercials with moral appeals made by children in an effort to persuade defaulting borrowers to repay their loans. See "Banks Make Emotional Appeals to Get Borrowers to Repay Loans " *Live Mint*, October 2016. Some firms in the United States and Europe have also used religious content in messages to their customers. See, for example, "Alaska Airlines Ends Decades-old Prayer Card Tradition," *Reuters*, January 2012.

debts by someone who is able to repay is an injustice"). Additional treatments, designed to unpack the mechanism through which moral appeals affect repayment, remove the first component, allowing us to isolate the moral statement from explicit references to religion and test whether an appeal to a moral norm without reference to a moral authority can affect behavior. Hence, while we use an Islamic credit card to obtain a setting in which content and framing of the moral appeal are natural, our experimental design allows us to identify the impact of religious context separately from the impact of the moral appeal itself.

We document a strong effect of moral appeals on debt repayment. In our preferred specification, receiving the moral message raised the share of customers meeting their minimum payment by about 15% above the 34% of customers making the minimum payment by the deadline in the control group.[10] In order to assess the economic magnitude of our main result, we benchmark the impact of the moral message against the effect of direct financial and reputational incentives. Our first benchmark is a financial incentive treatment in which late-paying customers were sent a message offering a cash rebate equal to 50% of the outstanding minimum payment – or 5% of the customer's current debt – for making a payment before the deadline.[11] The cash rebate treatment increased repayment by 7% relative to the control group, which is about half of the effect of receiving the moral message. Since a text message had to be sent in both cases, but the rebate had additional costs, the moral appeal was significantly more cost-effective than these direct financial incentives. Our second

---

[10]In a typical month, approximately 90% of credit card customers made the required minimum payment before the time of our intervention, which always occurred two days before the end of the grace period. Prior to the experiment, only about 5% of customers remained more than one month overdue on their minimum payment each month.

[11]We offered a cash rebate, instead of a discount on current payments, to avoid liquidity constraint effects. Hence, the treatment focuses on customers' willingness, rather than ability to repay.

benchmark examines the effect of informing customers about the reputational consequences of non-repayment. To do so, a group of customers received text messages that informed them about the existence of the Indonesian credit registry and the negative consequences of being reported for late repayment. The reputational incentive treatment raised the probability of meeting the minimum repayment by nearly 30%. These results suggest that both moral and reputational considerations affect repayment behavior in our setting.[12]

To better understand how the moral message affects repayment decisions, we conduct a series of additional interventions that allow us to test and rule out a number of alternative mechanisms. First, is the impact of the moral message simply due to a reminder effect? In order to address this possibility, a group of customers were sent another simple reminder message that did not contain a moral appeal. This message had no significant effect on repayment, ruling out this channel.

Second, does the moral message work because it primes customers on religion or evokes a religious frame of mind? To answer this question, a group of customers were sent a religious placebo message, which included a quote from the Prophet that was taken from the same religious text as the moral message but made no reference to debt repayment. This message also had no effect on repayment. It is also worth noting that the simple reminder and the religious placebo message were both newly designed text messages that had never before been received by any of the bank's customers. The fact that neither of these messages affects repayment also rules out that the effect of the moral appeal is due to the surprise of

---

[12]Our result that customers care and respond to reputational incentives is consistent with recent findings from the literature (see, for example, Liberman, 2016). In Section 1.4.3, we provide evidence suggesting that intensive margin effects (i.e. amount repaid) are stronger in the moral incentive group than in the reputational incentive group.

receiving a novel or attention-grabbing message.

Third, does the moral appeal only work in a religious context or when using language that has a religious connotation? To answer this question, the bank sent two additional variations of the moral message. While the original moral incentive message explicitly quoted the Prophet, cited the religious text from which the quote was taken, and employed a word of Arabic origin for "injustice" that is typically used in a religious context, the two additional variations of the moral message removed the religious elements of the moral appeal. The first message omitted the reference to the Prophet and the religious text from which the quote was taken. The second message additionally replaced the Arabic-origin word for "injustice" used in the moral message with the standard Indonesian word, which has no religious connotation. The first message allows us to test whether invoking a credible religious source increases the effectiveness of the moral appeal, and sheds light on the mechanism through which references to a moral norm affect behavior. The second message tests whether the moral appeal without religious connotation affects repayment.

We find that all variations of the moral appeal have the same effect. That is, a non-religious moral statement is just as powerful as the same moral statement identified as a quote from the Prophet and attributed to a well-known religious text. This indicates that either customers already associated the moral appeal contained in the message with religion (and potentially with the Prophet) or that the pure moral statement was sufficient to trigger repayment. To disentangle these two explanations, we conducted an end-line survey in which a bank employee read the non-religious version of the moral message to customers in the control group and asked if they associated the message with religion. The vast majority of

respondents did not associate the quote with a religious source. This finding indicates that our results are driven primarily by the moral appeal, rather than the religious nature of the message.

Fourth, does the moral appeal work because receiving a strongly worded message signals that the bank is committed to enforcing debt collection? To test this possibility, we surveyed customers who had received either no message, the basic reminder, or one of the versions of the moral message one day after the payment deadline. Customers were asked "How committed do you think [bank name] is to collect debt from delinquent customers on a scale from 1 to 5 (where 1 is not committed and 5 is very committed)?" There is no statistically significant difference in the response to this question between customers assigned to the different treatments.

Finally, does the moral message work only once, for example because it is novel or because it conveys new information, or would it work if it were sent repeatedly? To explore this question, the bank sent the moral message to consumers who reappeared on the late payers list and had already received a moral message once. We find that moral messages are effective even when they are resent after as little as two months. Of course, customers who appear on the list a second time are a selected sample, so a clean comparison with the effect of the first moral message is not possible. To partly address this issue, we control for a number of observables to address the differential selection of repeated late-paying customers. We find that the effect of receiving the moral message a second time is nearly identical to the effect of receiving the message for the first time.

This suggests that the effect of the moral message cannot be explained by the novelty of

the message. Our findings thus also rule out an explanation for the effectiveness of the moral message based on new information, such as customers learning about a religious teaching or social norm that they were not previously aware of. The repayment rate does not change when the information content of the message is reduced either by excluding religious language or by excluding the reference to a well-known religious text. Moreover, the effect of receiving the same message for the second time is nearly identical to the effect of receiving it for the first time. If the effects were driven by the provision of new information, the message would affect repayment only when a late-paying customer receives it for the first time. While we cannot test for this channel directly, our results appear to be consistent with an inattention interpretation, in which the moral incentive message temporarily draws customers' attention to moral considerations. In line with this interpretation, we also show that the effect of the moral message is not persistent. That is, receiving the message once does not affect repayment in subsequent months.[13]

Taken together, these findings are consistent with the idead that even when making important financial decisions, people experience a cost from consciously violating a moral norm, so that a moral appeal can affect behavior even when it is not associated with an explicit threat of punishment or negative financial consequences.

This paper relates to several strands of the literature. First, our work contributes to a literature on consumer financial behavior that has examined debt accumulation and re-payment decisions (see Agarwal et al., 2009a,b; Bertrand and Morse, 2011; Zinman, 2015). Several studies in this literature have found that consumers often make suboptimal financial

---

[13]More broadly, our findings relate to a recent line of research that models what individuals pay attention to, and how this influences their decisions (Bordalo et al., 2013; Kőszegi and Szeidl, 2013; Gabaix, 2014).

decisions (Stango and Zinman, 2009, 2015; Dobbie and Skiba, 2013; Dobbie and Song, 2016) and have explored how non-traditional regulation and incentives can help consumers make better financial choices.[14] There is also evidence to suggest that moral considerations play a potentially important role in debt repayment decisions. Guiso, Sapienza, and Zingales (2013) use survey data to study attitudes toward strategic default on mortgages among households in the United States. They find that 82% of respondents believe that it is morally wrong to engage in strategic default, and that these respondents are about 10 percentage points less likely to declare strategic default on their mortgages.

Second, our work is related to a large literature on non-monetary incentives (Frey, 1997; Akerlof and Kranton, 2000a; Gneezy, 2005; Bénabou and Tirole, 2003, 2006). In particular, we shed light on how moral appeals affect an important financial decision. Moral appeals, directed to the audience's sense of what is right and proper, are among the most common strategies of persuasion. Many companies advertise their support for fair trade or charitable causes to affect consumer choices. Most closely related to our setting, a number of banks have used television commercials with moral suasion content to get delinquent borrowers to repay their debt, and financial consumer protection agencies have used moral appeals in to encourage responsible financial behavior.[15] Although moral appeals are widely used in this context in practice, there is relatively little evidence about how and especially why they work. By examining how a direct moral appeal affects debt repayment and exploring the underlying mechanisms, our paper contributes to a recent experimental literature on the

---

[14]For examples of non-traditional incentives and regulation intended to improve financial decisions see, for example, Madrian and Shea, 2001; Benartzi and Thaler, 2004; Campbell et al., 2011; Agarwal et al., 2014.

[15]See, for example, "Banks Make Emotional Appeals to Get Borrowers to Repay Loans " *Live Mint*, October 2016.

effects of moral suasion and normative incentives (see Dal Bó and Dal Bó, 2014; Pruckner and Sausgruber, 2013; Hallsworth et al., 2014, 2015; Fellner et al., 2013; Ito et al., 2015).

Beyond helping to understand the impact of moral suasion in financial decisions, our work also relates to a literature on religion and economic behavior (see Iannaccone, 1998; Barro and McCleary, 2006, Clingingsmith, Khwaja, and Kremer, 2009, Cantoni, 2015; Campante and Yanagizawa-Drott, 2015; Bénabou et al., 2015; Benjamin et al., 2016). Identifying the effect of moral appeals linked to religion is difficult because religious activities often combine moral, instrumental, and social motivations. For example, people may go to church because they believe it is the "right thing to do," but they may also do so for indirect material or social benefits, such as socializing with others or signaling one's beliefs or shared values. For example, people may go to church because they believe it is the "right thing to do," but may also do so for indirect material or social benefits, such as socializing with others, or signaling one's beliefs or shared values. Our paper adds to this literature by providing evidence that moral motivations associated with religion can drastically affect behavior in a setting where the social interactions usually associated with religion are absent.[16]

The remainder of the paper proceeds as follows. In Section 1.2, we describe the setting and experimental design. Section 3.3.3 presents the main results. Section 1.4 discusses the interpretation of our findings, and Section 1.5 concludes.

---

[16]Many laboratory experiments have also shown that religious primes can induce prosocial behavior, increasing the amount shared in dictator games (Shariff and Norenzayan, 2007), reducing cheating (Randolph-Seng and Nielsen, 2007; Mazar et al., 2008), and increasing charitable donations (Pichon et al., 2007). Other work shows that priming religion increases punishment of unfair behavior, but only among religiously committed subjects (McKay et al., 2011; Laurin et al., 2012).

1.2 Experimental Design

1.2.1 The Credit Card

We design a natural field experiment with the universe of late-paying customers of Indonesia's most popular Islamic credit card. The credit card is issued by one of the country's leading Islamic banks, which offers credit cards as part of its portfolio of Islamic consumer finance products. Originally introduced in 2009, the card had approximately 200,000 customers at the time of our experiment.

The credit card features are designed to comply ith the principles of Islamic *Shari'a* law which, among other prescriptions, , prohibits charging interest and investing in activities considered contrary to the principles of Islam. In order to be fully consistent with Islamic law, the features of the card are based on a *fatwa* (legal decree) issued in 2006 by the *Indonesian Council of Islamic Scholars* that lays out the guidelines under which banks can offer *Shari'a* compliant credit cards. Following these guidelines, the credit card is structured as an *Ijara* fee structure contract, which means that customers pay a fee for the transaction services provided by the card instead of a variable interest rate. Customers are charged fixed annual fees of Rp 120,000 (US\$ 10) for a basic card, Rp 240,000 (US\$ 20) for a gold card, and Rp 600,000 (US\$ 45) for a platinum card, plus a monthly membership fee of 2.75% of the customer's credit limit. This monthly fee can be partially or fully waived through a "cash rebate," which is proportional to the customer's available credit and can range from zero to the total amount of the monthly fee.[17] The monthly fee is waived entirely if there is

---

[17]The cash rebate is calculated as follows: *cash rebate = 2.75% × (credit limit - amount outstanding).* The net monthly fee is the monthly membership fee minus the cash rebate, that is, 2.75% × *amount outstanding.*

no outstanding debt.

There is a monthly billing cycle, with a billing date on the eighteenth day of each month. The minimum monthly payment, equal to either 10% of the customer's total outstanding balance or Rp 50,000 (whichever amount is higher) plus eventual arrears and overdrafts, is due on the eighth day of the following month. Customers who do not meet the minimum payment by the due date receive a text message from the bank on the following day. The bank grants late-paying customers a grace period of ten days, which ends on the eighteenth day of each month (we refer to this date as the "repayment deadline"). Customers who do not make the minimum payment by this date are considered "delinquent" and are reported to the Indonesian credit registry, the *Sistem Informasi Debitur*, which all banks in Indonesia consult before issuing credit. On the same day, they receive a phone call from the bank. They are charged a nominal late payment fee, ranging from Rp 15,000 to Rp 35,000 and the card is automatically blocked.[18] Once the customer makes the minimum payment, the card is immediately unblocked. If a customer's minimum payment remains outstanding for more than 90 days after the due date, the card is permanently blocked and the account is closed. Accounts that remain more than 120 days overdue are sent to the bank's collections department and, eventually, an outside collection agency. Figure 1.1 summarizes the credit card billing cycle and the timeline of our intervention.

---

[18]Late payment fees increase over time. For example, customers who are more than 30 days late are charged additional fees ranging from Rp 20,000 to Rp 50,000.

1.2.2 Sample Population and Random Assignment

The population for our experiment comprises the 14,429 credit card customers who were more than one week late on their minimum payment at least once during one of the six months between February 2015 and April 2016 in which the experiment was carried out.[19] Because some customers were late more than once during this period, there are 23,520 observations in our sample frame.[20]

The experiment was conducted in six waves, coinciding with the monthly credit card repayment cycle.[21] Each month, the bank shared with us the list of customers that had not made the minimum required payment by the sixteenth day of the month, or two days before the final repayment deadline. In the main experiment, we excluded from this list all customers who had previously received a text message treatment. Customers assigned to the control group in a previous month remained in the sample and could either be assigned to one of the treatments or form part of the control group again. For example, in March 2015, 4,803 customers were more than a week late. Out of these, 1,018 had previously received a treatment message and were thus excluded from the sample; the remaining 3,785 customers

---

[19]The experiment was conducted in February, March, May, and June 2015, and February and April 2016. We originally planned to have a treatment group receiving restructuring offers in April 2015, but the partner bank was not able to operationalize this. Upon agreement with the bank, we then decided to pause our main intervention in April 2015 and to resume it in May 2015. We also ran a small pilot with 250 customers in January 2015 that had results similar to those in our main intervention.

[20]In the universe of 14,429 customers, 8,691 were late only once, while the remainder appeared in our sample more than once: 3,052 customers were late twice, 1,414 were late three times, 579 four times, 191 five times, and 52 were late in all six months.

[21]The first two waves of the experiment were conducted in February and March 2015. The last three waves were conducted in June 2015, and February and April 2016. As part of a parallel experiment for a second paper, we had two other treatment groups with customers receiving multiple text messages on the same day. We excluded those 2,200 observations from our analysis. Results are unaffected when these observations are included, and are displayed in Table A.4 in the Supplemental Appendix. In the notes to Table A.4 we also discuss some design and implementation issues which affect that interpretation of the effect of these additional treatments.

were assigned to one of the treatment conditions or the control group. Following this process, we obtain a dataset that includes 13,428 observations, representing 12,104 unique credit card customers.[22]

Eligible customers were randomly assigned to one of several treatment conditions or to a control group. As part of the bank's standard communications policy, all customers received a neutral text message reminder one day after they had missed the minimum payment. The 4,120 customers assigned to the control group received no other text from the bank, while the 9,308 customers assigned to one of the treatment conditions received additional information through a text message sent two days before the repayment deadline. All treatments were randomly assigned at the individual customer level and delivered through text messages, using the bank's existing customer notification system.[23] Appendix Figure A.2 summarizes the experimental design.

In February and April 2016, we conducted a separate follow-up experiment with the 898 customers who reappeared on the list of late payers and had previously received the moral message as part of the main experiment. The experiment was designed to test if the moral message only works the first time it is sent, for example because it is novel or conveys new information, or if sending the message repeatedly can still affect repayment. Following the

---

[22]Of these 13,428 observations, 10,903 customers appear on the list of late-payers only once, 1,088 appear twice (the first time in the control group), 104 appear three times (the first two in the control group), 6 appear four times (the first three in the control group), and 1 customer appears 5 times (the first four times in the control group). Although this approach does not affect the internal validity of our analysis, it could potentially reduce the representativeness of our sample, since in a given month, customers who received a previous treatment message could have been part of the list of late payers if they had been assigned to the control group instead. However, given that the effect of our treatments is very similar for subjects appearing in the sample for the first time and those previously assigned to the control group, re-weighting the sample to correct for the probability of being excluded does not affect our results.

[23]All messages were in Bahasa Indonesia, the official language of Indonesia, which is also the standard language used by the bank in all of its customer communications.

same procedure and timing as above, recurrent late payers were randomly assigned either to a control group or to a repeated message treatment group.[24] The 450 customers assigned to the control group again only received a neutral reminder one day after they missed the minimum payment. The 448 customers assigned to the repeated moral message treatment group received a moral message identical to the one they had previously received. As in the main experiment, this message was sent two days before the repayment deadline.

1.2.3 Experimental Treatments

Control Group

A total of 4,120 customers were assigned to the control group, which forms the basis of comparison throughout the experiment. Customers in this group received a single reminder one day after they had missed the required minimum monthly payment:

> *Your [name of the card] has reached the due date. Please make a payment at your earliest convenience. If you have already paid, ignore this text. Call [customer service number].*

While all other customers received an additional message from the bank two days before the repayment deadline, customers in the control group received only this initial reminder.

---

[24]We stratify on how recently the customer had received the first moral message: 364 customers were treated two months before reappearing in the late-payer list, while the other 534 customers were treated for the first time between eight and fourteen months before.

Moral Incentives

To test the impact of moral appeals, we assigned 2,244 participants to the *moral incentive* treatment condition. In addition to the basic reminder sent to all customers who missed the due date, these customers received an additional message drawing attention to the religious implications of not repaying their debts. The message quotes from the *Shahih al-Bukhari*, one of the main religious texts of Sunni Islam, which reports of the teachings, deeds, and sayings of the Prophet Muhammad and serves as one of the main sources for the interpretation of Islamic law. The quote highlights the religious doctrine on repayment of debts and asks the customer to repay her outstanding balance:

> *The Prophet (Peace and blessings be upon Him) says: "non-repayment of debts by someone who is able to repay is an injustice" (Imam al-Bukhari). Please repay your credit card balance at your earliest convenience. Call [customer service number].*

To better understand the mechanisms underlying the impact of moral appeals, the bank sent two additional variations of this treatment, which varied the degree of its religious content. The first variation of the message (the *implicit moral incentive* condition) removed the reference to the Prophet and the text from which the quote was taken, but kept the Arabic-origin word for "injustice" from the original quote, which may be associated with religion. This message, assigned to 1,186 customers, reads:

> *Non-repayment of debts by someone who is able to repay is an injustice. Please repay your credit card balance at your earliest convenience. Call [customer service*

number].

The second variation of the message (the *non-religious moral incentive* condition), which was assigned to 1,180 customers, not only omitted the reference to the Prophet and the source of the quote, but also replaced the Arabic-origin term for "injustice" (*kezaliman*) with the standard Indonesian word (*ketidakadilan*), which has no religious connotation.

The first variation of the moral message allows us to test whether a moral appeal is strengthened by invoking a credible religious source. The second message tests whether receiving a moral statement without any religious connotation can affect the repayment decision.

Direct Financial Incentives: Cash Rebate

To benchmark the effect of moral appeals against financial incentives, we implemented two treatment conditions. The first treatment consisted of a direct one-time financial incentive in the form of a large cash rebate. In this *cash rebate incentive* condition, the bank sent the standard reminder on the due date and an additional message two days before the repayment deadline in which customers were offered a rebate equal to 50% of their minimum payment (5% of their total outstanding balance), if they made the required minimum payment by the deadline. The rebate would then be credited to their account in the next billing cycle.[25] This message, assigned to 336 participants in June 2015, reads as follows:

---

[25]We worked with the bank to design a rebate that customers would easily understand based on their previous experience. In general, clients in our sample are familiar with the concept of rebates: they have been offered similar incentives before, and cash rebates are an inherent feature of the card's pricing scheme, described in Section 1.2.1.

This month, make your credit card payment to get a cash rebate equal to 50% of your minimum payment on your next statement. Please repay your card balance at your earliest convenience. Call [customer service number].

Indirect Financial Incentives: Credit Reputation

The second benchmarking treatment consisted of a message highlighting the negative effect of non-repayment on a customer's credit reputation and the ability to obtain credit in the future. In this *credit reputation incentive* condition, customers received the standard reminder on the due date and an additional message two days before the repayment deadline. The message stated that non-repayment will result in the customer being reported to the Indonesian credit registry, the *Sistem Informasi Debitur*, which will diminish the customer's access to credit in the future. This message was assigned to 2,000 customers and reads as follows:[26]

Late payments are reported monthly to Bank Indonesia Sistem Informasi Debitur (SID), which all banks consult. This will diminish your ability to get credit in the future. Please repay your card balance at your earliest convenience. Call [customer service number].

---

[26]We designed two variations of this text message and randomly assigned 1,000 customers to each of two subgroups. The first subgroup received the message in the main text. The second group received a text that says "*Late payments are reported monthly to Bank Indonesia Sistem Informasi Debitur (SID), which all banks can consult. Please repay your card balance at your earliest convenience. Call [customer service number].*" We pool these two treatments in our analysis, since their effect on repayment is not statistically different.

Placebo: Simple Reminder

We assigned 1,362 customers to the *simple reminder placebo* treatment condition. Customers in this treatment received the standard reminder on the due date and an additional neutral reminder two days before the repayment deadline.[27] This second reminder is similar to the message sent to all customers who miss the due date and makes no reference to the moral or financial implications of non-repayment:

> *The due date of your [name of the card] bill was on [due date] and your payment has not been received yet. Please repay your credit card balance at your earliest convenience. Call [customer service number].*

This treatment tests how receiving a second reminder affects repayment through channels such as limited attention and memory. Comparing its effect to that of moral incentives allows us to distinguish the impact of moral appeals from the effect of receiving additional reminders.

Placebo: Religious Message

Finally, we assigned 1,000 customers to a *religious placebo* treatment. Customers in this group received the standard message on the due date and an additional message with a quote from the Prophet taken from the same source used in the *moral incentive* treatment two days before the repayment deadline. However, in contrast to the moral incentive message,

---

[27]A number of customers were included in this treatment in the last wave of the experiment to compare the effect of the moral incentive to that of a simple reminder on outcomes measured in a phone survey. The survey asked whether customers would like to receive the same text message again, and how committed they thought the bank is at collecting debt. The survey instrument is available in the Supplemental Appendix.

this quote made no reference to financial matters or debt repayment:

> *The Prophet (Peace and blessings be upon Him) says: "When Allah wishes good for someone, He bestows upon him the understanding of the Book" (Imam al-Bukhari). Please repay your credit card balance at your earliest convenience. Call [customer service number].*

This treatment allows us to test whether moral appeals work because they highlight the moral implications of a specific action, the non-repayment of debts, or simply because they remind recipients of the religious nature of their contract with the bank or evoke a religious frame of mind.

1.2.4 Data and Summary Statistics

The dataset we use in our analysis combines the results from the experiment, administrative data from our partner bank, and information from a number of follow-up surveys phone surveys administered to the bank's customers.

Administrative Data

We first obtained bank data on customer characteristics (age, gender, religion, province of residence, and monthly income) for the universe of late-paying customers participating in the experiment. Table 1.1 reports summary statistics and presents a test of random assignment.[28] The median credit card customer in our sample is male, 41 years old, has a monthly income

---

[28]See Table A.2 in the Supplemental Appendix for summary statistics and a test of random assignment for the follow-up experiment.

of Rp 5,000,000 (US\$ 375), a credit limit of Rp 10,000,000 (US\$ 750) and an outstanding debt of Rp 7,739,015 (US\$ 580) on the credit card.[29] As expected from random assignment, the sample is well balanced across all baseline characteristics.[30]

In a second step, the bank shared data on credit card repayment for customers in our sample after each wave of the experiment as well as historical repayment data covering the 12 months before our intervention. In the monthly repayment data, we observe whether the customer made the required minimum monthly payment by the deadline, which is the main outcome of interest for our analysis. The bank also provided further financial data for the customers in our sample. In particular, we collected data on savings account balances for all customers in the main experiment who also have an account with our partner bank.[31]

Survey Data

We combine data from the experiment with information from a number of phone surveys administered to the population of credit card customers.[32]

The main survey, conducted in June and July 2015, asked respondents about their level of religiosity and their familiarity with the quote used in the three variations of the moral incentive treatment condition. The same survey was also administered to a randomly drawn

---

[29]For comparison, Indonesian per capita income was US\$3,491 (approximately US\$ 291 per month) at the time of the experiment (World Bank, 2014).

[30]Our sample is also very similar to the universe of the bank's credit card customers along most observable dimensions. Late payers are only marginally more likely to be female (40% female versus 37% male) and, on average, have a slightly lower credit limit (Rp 13.5 million versus Rp 14.7 million).

[31]The bank's customers are not required to have a checking or savings account to open a credit card. The most common deposit account within the bank is a liquid savings (*tabungan*) account. In our 2015 sample, 30 customers had a checking account and 1,088 customers have a savings account at the bank at the time of the experiment.

[32]The survey instruments are available in the Supplemental Appendix.

sample of the bank's credit card customers all over Indonesia who were not late in their payments during the study period. We use the results from this survey to construct a measure of local religiosity for the regions in which credit card customers reside. The bank also shared with us the results of an earlier survey, conducted in December 2014 with a random sample of credit card customers not included in our sample. This survey contains broader questions credit and repayment. We use this survey to measure general knowledge about the Indonesian credit registry.

An additional survey was administered one day after the repayment deadline in April 2016 to a random sample of credit card customers who had participated in the experiment that month. The purpose of this survey was to test whether the moral appeal signals that the bank is particularly committed to collected its debt, whether receiving it causes any disutility to customers, and to measure whether the reputational message increases knowledge about the credit reporting system. Respondents in this survey had received either no message, the basic reminder, or one of the versions of the moral message. The survey first asked these customers how committed they thought the bank was to collect debts. Second, it asked whether they wished to receive text messages like the one they had received a few days earlier in the future. Third, customers were randomized in two groups: those in a treatment group were read the content of the reputational incentive message, while those in a control group were not given any information. All customers were then asked questions about the Indonesian credit registry and their beliefs about the consequences of non-repayment.[33]

---

[33]The survey conducted in June and July 2015 was administered to 2,273 participants of our experiment and to other 567 randomly selected customers. The survey conducted in December 2014 was administered to 223 randomly selected customers. The survey conducted in April 2016 was administered to 95 randomly selected participants of the experiment that month, stratified by treatment group.

Main Outcome of Interest

Our main outcome variable is a dummy, indicating whether a customer has made the required minimum payment by the eighteenth day of the month (the repayment deadline). Note that we are limited in our ability to evaluate outcomes measured after this deadline, as we no longer have full experimental control after this date: If a customer fails to make a payment by the deadline, her account is automatically deactivated, she is reported to the credit registry and may receive phone calls from the bank. In particular, the bank may expend greater effort calling customers in treatment groups with lower average repayment, so that the bank's actions might interact with a customer's treatment status. Hence, the impact of our intervention on outcomes other than repayment, observed after the deadline, may not be causal and must be interpreted with caution.

1.2.5 Estimation

Since treatment status was randomly assigned, our identification strategy is straightforward. We identify experimental treatment effects using regressions of the form:

$$Y_i = \alpha + \sum_c \beta_c I_{c,i} + \gamma' \mathbf{X}_i + \epsilon_i, \tag{1.1}$$

where Y_i is an indicator for customer i repaying an amount equal to or greater than the required minimum payment within the deadline. The variables $I_{c,i}$ are indicators for customer i being in category c, where c denotes the experimental treatment condition to which customer i was assigned. In some specifications, we additionally include a vector of control

variables, $\mathbf{X}_i$, which contains either month fixed effects only, or month fixed effects as well as a set of customer and account characteristics. In all regressions, the omitted category is the control group, which received only a basic reminder on the due date but no second text message two days prior to the deadline.

1.3 Results

1.3.1 Main Result: Moral Incentives

We begin by reporting the effect of the moral incentive treatment compared to the control group across all waves of the experiment, shown in Table 1.2, column (1). We do not include any controls, so that these numbers represent raw repayment rates. Compared to the control group, the share of customers making a payment equal to or greater than the required minimum payment by the deadline increases by 13% (from 34% to 38%) under the moral incentive treatment condition.[34] The difference in repayment rates is significant at the 1 percent level (p-value=0.000). In column (2), we add month fixed effects, and in column (3) we add customer-level covariates. The results remain very similar across all specifications (with treatment effects ranging from 13% to 15% above the baseline effect of 34% in the control group), indicating that the randomization was successful.

The treatment effect is similar for men and women, and also does not differ by age, religion, or whether a customer has appeared on the list of late payers at least once in the

---

[34]DellaVigna and Gentzkow (2010) report persuasion rates from a variety of settings. Using the same calculation as in their paper, the implied persuasion rate in our experiment is approximately 6%.

year before our intervention.[35] The treatment effect is stronger for customers with a lower debt-to-income ratio.[36]

1.3.2 Benchmarking the Moral Incentive Effect

Moral versus Financial Incentives

To assess the economic significance of the moral incentive effect, we use two benchmarking treatments. In the first benchmarking treatment, the bank sent text messages to a random subset of customers offering them a substantial cash rebate in the next billing cycle if they made the required minimum payment in the current month by the deadline. This cash rebate amounted to 50% of the minimum payment to be made, which is equivalent to a 5% reduction of the customer's total outstanding credit card debt. The median rebate offered was Rp 380,000 (US\$28), which is equal to 8% of monthly earnings for the median customer in our sample (the average rebate offered was Rp 500,000).

The results are reported in Table 1.3. In column (1), we present raw repayment rates, restricting our sample to the fourth wave of the intervention to keep the time period constant across treatments. In column (3), we add month fixed effects and include observations from all months of the intervention. In column (4), we also include individual controls. Across all specifications, we find that providing financial incentives increases repayment

---

[35]Fewer than 10% of customers in our sample are non-Muslim, so that it is not possible to estimate this effect precisely. We discuss heterogeneity by *local religiosity* in Section 1.4.1. Heterogeneous treatment effects are reported in Table A.3 in the Supplemental Appendix.

[36]However, this pattern also holds for the reputational incentive treatment, and is therefore suggestive of financial constraints, rather than a reaction to the conditional statement "non-repayment of debts *by someone who is able to repay* is an injustice" in the moral message.

rates. Although the magnitude of the effect of financial incentives is lower than that of the moral message, we cannot rule out that the two effects are the same under conventional significance levels due to the limited sample size. However, the p-value of the one-sided test that that the cash rebate treatment coefficient is larger than the moral incentives treatment coefficient is 0.104 in the specification with fixed effects, and 0.055 in the specification that includes fixed effects and controls. These results suggest that providing moral incentives can be more powerful than providing substantial one-time financial incentives, especially in terms of cost-effectiveness: the average rebate offered to clients who responded to the message by making a payment was Rp 580,000 (US$43). By contrast, sending text messages with a moral appeal comes at practically no cost to the bank.

Moral versus Reputational Incentives

In addition to the direct financial consequences of non-repayment, customers might also care about material incentives with a longer time-horizon, such as their reputation in the credit market. To evaluate this hypothesis, one group of customers was sent a text message informing them about the Indonesian credit registry, the *Sistem Informasi Debitur*, and the consequences of being reported for non-repayment instead of the moral message. Specifically, the credit reputation message informed customers that all banks consult the credit registry before issuing credit, so that non-repayment of credit card debt has adverse consequences on future access to credit.

Evidence from a survey with 223 clients drawn from the universe of bank customers suggests that overall knowledge about the Indonesian credit registry is limited. About 75%

of respondents report that they do not know about the credit registry, and most clients demonstrate to have substantial misconceptions about the consequences of a bad credit record. For example, 34% of respondents think it will make them unable to open a deposit account, 48% think they will have to appear in front of a judge –neither of which are true, and 22% of respondents think it will have no consequences on their ability to obtain credit in the future, which is also false, since all banks in Indonesia use the credit registry to screen customers.[37]

Results from the credit reputation treatment are also reported in Table 1.3. When looking at raw repayment rates in Table 1.3, column (2), informing customers about the credit registry raises the probability of meeting the minimum repayment by the deadline by 29% (as opposed to 18% for moral incentive messages sent during the same months). To gain a better understanding of how the reputational incentive treatment affects customers' decisions, a small survey was conducted in April 2016. Customers were randomized in two groups: customers in a treatment group were read the content of the reputational incentive message, and customers in a control group were not given any information. All participants were then asked some questions about the Indonesian credit registry. The results from the survey suggest that late paying customers are poorly informed about the functioning of the credit registry, and that the reputational incentive message does not increase their knowledge of how the registry functions. Instead, the message seems to simply make customers think

---

[37]The survey referred to the *Sistem Informasi Debitur* and not to the credit registry in general, so that some clients might be aware about the existence of a credit registry, but not its actual name. The relatively large effect of the credit reputation treatment might in part be due to the message signaling to customers that the bank is serious about actually reporting delinquent customers to the registry.

that the consequences of being reported to the credit registry are more severe.[38] Taken together, the benchmarking results indicate that both moral and reputational incentives affect repayment decisions in our setting.

1.3.3   Ruling out Other Channels

Our results establish that receiving a moral message substantially increases the repayment rate. However, there are several mechanisms other than moral suasion that could explain this effect. In this section, we present a number of tests to evaluate alternative channels and show which of these potential explanations can be ruled out.

Reminding Customers

First, receiving a text message might increase repayment rates simply because it acts as a reminder, irrespective of whether the message contains a moral appeal or not (see, for example, Karlan et al., 2015). To address this possibility, we compare repayment in the moral incentive treatment group to repayment among customers assigned to the simple reminder placebo treatment, which consisted of a basic non-religious reminder that made no reference to morality or religion and was sent at the same time as the moral message. The results, displayed in Figure 1.2 and reported in Table 1.4, show that receiving the simple reminder has no effect on repayment. The raw repayment rate is 35% in the group receiving the basic

---

[38]Compared to the control group in the survey, the share of customers correctly stating that being reported will have consequences for their ability to get credit from other banks increases from 38% to 49% (p-value=0.18). However, exposure to the text from the reputational treatment also increases the share *incorrectly* reporting that: (i) they will have problems opening a deposit account (from 27% to 49%, p-value=0.007); (ii) they will have problems getting credit from the same bank (from 38% to 50%, p-value=0.18); (iii) they will have to appear in front of a judge (from 19% to 25%, p-value=0.421).

reminder, compared to 34% in the control group. The p-value of the difference between the simple reminder and the control is 0.714, and the p-value of the difference between the simple reminder and the moral message is 0.013. We can therefore rule out that the moral message works simply because it reminds customers to repay their debt.

Priming Religion

Second, receiving a text message with religious content could affect the repayment decision through priming effects, which are also unrelated to moral suasion. The moral message might, for example, remind recipients of the religious connotation of the credit contract or evoke a religious frame of mind more generally. To rule out this possibility, we compare repayment in the moral incentive treatment group to repayment among customers who received the religious placebo message. The religious placebo message contains a quote from the Prophet that is taken from the same religious text as the quote used in the moral message but makes no reference to debt repayment. The results, also displayed in Figure 1.2 and reported in Table 1.4, show that the religious placebo message has no effect on the repayment rate. The raw repayment rate is 35% in the group receiving the religious reminder and nearly identical to the repayment rate in the control group. The p-value of the difference between the religious placebo and the control is 0.889, and the p-value of the difference between the religious placebo and the moral message is 0.007, indicating that the effect of the moral message is also not driven by priming on religion.

Novelty of the Message

Third, customers may respond to the message not because of its moral content, but because it is novel or attention-grabbing. To test for this possibility, we consider repayment rates under different text message treatments that use new content. Note that several of the messages that were sent to credit card customers as part of the experiment –including the simple reminder, religious placebo, and financial reminder messages– were specifically designed for the study, and had never been received by any of the bank's customers before. The fact that none of these messages had a statistically significant effect on repayment allows us to rule out that the effect of the moral message is explained by the novelty of the message. We can also rule out the possibility that receiving a message with a quote from the Prophet is particularly attention-grabbing. The religious placebo message also uses a quote from the Prophet, which is very similar to that in the moral incentive treatment and taken from the same religious text. However, as we show above, this message has no effect on repayment.

Signaling the Bank's Commitment to Debt Collection

Finally, since customers had previously received a text message at the time of the due date, receiving a second message could be perceived as a signal that the bank is particularly committed to debt collection, which could affect repayment rates independent of the moral appeal. To address this possibility, we conduct the following test. In April 2016, the bank the sent placebo messages and the three variations of the moral message discussed above to customers never treated before. Another group of customers was randomly assigned to a control group and received no message. We conducted a phone survey with customers in

both groups the day after the payment deadline and asked "How committed do you think [bank name] is to collect debt from delinquent customers on a scale from 1 to 5 (where 1 is not very committed, and 5 is very committed)?" The percentage of respondents that answered 4 or 5 is 76% in the control group, 67% in the basic reminder group, and only 59% among customers that received a moral message (the p-value for the test of equality of all three coefficients is 0.30, and the p-value of the test of equality between respondents in the control group and the treatment group is 0.12). Hence, there is no evidence to suggest that receiving the moral message is perceived as a signal that the bank is now more committed to enforce debts.[39, 40]

1.4 Interpreting the Results

1.4.1 What Drives the Moral Appeal

The evidence from the previous section rules out several mechanisms that are unrelated to moral suasion but could generate higher repayment rates in response to the moral message. We now explore competing hypotheses for the effectiveness of the moral appeal and present tests to distinguish between these alternative explanations.

---

[39]We also obtained the repayment history of all clients in our sample from the partner bank, and use this information to test whether the response to the moral incentive treatment differs, depending on whether a customer appears on the list of late payers for the first time or has been delinquent before. We find no evidence that this is the case.

[40]These findings also relate to those in Hallsworth et al. (2015), where a reframing of debt non-repayment from an error of "omission" to an error of "commission" increased of repayment of tax debt. The authors in that paper also find that the act of commission is associated with greater beliefs about punishment for non-repayment, and this change in beliefs about punishment is their preferred interpretation for their results. In our setting, since there are no changes in beliefs about punishment from the bank, the findings indicate an association of commission with greater moral costs. This suggests that the mechanism of moral penalties might also be at play, in addition to the main channel proposed in that paper.

Religious Connotation of the Message?

The first possibility is that individuals indeed respond to the moral content of the message, but that this effect arises only because the moral appeal is delivered in a religious context. This seems plausible, given that the original moral incentive message explicitly quoted the Prophet and cited the religious text from which the quote was taken. Moreover, the original moral incentive message used a word for "injustice" that is of Arabic origin, and is often used in a religious context.

In order to distinguish the religious context of the moral message from the effect of the moral appeal, the bank sent two additional variations of the moral message to a randomly chosen subset of credit card customers. The first message was identical to the main treatment, but omitted the name of the Prophet and the source of the quote. The second variation of the moral message omitted the name of the Prophet as well as the source of the quote and additionally replace the Arabic-origin word for "injustice" with the standard Indonesian word, which has no religious connotation. Hence, the first message tests, whether adding a credible religious source adds power to the impact of a moral appeal. The second message tests if simply receiving a simple moral appeal without any religious connotation affects repayment decisions.

The results are displayed in Figure 1.3, and reported in regression format in Table 1.5. In the months in which the three variations of the moral message were sent raw repayment rates are similar for all three variations of the moral incentive condition. This could indicate that either customers already associated the moral appeal contained in the message with religion (and potentially with the Prophet), or that the pure moral statement was indeed

sufficient to trigger repayment. To disentangle these competing hypotheses, we conducted a follow-up phone survey with a random sample of credit card customers. In this phone survey, the message with the standard Indonesian word for "injustice" and without reference to the Prophet was read to customers, who were then asked to indicate its source.[41] The vast majority of clients were not immediately aware of the religious origin of the message. When asked "Who do you think might have said this phrase?", out of 5 given options, 77% chose "I don't know", whereas only 19% associated the phrase with religious figures or institutions (including the bank itself). These findings suggest that the higher repayment rate was not due to an implicit religious association with the message. These results also corroborate the view that our sample is relatively secular; most clients did not recognize the Islamic doctrine on non-repayment of debts.

The follow-up survey helps us further clarify the role of religiosity in explaining the effects. In the survey, respondents were asked about the importance of religion and the rules of Islamic law in their life, using a 1-5 Likert scale. The survey also asked customers to rank the relative importance of family, work, friends and religion. Because of the small sample size of the survey, we cannot directly use this measure to assess the individual-level heterogeneity of treatment effects.[42] Instead, we use it to construct province-level indicators of religiosity. To do so, we split the sample according to the share of respondents who identified as very religious in each province and compare treatment effects for customers in locations classified

---

[41]None of the customers in this sample had previously received any of the moral incentive text messages.

[42]This survey was administered to 2,840 customers. Among them 2,273 participants of our experiment and 567 randomly selected customers of the bank that did not participate in the experiment.

as more or less religious according to this measure.[43] In the less religious half of provinces, the moral message (all versions) increased repayment rates by 3.8 percentage points, or by 11% (p-value=0.003). In the more religious half of provinces, the effect of the moral message was significantly larger, with an additional 4.7 percentage point increase in the likelihood of repayment. The p-value of the interaction bentween the moral message and a dummy for local religiosity above the median is 0.047.[44] Interestingly, these patterns are similar for the religious, implicit, and non-religious version of the moral message. This suggests that customers in more religious regions may be more responsive to moral appeals in general, but supports the conclusion that the effect of the moral appeal is not driven by religion.

Provision of New Information? The Impact of Repeated Messages

We also explore whether the moral message works only when it is sent for the first time – for example, because it conveys new information – or if it works if it is sent to customers who have received the message before. To address this question, we conducted a follow-up experiment with a sample of customers who had already received the moral message once and reappeared on the list of late payers. In February and April 2016, customers in this group were either sent the same version of the moral message that they had previously received for

---

[43]Customers are identified as very religious if they answered "Extremely Important" to both the question about religion and the question about the rules of *Shari'a* law, and if they ranked religion as the most important thing in their life among all the choices given.

[44]We find the same result using data from the *Indonesian Family Life Survey* (IFLS). Using responses to the question "How religious are you?" we built a similar measure of local religiosity and ranked provinces with respect to the proportion of respondents who describe themselves as "very religious". The results are quite similar: the moral message leads to a 3 percentage point increase in repayment (a 9% increase compared to the control group) in the less religious half of provinces (p-value=0.104), and and additional 3 percentage point increase in the more religious half of provinces (an 18% increase compared to the control group). The p-value of the interaction between the moral incentive treatment and a dummy for local religiosity above the median is 0.131.

a second time, with a lag of either two months or approximately one year, or were assigned to a control group that received no additional message.

Table 1.6 reports the results, pooling across different versions of the moral incentive messages. We find suggestive evidence that repeated moral messages still affect repayment, and that the size of the effect is not lower among customers to whom the moral message is sent for a second time. In the specification without individual covariates and month fixed effects, reported in Table 1.6, column (1), the effect of the repeated moral message is 0.041 (p-value 0.175).[45]

We next compare the effects of the first and the second moral messages. In order to do so, we pool the sample from the repeated message experiment with the data from the main experiment. This requires some caution, since there are likely to be selection issues. In particular, customers who show up on the list of late payers for a second time are likely to be different from those who appear on the list for the first time. Indeed, we find that while the two samples are well balanced on demographics, customers in the repeated message sample have lower income and credit limits, and are more likely to have been more than thirty day past due at least once in the previous year.[46] For this reason, it is important to include individual covariates to address this potential selection problem.[47] The results are reported

---

[45]There is suggestive evidence that the effects do not vary depending on the time lag between the first and the repeated message. Sending the moral message to customers who already received the same message one year before increases repayment by 0.040 compared to sending no message (p-value 0.323). Sending a moral message to customers who received the same message two months before increases repayment by 0.044 compared to sending no message (p-value 0.346). However, the sample sizes are too small to estimate effects separately by time since the first message.

[46]See Appendix Table A.5 for details.

[47]Another possible concern is the presence of differential selection due to the treatment. However, we do not find any evidence of this type of selection: the proportion of customers showing up on the list of late payers a second time in 2016 after having appeared in the sample of our main experiment in 2015 is

in Table 1.6, column (4). The point estimate of receiving the moral message for the first time is 0.045. With a point estimate of 0.043, the effect of receiving the moral message for a second time is nearly identical, and both effects are statistically significant. The p-value of a test of equality of the two effects is 0.955.[48]

The result that the moral message affects repayment even when it is sent repeatedly rules out the possibility that the message affects repayment by conveying new information. The finding that even a moral message with no reference to religion affects repayments already indicates that the effect is not driven by the recipient learning about a religious teaching that they were not previously aware of. Similarly, the effect cannot be explained by the customer learning that non-repayment of debts can be considered immoral. In both cases, the message would affect repayment only when this information is conveyed to a delinquent customer for the first time. Though we cannot directly test for it, our results are consistent with an inattention interpretation, in which the moral incentive message temporarily draws customers' attention to the moral dimension of the repayment decision.

.251 among those receiving a moral message and .242 among controls (p-value of the difference .345). So, the fact of being late again after a few months from our intervention is likely due to some negative shock independent of treatment status (possibly a negative income or liquidity shock, which is in line with these customers having lower income, credit limit and being more likely to have been more than 30 days past due in the past).

[48]Note that here we are attempting to compare the size of the effect of a message sent to customers who have never seen it before and are late for a first time, to the size of the effect in the (selected) sample of customers who have seen the message before and are late a second time. While both estimates can be interpreted causally, we cannot causally evaluate the effect of repeated messages on the non-selected sample since no further messages are sent to customers who are not late a second time.

1.4.2 Disutility from Receiving the Message

The results indicate that moral incentives are effective at getting customers to repay their credit card debt. However, so far, it is unclear if this comes at a utility cost to customers. To answer this question, in the last month of the intervention the bank called back customers who had received either the simple reminder or one of the versions of the moral incentive message as part of a follow-up survey one business day after messages had been sent out (messages were sent on Friday and surveys were conducted the next Monday). These customers were asked the question "[Bank name] is sending reminder messages to its customers to help them make their payment on time. You received one of this messages last week. Would you like to receive the same message again in the future?" The percentage of customers who reported that they would like to receive the message again was 80% among those who had just received a placebo reminder, and also 80% among those who had received one of the variations of the moral message. The fact that a large majority of customers would like to receive a similar message again suggests that receiving a moral appeal does not create disutility to the recipients. Moreover, receiving a message containing a moral appeal does not seem to create differential disutility, compared to a simple reminder.

As an additional test, we also examine whether sending a moral appeal could negatively affect the bank by reducing card usage or transaction volumes (perhaps because customers are dissatisfied with the bank after receiving the message, or want to avoid receiving a similar moral appeal in the future). We find that this is not the case. In the 30-day window after the intervention, the average amount spent is Rp 1,217,169 for customers that received the moral message, and Rp 1,260,626 for customers in the control group (p-value 0.699).

The probability of card usage during this time period is .448 and .441 respectively (p-value 0.691).[49]

1.4.3 Additional Results and Extensions

Impact in Later Months

We next examine the persistence of the moral message effect in the sample of customers that received the moral message only once, restricting ourselves to the waves of the experiment conducted in 2015.

It is first worth noting that a sizable share of customers who are late in making repayments in a given month during our sample period appear again in the list of late-paying clients a month later. Among clients in the control group, the average probability of showing up in the list the following month is 0.31. There is also some evidence of income effects: individuals in the control group who make a payment are 7 percentage points more likely to appear again on the list of delinquent customers, than individuals who do not make a payment.

One obstacle we face when trying to examine the persistence of the moral incentive effects is the fact that we no longer have full experimental control after the repayment deadline. Once the deadline has passed, customers that have not made a payment are reported to the credit registry and the bank's collection team attempts to call delinquent clients. It is possible that customers react differently to a given phone call if they have previously received a treatment. Moreover, the bank itself can exert differential effort in calling different clients from different treatment arms. For instance, the bank might be more likely to call clients

---

[49]These results also hold for different time windows after receiving the message.

in the control group, or treatment groups with lower average repayment. Since we have no information on follow-up calls and effort by the bank after the deadline, any outcomes observed after the repayment deadline may not be causal and should be interpreted with caution.

With this caveat in mind, we can still attempt to assess whether the moral messages sent as part of our experiment had a persistent effect. Note that by raising the repayment rate at a given point in time, the moral incentive message may generate two counteracting effects on repayment in later months: First, it might be that the moral message generates greater incentives to make a payment in the month when it is received and in the following month. That is, the moral incentives themselves might be persistent. Second, an extra incentive to repay right away may also generate an income effect when compared to the control.[50] That is, the moral message will induce more clients to make a payment in the month the message is received, so that some of them may be less able to repay one month later. The impact of the treatment we observe in the months after the treatment is the combination of these two effects.

We observe that the likelihood of appearing again in the late paying list one month later is 1 percentage point higher for clients who received the moral incentive message the previous month (the difference is not statistically significant). Although we cannot isolate persistent effects of the moral message from income effects, we can still try to infer the size of these effects and assess the likelihood of finding persistent effects of the moral message

---

[50]We can rule out the possibility that the absence of persistent effects is driven by the fee structure of the card. Late fees are increasing in the number of days past due, but there are no escalating penalties for being late repeatedly.

in the absence of income effects. To approximate the size of income effects, we multiply the increase in the probability of repayment due to the moral incentive treatment by the increase in the probability of being late the next month after repaying in the control group. Abstracting from selection issues, we find that income effects can account for up to a 0.4 percentage point increase in the probability of late payment the following month in the moral incentive group, compared to the control. Subtracting this number from the higher likelihood of showing up in the list of late-payers the following month for the moral incentive group yields a persistent effect that is close to zero. Although we cannot make sharp predictions, the evidence suggests that the moral incentive effects did not persist until the following month. This is consistent with the interpretation that the moral appeals sent as part of our experiment affect repayment by highlighting the moral component of the repayment decision, thus temporarily drawing customers' attention to moral considerations.

Impact on Savings Account Balances

To better understand how customers make payments in response to the experimental treatments, we next examine the effect of repayment on savings account balances. For this purpose, we obtained detailed data on savings account balances for participants of our experiment from our partner bank. We have access to customers' daily balances on their *tabungan* (Indonesian for "savings") accounts. These are the most common type of deposit account among clients of our partner bank, and have all characteristics of a standard liquid savings account. Since credit credit card customers are not required to also have another account with the bank, savings account balances are available for only 13% of customers in

our sample, which may give rise to selection issues.

We find that meeting the minimum payment increases the likelihood of a reduction in customers' savings account balances, suggesting that customers are using their savings account balances to repay more expensive credit card debt. More specifically, among those who met their minimum repayments in response to receiving one of our messages, 22% reduced their savings account balance between the sixteenth and eighteenth day of the month. Among those who did not repay, only 8% had a reduction of their savings account balance over the same time period. The difference is significant at the 1 percent level (p-value=0.000). However, we do not have sufficient statistical power to detect differences in savings balances across the treatment arms of our intervention, so that we consider this evidence as merely suggestive.

Impact on the Intensive Margin of Repayment

We can further unpack the mechanism through which moral incentives affect behavior by examining the intensive margin of repayment, that is, the amount repaid conditional on meeting the minimum payment. Since each of our treatments may induce a different subset of consumers to repay, looking at the intensive margin of repayment in isolation induces selection problems. In fact, because customers with a lower average willingness to repay might make a payment if they were included in one of the moral incentive treatment groups, a simple comparison between treatment and control groups would most likely understate the intensive margin effect.

To avoid this selection problem, we impute zeros for all customers who did not make a

payment and analyze the combined effect of our treatments on the intensive and extensive margin. These are unconditional means, and therefore *not* subject to selection issues. The average amount repaid in the control group is Rp 637,819, and expected repayment in the moral incentive group is slightly higher than in the reputational incentive group at Rp 745,352 versus Rp 713,437 (p-value=0.65).[51] The share of customers that repay substantially more than, i.e. more than twice, the amount required to avoid being reported to the credit registry is significantly higher in the moral incentive group, compared to the reputational message group (23% versus 19%, p-value=0.08). This result suggests that customers in the reputational incentive treatment act much more strategically in response to the message, and are more likely to repay only the required 10% of their outstanding balance. In contrast, customers receiving the moral message are more likely to repay an amount significantly higher than the minimum payment, required not to be reported to the credit registry.

Therefore, while the effect of the reputational message on the extensive margin is slightly larger than that of the moral incentive, their effect on the intensive margin of repayment differs. There are two possible channels at play: moral hazard and adverse selection. With moral hazard, ex-ante identical individuals will respond differently to each message. For example, after receiving a message stating that the bank reports all customers who fail to meet the minimum payment to the credit registry, an individual may exert effort to meet the minimum payment (but will not make a payment exceeding this amount). On the other hand, that same individual may decide to repay even more than the required minimum

---

[51]For these comparisons, we restrict the sample to customers late in February, March and May 2015 since there are the only months when the reputational incentive message was sent. If we consider the whole sample, the average amount repaid in the control group is Rp 615,835 and in the moral incentive group is Rp 725,169.

amount after being reminded of the injustice of failing to repay her debt. In the presence of adverse selection, individuals who respond to each message are different ex-ante. For example, customers who respond to the threat of being reported to a credit registry may be more strategic to begin with.[52] Since meeting the minimum repayment is voluntary, we cannot disentangle moral hazard from adverse selection in our setting. Note, however, that while moral hazard and adverse selection have different welfare implications, they have the same implication for the effectiveness of moral and material incentives in our setting: while material incentives are effective in inducing people to meet the minimum payment, few people pay more than the minimum. In contrast, moral incentives induce slightly fewer people to meet the minimum repayment. However, more of those who repay exceed the minimum required amount.

1.5 Conclusion

While moral considerations may influence many important economic decisions, economists have typically focused on monetary incentives as the main determinant of behavior.

In this paper, we provide evidence that non-pecuniary moral incentives can strongly affect a financially important and recurrent economic choice: the decision to repay one's debts. In our setting, moral appeals are more effective than substantial monetary incentives as a means to encourage debt repayment. We find that the impact of our intervention on behavior is driven by responses to the moral appeal and use a number of placebo treatments to rule out

---

[52]See, for example, Einav et al. (2013) for evidence of such "selection on moral hazard" in health insurance.

competing explanations, such as reminder effects, novelty of the message, priming religion, signaling the lender's commitment to debt collection, and the provision of new information.

In our setting – an Islamic credit card –, moral appeals similar to the ones we study are natural and common. This allows us to design a real-stakes experiment in which we exogenously vary the religious content of the moral appeal, so that it is possible to experimentally isolate the response to the moral appeal from the response to its religious context. We find that changing the religious content of the appeal has no effect on repayment. In particular, non-religious messages similar to messages used by commercial banks in other settings also encourage debt repayment.

These findings are consistent with the idea that, even in the case of financial decisions, people experience a cost from consciously violating a moral norm. A moral appeal can therefore affect behavior, even when it is not associated with an explicit threat of punishment or negative financial consequences. The relative importance of monetary and non-monetary considerations in such decisions is of course context-dependent. Studying how moral incentives operate in other settings is therefore an important avenue for future research.

1.6 Figures and Tables

Figure 1.1: **Timeline of Events**
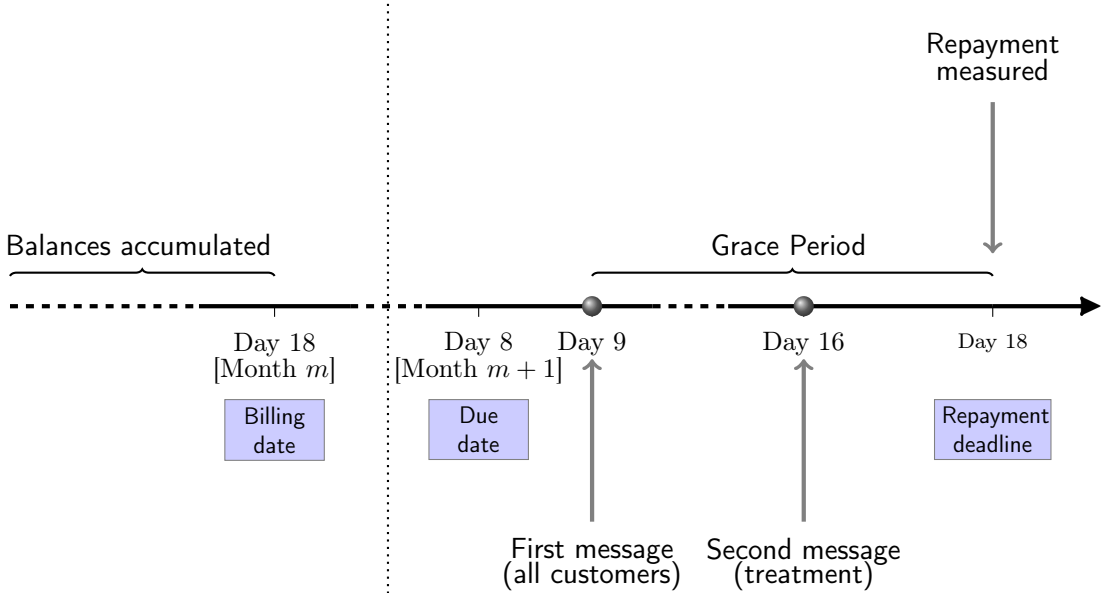
Notes: The figure shows the credit card billing cycle and timing of the intervention. Customers receive their monthly statement on the *eigthteenth* day of each month. The due date is on the *eight* day of the following month. One day later, the bank sends a simple reminder message to *all* late-paying customers. The repayment deadline is on the *eighteenth* day of the month, at the end of a 10-day grace period. In the morning of the *sixteenth* day of the month (two days before the repayment deadline), randomly assigned reminder messages are sent to customers assigned to one of the treatment groups. Repayment is observed at the time of the final deadline, which is midnight of the *eighteenth* day of the month.

Figure 1.2: **Treatment Effects**
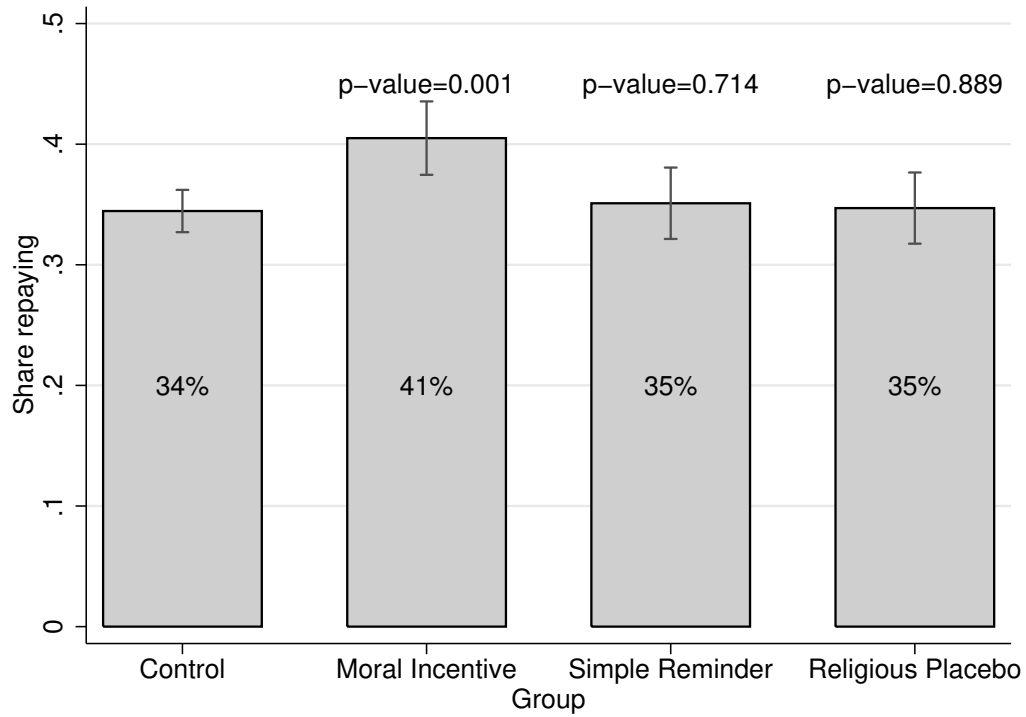
Notes: This figure presents the means and 95 percent confidence intervals of the raw repayment rates for the sample of customers assigned to one of the four following groups: control, moral incentives, simple reminder, and religious placebo (these two treatments have not been run simulateously in Waves IV, V and VI, so customers late in June 2015, and February and April 2016 are excluded from the sample analyzed in this figure). There are 1000 observations in each of the treatment groups, and 2821 customers in the control group. For each treatment we report the *p*-value of a test of equality of the means in the treatment and in the control.

Figure 1.3: **What Drives the Moral Appeal? Religious Connotation**
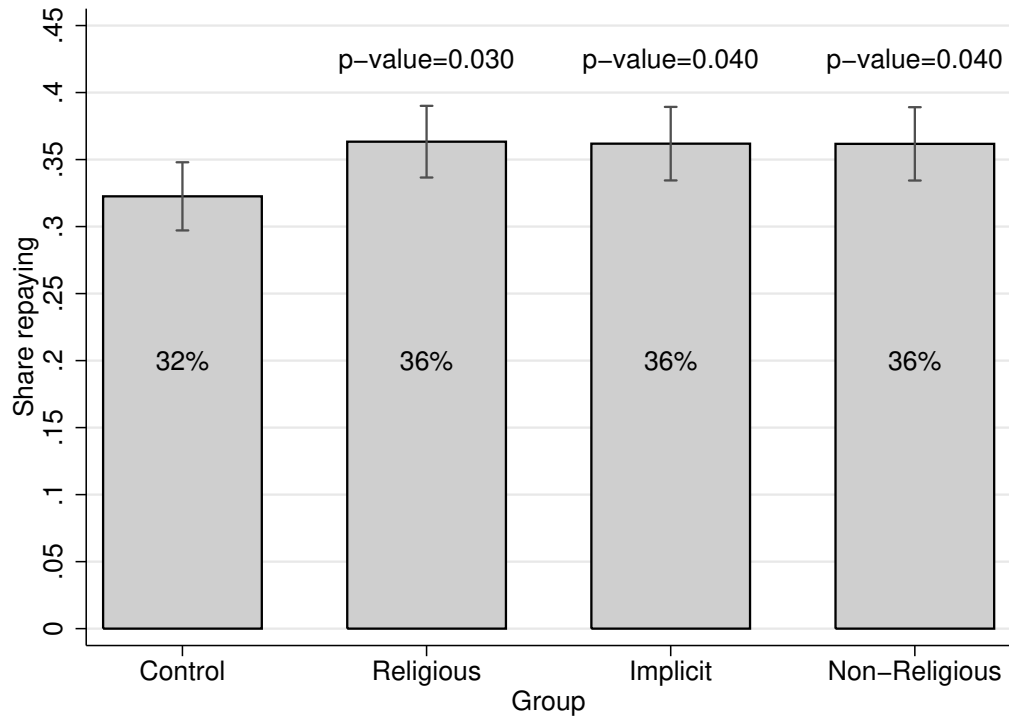
Notes: This figure presents the means and 95 percent confidence intervals of the raw repayment rates for the sample of customers assigned to one of the four following groups: control, religious moral incentives, implicit moral incentives, and non-religious moral incentives (these last two treatments have been run only in Waves IV, V, and VI, so customers late in February, March, and May 2015 are excluded from the sample analyzed in this figure). There are respectively 1244, 1186, and 1180 observations in the religious moral incentives, implicit moral incentives, and non-religious moral incentive treatment groups, and 1299 customers in the control group. For each treatment we report the *p*-value of a test of equality of the means in the treatment and in the control.

Table 1.1: **Balance and Treatment Cell Size**

	Full Sample	Moral Incentive	Simple Reminder	Religious Placebo	Credit Reputation	Control Group	p-value
Panel A1: Waves I, II, and III Balance of Covariates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	42.03	42.36	42.10	41.73	41.99	42.03	0.631
	[9.071]	[9.317]	[8.776]	[8.717]	[9.092]	[9.195]	
Female	0.40	0.40	0.41	0.41	0.39	0.40	0.914
	[0.489]	[0.490]	[0.491]	[0.491]	[0.488]	[0.489]	
Muslim	0.92	0.92	0.91	0.91	0.92	0.92	0.427
	[0.273]	[0.271]	[0.286]	[0.289]	[0.271]	[0.264]	
Annual Income	151.67	135.51	185.73	134.86	177.65	132.85	0.418
(Rp, million)	[836.968]	[175.295]	[1242.218]	[187.644]	[1369.992]	[201.640]	
Credit Limit	13.55	13.93	13.28	13.77	13.38	13.55	0.438
(Rp, million)	[9.338]	[9.708]	[8.652]	[9.444]	[9.272]	[9.448]	
Panel A2: Waves I, II, and III Treatment Cell Size							
Wave I	2871	400	400	400	800	871	
Wave II	2985	400	400	400	800	985	
Wave III	1965	200	200	200	400	965	
Total	7821	1000	1000	1000	2000	2821	

	Full Sample	Moral Incentive [Religious]	Moral Incentive [Implicit]	Moral Incentive [Non-Religious]	Cash Rebate	Control Group	p-value
Panel B1: Wave IV Balance of Covariates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	42.24	41.82	42.70	41.98	42.31	42.38	0.764
	[9.491]	[9.170]	[9.415]	[9.137]	[9.196]	[10.477]	
Female	0.39	0.42	0.38	0.38	0.37	0.40	0.703
	[0.488]	[0.494]	[0.486]	[0.487]	[0.482]	[0.490]	
Muslim	0.92	0.93	0.91	0.90	0.93	0.92	0.517
	[0.271]	[0.253]	[0.281]	[0.302]	[0.253]	[0.265]	
Annual Income	134.64	121.99	132.46	138.35	152.25	128.27	0.345
(Rp, million)	[189.589]	[192.350]	[154.065]	[187.183]	[233.037]	[172.253]	
Credit Limit	13.56	13.15	13.13	14.20	13.87	13.44	0.569
(Rp, million)	[9.834]	[10.587]	[9.360]	[9.525]	[9.867]	[9.803]	
Panel B2: Wave IV Treatment Cell Size							
Wave IV	1687	336	336	336	336	343	
Total	1687	336	336	336	336	343	

	Full Sample	Moral Incentive [Religious]	Moral Incentive [Implicit]	Moral Incentive [Non-Religious]	Simple Reminder	Control Group	p-value
Panel C1: Wave V and VI Balance of Covariates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	41.61	41.73	41.80	41.36	40.95	41.79	0.557
	[9.722]	[10.093]	[9.481]	[9.639]	[9.954]	[9.562]	
Female	0.39	0.41	0.37	0.40	0.43	0.36	0.087
	[0.488]	[0.492]	[0.483]	[0.489]	[0.496]	[0.481]	
Muslim	0.90	0.88	0.90	0.92	0.88	0.89	0.087
	[0.306]	[0.326]	[0.295]	[0.279]	[0.321]	[0.314]	
Annual Income	158.51	141.85	205.76	159.15	160.08	131.45	0.379
(Rp, million)	[966.064]	[556.385]	[219.339]	[1942.643]	[609.755]	[184.891]	
Credit Limit	13.87	13.68	14.00	13.73	13.59	14.17	0.786
(Rp, million)	[10.257]	[10.143]	[10.037]	[10.530]	[9.967]	[10.441]	
Panel C2: Wave V and VI Treatment Cell Size							
Wave V	2106	546	482	488	0	590	
Wave VI	1814	362	362	362	362	366	
Total	3920	908	850	844	362	956	

Notes: Panel A1 reports summary statistics for the sample and presents a test of random assignment for waves I, II, and III. Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample. Columns (2) to (6) report the mean level of each variable, with standard deviations in brackets, for all the experimental conditions. Column (7) reports the p-value of a test that means are the same in all the experimental conditions. Panel A2 reports treatment cell sizes by month. Panels B1 and B2 replicate this for wave IV. Panels C1 and C2 replicate for waves V and VI.

Table 1.2: **Moral Incentive Effects**

Dependent variable	Dummy: customer repaid within the deadline		
	(1)	(2)	(3)
Moral Incentive	0.044***	0.052***	0.051***
	[0.013]	[0.013]	[0.013]
Mean Repayment Control Group		0.34	
Month fixed effects	No	Yes	Yes
Controls	No	No	Yes
Waves	All Waves	Full Sample	Full Sample
N	6364	13428	13428
R^2	0.002	0.011	0.057

Notes: Column (1) restricts the sample to customers assigned to the moral incentive treatment or to the control group. Column (2) and (3) use the whole sample. Columns (1) presents OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on treatment group dummies. The control is the omitted group, for which we report the mean repayment rate. Column (2) replicates and adds month fixed effects. Column (3) replicates and adds individual covariates (age, gender dummy, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1.3: **Benchmarking Moral Incentives: Cash Rebate and Credit Reputation**

Dependent variable	Dummy: customer repaid within the deadline			
	(1)	(2)	(3)	(4)
Moral Incentive	0.054	0.060***	0.052***	0.051***
	[0.036]	[0.018]	[0.013]	[0.013]
Cash Rebate	0.021		0.014	0.003
	[0.035]		[0.030]	[0.029]
Credit Reputation		0.098***	0.102***	0.104***
		[0.014]	[0.014]	[0.013]
Moral Incentive -	0.033		0.038	0.047
Cash Rebate	[0.036]		[0.030]	[0.029]
	(0.185)		(0.104)	(0.055)
Moral Incentive -		-0.038**	-0.051**	-0.053***
Credit Reputation		[0.019]	[0.016]	[0.016]
Mean Repayment Control Group	0.30		0.34	
Month fixed effects	No	No	Yes	Yes
Controls	No	No	No	Yes
Waves	Only Wave IV	Waves I, II, and III	Full Sample	Full Sample
N	1015	5821	13428	13428
R^2	0.002	0.008	0.011	0.057

Notes: Column (1) restricts the sample to customers late in June 2015 and assigned to one of the three following groups: moral incentives, financial incentives (this treatment has been run only in Wave IV) and control. Column (2) restricts the sample to customers late in February, March or May 2015 and assigned to one of the three following groups: moral incentives, reputational incentives (this treatment has not been run in waves IV, V and VI) and control. Column (3) and (4) use the whole sample. Column (1) and (2) present OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on treatment group dummies. The control is the omitted group, for which we report the mean repayment rate. Column (3) replicates and adds month fixed effects. Column (4) replicates and adds individual covariates (age, gender dummy, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). "Moral Incentive - Cash Rebate" gives the difference between the coefficient on "Moral Incentive" and the coefficient on "Cash Rebate." P-value for the test of inequality "Moral Incentive < Cash Rebate" in parenthesis. "Moral Incentive - Credit Reputation" gives the difference between the coefficient on "Moral Incentive" and the coefficient on "Credit Reputation." Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1.4: **Ruling Out Other Channels**

Dependent variable	Dummy: customer repaid within the deadline		
	(1)	(2)	(3)
Moral Incentive	0.060***	0.052***	0.051***
	[0.018]	[0.013]	[0.013]
Simple Reminder	0.006	0.023	0.022
	[0.018]	[0.015]	[0.015]
Religious Placebo	0.002	0.006	0.010
	[0.018]	[0.017]	[0.017]
Moral Incentive -	0.054**	0.029*	0.028*
Simple Reminder	[0.022]	[0.017]	[0.017]
Moral Incentive -	0.058***	0.045**	0.041**
Religious Placebo	[0.022]	[0.019]	[0.019]
Mean Repayment Control Group		0.34	
Month fixed effects	No	Yes	Yes
Controls	No	No	Yes
Waves	Waves I, II, and III	Full Sample	Full Sample
N	5821	13428	13428
R^2	0.002	0.011	0.057

Notes: Column (1) excludes customers late in June 2015, February 2016 and April 2016, and restricts the sample to customers assigned to one of the four following groups: moral incentives, simple repayment reminder, religious placebo (these treatments have not been run simultaneously in Wave IV, V and VI) and control. Column (2) and (3) use the whole sample. Columns (1) presents OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on treatment group dummies. The control is the omitted group, for which we report the mean repayment rate. Column (2) replicates and adds month fixed effects. Column (3) replicates and adds individual covariates (age, gender dummy, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). "Moral Incentive - Simple Reminder" gives the difference between the coefficient on "Moral Incentive" and the coefficient on "Simple Reminder." "Moral Incentive - Religious Placebo" gives the difference between the coefficient on "Moral Incentive" and the coefficient on "Religious Placebo." Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1.5: **What Drives the Moral Appeal? Religious Connotation**

Dependent variable	Dummy: customer repaid within the deadline		
	(1)	(2)	(3)
Moral Incentive	0.041**	0.051***	0.051***
	[0.019]	[0.013]	[0.013]
Implicit Moral Incentive	0.039**	0.041**	0.039**
	[0.019]	[0.018]	[0.018]
Non-Religious Moral Incentive	0.039**	0.040**	0.038**
	[0.019]	[0.018]	[0.017]
Moral Incentives -	0.001	0.011	0.011
Implicit Moral Incentive	[0.019]	[0.018]	[0.018]
Moral Incentives -	0.002	0.011	0.012
Non-Religious Moral Incentive	[0.020]	[0.018]	[0.017]
Mean Repayment Control Group	0.32	0.34	
Month fixed effects	No	Yes	Yes
Controls	No	No	Yes
Waves	Wave IV, V, and VI	Full Sample	Full Sample
N	4909	13428	13428
R^2	0.001	0.011	0.057

Notes: Column (1) restricts the sample to customers late in June 2015, February 2016 or April 2016 and assigned to one of the four following groups: moral incentives, moral incentives without quoting the Prophet, moral incentives without religion connotation (these last two treatments have been run only in Wave IV, V, and VI) and control. Column (2) and (3) use the whole sample. Column (1) presents OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on treatment group dummies. The control is the omitted group, for which we report the mean repayment rate. Column (2) replicates and adds month fixed effects. Column (3) replicates and adds individual covariates (age, gender dummy, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). "Moral Incentives - Implicit Moral Incentive" gives the difference between the coefficient on "Moral Incentives" and the coefficient on "Implicit Moral Incentive." "Moral Incentives - Non-Religious Moral Incentive" gives the difference between the coefficient on "Moral Incentives" and the coefficient on "Non-Religious Moral Incentive." Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1.6: **The Effect of Repeated Moral Messages**

Dependent variable	Dummy: customer repaid within the deadline			
	(1)	(2)	(3)	(4)
Repeated Moral Incentive	0.041	0.041	0.036	0.043*
	[0.030]	[0.030]	[0.031]	[0.025]
First Moral Incentive				0.045***
				[0.011]
Repeated Moral Incentive -				-0.001
First Moral Incentive				[0.025]
Mean Repayment		0.30		0.33
Control Group				
Month fixed effects	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Waves	Waves V and VI	Waves V and VI	Waves V and VI	Full Sample
N	898	898	898	14326
R^2	0.002	0.006	0.071	0.056

Notes: Column (1), (2) and (3) restrict the sample to customers part of the follow-up experiment, that is customers late in February 2016 or April 2016 and treated in a previous wave with a moral message. Column (4) uses the whole sample (both the main experiment sample and the follow-up experiment sample). Column (1) presents OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on a dummy for having received one of the three version of the moral message after being treated in a previous wave with the same moral message. The control is the omitted group, for which we report the mean repayment rate. Column (2) replicates and adds month fixed effects. Column (3) replicates and adds individual covariates (age, gender dummy, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). Column (4) replicates, adds a dummy for having received one of the three version of the moral message for the first time, and other treatment group dummies. "Repeated Moral Incentive - First Moral Incentive" gives the difference between the coefficient on "Repeated Moral Incentive" and the coefficient on "First Moral Incentive." P-value for the test of inequality "Moral Incentive < Cash Rebate" in parenthesis. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Chapter 2

Status Goods: Experimental Evidence

from Platinum Credit Cards

2.1 Introduction

Social image concerns affect many important behaviors, from donations to political behaviors to student effort (Perez-Truglia and Cruces, Forthcominga; DellaVigna et al., 2012a, 2017a; Bursztyn and Jensen, 2015a; Enikolopov et al., 2017a). A fundamental economic behavior – consumption – may also be shaped by social image concerns. Specifically, a desire to signal high income or wealth may cause consumers to purchase *status goods*.[1] In theory, such conspicuous consumption can impose negative positional externalities, and lead to wasteful spending in a consumption rat race.[2] Empirical research has highlighted the potential role of conspicuous consumption in important economic phenomena such as the wealth gap between Blacks and Whites in the United States (Charles et al., 2009), bankruptcy decisions (Agarwal et al., 2016, 2017), and large expenditures on weddings (Bloch et al., 2004) and festivals (Rao, 2001) among the poor in developing countries.[3]

However, directly testing for status concerns in consumption is challenging. With observational consumption data, it is difficult to fully separate unobserved consumption utility from a desire to signal high income. For example, a person who buys a Ferrari and an Armani suit could simply have a particularly strong taste for nice cars or fashionable clothes. Moreover, such consumption decisions could be driven by self-image and identity, rather than social image. That is, consuming the types of goods associated with wealth might

---

[1]See Veblen 1899; Duesenberry 1949, and Bagwell and Bernheim 1996.

[2]See, for example, Frank 1985, Banerjee 1990 and Hopkins and Kornienko 2004.

[3]In fact, the role of income-signaling in consumption was already pointed out by Adam Smith in the *Wealth of Nations*: "A linen shirt, for example, is strictly speaking, not a necessary of life. [...] But in the present times, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct" (Smith, 1776).

provide an individual with psychic utility, even if that consumption was invisible to others (Akerlof and Kranton, 2000b). More generally, self-image or identity and the demand for status could be deeply connected, and it remains an open question whether self and social image are substitutes or complements. Nor do we understand whether the demand for social image is purely instrumental, or instead is also valued for hedonic reasons.

In this paper, we (i) provide the first field-experimental evidence of the existence of status goods, and shed light on the magnitude of the demand for status; (ii) test for the associated positional externalities; and (iii) show that the demand for status is in part psychological (as opposed to purely instrumental), with social image acting as a substitute for self-image. We work with a large bank in Indonesia to design three related experiments that market the bank's popular platinum credit cards. The credit cards in our experiment are widely-recognized throughout Indonesia.[4] They are typically restricted to high-income customers, and come with a number of instrumental benefits, such as a higher credit limit and discounts on the purchase of luxury goods. Our sample consists largely of urban, (upper) middle-class bank customers. We consider this an important context in which to study conspicuous consumption. The developing world is experiencing rapid economic growth and urbanization – precisely the conditions under which Veblen originally theorized conspicuous consumption would be most important. Recent estimates suggest that approximately 130 million of 330 million global luxury good consumers are located in emerging markets.[5] In Indonesia, for instance, there are an estimated 74 million middle-class consumers, and this number is

---

[4]We confirm that the cards are viewed as prestigious, using survey evidence presented below.

[5]See "Luxury Goods Worldwide Market Study" *Bain & Company*, 2014. See Banerjee and Duflo (2008) for evidence on the consumption patterns of middle-class households around the world.

expected to double by 2020.[6] Such individuals are obtaining access to credit cards and a broader set of visible consumption and luxury goods.

Isolating the demand for status. The first experiment shows that a substantial part of the demand for the platinum card is explained by the desire to own the prestigious card itself, beyond the tangible benefits and services it comes with. The innovation of this experiment is to engineer a control product which holds constant all the instrumental benefits of the platinum credit card, while stripping away the associated status component. Specifically, we offer paid credit card upgrades to a population of bank customers. In a control group, customers are offered all the financial services and instrumental benefits of the platinum card, made available as a benefits upgrade on a nondescript credit card. In a treatment group, customers are instead offered an upgrade to an actual platinum card. In both groups, customers are truthfully told that they were randomly selected to receive the offer, to avoid providing information about their relative income and status.

We find that demand for the platinum card (21% take-up at market price) is substantially higher than demand for the instrumental benefits it comes with (14% at the same price), providing prima facie evidence of demand for the status aspect of the card.[7] The difference in demand for the two offers (7 percentage points) is economically meaningful: offering a 25% price discount on the instrumental benefits package in the control group increases

---

[6]See "Indonesia's Rising Middle-Class and Affluent Consumers". *Boston Consulting Group*, 2013.

[7]In a slightly modified variation of the platinum card script we instead truthfully informed customers that they were selected as a result of being among the bank's top customers. In principle, this might boost customers' self-image, yet it has little additional impact on take-up (23% compared to 21% for those informed they were selected at random). Note also that both scripts are truthful. The sample for this first experiment consisted of existing customers who are both selected to be higher income than the typical bank customer, *and* drawn randomly from the list of such customers, as well as randomly assigned to treatments. Thus, the customers are truly randomly selected and also truly chosen based on their income.

take-up by only 3.7 percentage points. Surveys and interviews of customers assigned to the control group suggest that the benefits package was fully credible. Despite believing that they would receive the exact same benefits and services as platinum card-holders, control group customers were less likely to accept the offer.

Demand for the status aspect of the card *decreases* with income. It is the relatively lower-income individuals in the sample who show the highest demand for the status aspect of the platinum card. By contrast, the richest customers show no differential demand for the actual platinum card compared to the instrumental benefits upgrade. Our interpretation is that richer individuals already have ways to signal their income, while the platinum credit cards are a more powerful (marginal) signaling tool for those with comparatively lower incomes.

Status-signaling in credit card transactions. Next, we analyze individual credit card transactions among a larger observational sample of customers to understand how the platinum card is used in practice, and whether this is consistent with social image motives. Exploiting the bank's assignment rules for credit limits and card types, we show that platinum card holders are more likely to use the card in social situations, such as spending in restaurants, bars and clubs, where the card is likely to be visible to others. This likely reflects platinum card holders substituting away from using other cards or cash for such expenditures, since a consumption recall survey reveals that actual restaurant visits do not differ between platinum and standard card holders. Moreover, the use of the platinum card for social signaling is in fact costly: while the card used in our study does not offer cash back rewards, at least 48% of platinum customers report owning other credit cards which *do* offer cash back at

restaurants and similar transactions.[8] Hence, the customers in our data forgo money in the form of cash back rewards each time they use the platinum card instead of other cards they may own to pay at a restaurant or bar. Taken together, these findings are consistent with the hypothesis that platinum cards are used to build social image.[9]

Valuation of social status. We can combine the results of the first experiment with the transaction data to shed light on how much consumers value social status in their social interactions. First, a simple calibration exercise utilizing the price variation from the first experiment suggests that consumers *value* the status aspect of the card by 218,000 Rupiah ($15.5) per year on average. Next, we note that the average user uses the card 4.8 times per year in social or visible situations. Suppose that the 'audience' infers that the card owner has the average annual income associated with the card: Rp. 522 million for platinum cards, and Rp. 215 million for gold cards, a difference of approximately Rp. 300 million. This implies that consumers, on average, value being seen as having Rp. 300 million higher income by Rp. 218,000/4.8 – approximately $3.2 – per interaction. While this back-of-the-envelope calculation depends on a number of assumptions, if consumers do value income signaling similarly in their other social interactions, the total value of status could be much higher. Products that could convey status to a larger audience (unlike a credit card) might thus command a substantial premium.

Positional externalities. Having established that status considerations play a substantial role in the demand and use of platinum credit cards, we turn to testing for 'positional

---

[8]Only 39% of non-platinum cardholders have other cards with such deals.

[9]Note that we cannot separate whether the greater use of the platinum card in social settings is a causal effect of the card, or whether those who care more about social image select differentially into the platinum card. We discuss this point in more detail in Section 4.

externalities' imposed by ownership of the cards (Frank, 2005). In a control group, current platinum card holders are offered an upgrade to a new, more-expensive but functionally identical, 'diamond card'. In the treatment group, customers receive the same offer, but are additionally informed that the income criterion for their existing platinum card – but not the new diamond card – has been recently reduced, so that some relatively lower-income customers now also qualify for the platinum card.[10] We find that providing this additional information nearly doubles take-up of the new diamond card. This result shows that the exclusivity of the platinum card matters for its demand, providing additional evidence in favor of a status-good model. Moreover, lower-income consumers weaken the status signal and thus impose a positional externality on higher-income consumers, even with instrumental benefits held fixed. Our finding supports the assumption underlying models of fashion cycles in status goods (Pesendorfer, 1995).[11]

Self-image and demand for status. In the final set of experiments, we provide indirect evidence that the demand for status is partly for hedonic reasons, rather than purely instrumental motives. In particular, we find that self-esteem – an important aspect of self-image – has a causal effect on customers' demand for status goods. To boost self-esteem, customers in a treatment group are asked to complete a self-affirmation task, in which they describe an event or achievement from their life which made them feel proud of themselves (Steele, 1988;

---

[10]This information is again truthful, as the income cutoff for the platinum card had in fact been recently reduced.

[11]It is worth emphasizing that the additional demand for the diamond card relative to the platinum card cannot be explained by customers using additional instrumental benefits as a cover to justify to others – or to themselves – why they are paying more for a good that provides more status. By holding fixed the instrumental benefits of the card, we also ensure that the results cannot be explained by differential inferences about the quality of the product, or about the suitability of specific benefits to different customer types.

Cohen et al., 2009; Hall et al., 2013).[12] A control group instead performs a placebo task, describing their media consumption habits. Both groups are then offered either an upgrade to the platinum card in the main condition, or the benefits upgrade in a placebo condition. While the experiment has limited statistical power, we find that the self-affirmation treatment does not affect take-up of the nondescript benefits upgrade, but reduces take-up of the platinum card by an economically meaningful but not statistically significant extent.

To build on this suggestive evidence, we conduct a higher-powered experiment with a parallel design on the online crowdsourcing platform *mTurk*. Instead of offering participants a platinum credit card or a placebo good, we elicit preferences between gift certificates for luxury brand clothing – a classic status good – versus low-end clothing, using an incentivized multiple price list procedure. We find a strong first-stage relationship between the self-affirmation treatment and a standard measure of self-esteem, and estimate a substantial and statistically significant *reduction* in willingness-to-pay for the status good as a result of receiving the self-affirmation treatment. That is, we find that having higher self-esteem results in lower demand for a conspicuous status good. More generally, our results suggest that having a more positive self-image reduces demand for social image; self and social image thus appear to be substitutes, rather than complements. The impact of a likely short-lived boost in self-esteem on the demand for status goods also suggests that at least part of the demand for status is driven by psychological rather than purely instrumental or economic reasons.

Taken together, our findings provide the first field experimental evidence on status goods.

---

[12]We show in an online experiment that this task temporarily increases self-esteem, but has no effect on the values that individuals cite as being most important to them.

We show that a desire to signal high social status, in isolation from instrumental utility or self-image considerations, can have a meaningful impact on consumption decisions. Moreover, the results from our second field experiment confirm an important prediction of these models, namely that the consumption of status goods creates a positional externality. Positional externalities can have important welfare effects by leading to wasteful consumption and inefficient innovation in the creation of status goods (Frank, 2005). By directly testing – and confirming– the key assumptions of status goods theories, our analysis suggests that the welfare and policy implications of these theories should be taken seriously. We view this causal evidence as supporting and complementary to the existing evidence from observational studies and natural experiments (Charles et al., 2009; Heffetz, 2011; Kuhn et al., 2011; Agarwal et al., 2016; Roth, 2014).

Using two entirely separate samples and products, we also show that higher self-esteem causally reduces demand for status goods. This evidence of substitution between self and social image may have implications beyond our setting. Factors lowering self-esteem –such as poverty, unemployment, or facing negative stereotypes– may magnify the effects of status-seeking behavior and increase susceptibility to social pressure more generally. Our finding might therefore shed light on related social phenomena, such as large wedding and festival expenditures by the poor in developing countries, and low-income, minority students conforming to harmful social norms at school.[13]

The remainder of the paper proceeds as follows. In Section 2, we describe our setting. In Section 3, we present the first field experiment, which isolates the demand for the social

---

[13]See, for example, the "acting white" hypothesis, Austen-Smith and Fryer, 2005.

status component of platinum credit cards. Section 4 presents the analysis of credit card transactions. In Section 5, we describe our second field experiment, establishing positional externalities. In Section 6, we present the final set of experiments, which examines the relationship between self-image and the demand for status goods. Section 7 concludes.

2.2   Setting: The Credit Card

The market for credit cards in Indonesia has several features that make it an especially attractive setting to study status goods. First, Indonesia is an important emerging market economy with a large and rapidly growing middle class. Credit cards are widely used, and premium credit cards have a comparatively high income threshold relative to median income, making them a credible and well-recognized signal of status and economic success. Second, working with a bank, we are able to vary the instrumental benefits and services offered with a given credit card. This allows us to construct control products that vary specific features of the credit card in order to distinguish demand for the instrumental benefits of the card from demand due to signaling motivations. Third, we can link each card to its full transaction history, to understand whether the use of the cards in an everyday setting is consistent with status-signaling motives.

We work with one of Indonesia's leading banks to conduct a series of field experiments. The bank has approximately 200,000 credit card customers across Indonesia and offers its credit card product in three tiers: classic, gold and platinum. The three tiers of the credit card are clearly vertically differentiated based on income. The platinum card has the high-

est income-eligibility criterion, followed by the gold card with the second highest income requirement and the classic card with the lowest income requirement. At the time of our experiment, a new customer was required to document an annual income of Rp 36 millions (US$2,556) to qualify for a classic card, an annual income above Rp 60 million (US$ 4,260) to qualify for a gold card, and an income above Rp 500 million (US$ 35,500) to be eligible for a platinum card.[14][15] Customers are charged a fixed annual fee of Rp 120,000 (US$ 9) for a basic card, Rp 240,000 (US$ 17) for a gold card, and Rp 600,000 (US$ 43) for a platinum card, plus a monthly membership fee equal to 2.75% of the customer's credit limit.[16]

Consistent with the eligibility requirements, only 10% of active credit card customers qualify for a platinum card, 72% of card customers have a gold card, and the remaining 18% qualify only for the classic card. The average (median) customer in the sample of active credit card clients has a reported annual income of Rp 154 million or US$ 10,934 (Rp 60 million or US$ 4,260). The bottom quartile of the credit card customer population is close to the median income of urban Indonesia, while the median credit card customer is in the top 15% of urban incomes in Indonesia. Even the lowest-income platinum card customers rank in the top percentiles of the Indonesian income distribution, so that qualifying for a platinum card plausibly serves as a strong signal of high (relative) income.

Importantly, the three tiers of the credit card also differ in their design, as shown in

---

[14]In November 2014, the eligibility threshold for the platinum card was reduced to Rp 300 million (US$ 21,300).

[15]The eligibility criteria for customers who are already clients of the bank can alternatively depend upon the client's deposit account balance, and on their credit history with the bank, say from consumer or housing loans.

[16]The annual fees are often waived for new customers as a result of various promotions and marketing initiatives.

Figure 2.1. Most notably, the platinum card is differentiated from the two lower tier cards in both color and design. It is dark purple and has the word *'Platinum'* printed in large cursive letters across the front of the card. All three tiers of the card are well-recognized and marketed throughout Indonesia using print, billboard, and online advertising that includes images of the cards.

To test for public recognition of the platinum card –a necessary condition for status signaling– we interviewed 113 randomly selected respondents at shopping malls in greater Jakarta, and presented them with pictures of the gold and platinum cards. The overwhelming majority of respondents (93 out of 113) ranked the cards correctly in terms of their income requirements. This provides prima facie evidence that the platinum card indeed signals high income and economic success relative to the gold card. Of course, this need not imply that status concerns are an important component of consumer demand for the platinum card, since the cards also differ in credit limit, price and other potentially valuable benefits. For example, the gold card has a credit limit between Rp 10 million (US$ 710) and Rp 30 million (US$ 2,130) , while the platinum card has a credit limit starting at Rp 40 million (US$ 2,840), and extending up to Rp 125 million (US$ 8,875) for the very highest-income clients. Platinum card customers also enjoy additional instrumental benefits: they can access premium airport lounges, receive cash-back discounts on international fashion brands, and are eligible for additional special offers and promotions available only to the bank's premium credit card customers.

While several features of the platinum credit card – the high income eligibility criteria, and the bold *'Platinum'* labeling – suggest the potential importance of status or income

signaling in demand for the card, this is clearly confounded with the differences in credit limit, instrumental benefits and price. In the following section, we report a field experiment designed specifically to remove these confounds and test for a demand for status in the context the platinum credit cards.

2.3 Experiment 1: *The Demand for Status*

In our first experiment, we test whether part of the demand for the platinum card is purely due to status motives. In order to isolate the status component of the card from its instrumental benefits, we engineer a control product which has exactly the same instrumental features as the platinum card, but lacks the visible appearance of the platinum card, thus striping away the status signaling aspect. We offer this card as a paid upgrade to existing bank customers in a randomly-assigned control group, and compare take-up to a treatment group in which customers are instead offered the actual platinum card itself. We utilize price variation to interpret the magnitude of demand for the status aspect of the card, and examine heterogeneity in the demand for status.

2.3.1 Theoretical Framework

To motivate our experiments and to interpret results, we adapt the framework of Bénabou and Tirole (2006) to our setting. In this framework, an individual exhibits social image concerns if her utility depends on the inferences that others make about her type, based on observable behavior.

Formally, consider an individual i, undertaking an observable action $a_i \in \{0, 1\}$, which may reveal information about her type. In our case, individual i is offered a status good and decides whether or not to take up the offer, and we let $a_i = 1$ denote the case in which individual i purchases the status good, and $a_i = 0$ the case in which she does not purchase the status good. Since status goods are assumed to be visible to others, they may reveal information about i's type. More specifically, it might allow others to make inferences about individual i's income $y_i \in \{l, h\}$, where $y_i = h$ indicates that i is high-income.

We assume that, in our setting, it is socially desirable to be viewed as wealthy by others, so that an individual's utility function includes the social image term S_i, which we define as:

$$S(a_i) = \lambda \mathrm{Pr}_{-i}\left(y_i = h \mid a_i\right). \tag{2.1}$$

In this equation, $\mathrm{Pr}_{-i}\left(y_i = h \mid a_i\right)$ represents the posterior probability that the members of individual i's reference group think that her income is h, conditional on observing individual i undertaking action $a_i \in \{0, 1\}$. The parameter λ measures how much individuals care about being perceived as being of type h. In the context of our first experiment, the null hypothesis that individuals care only about the material benefits of a product and derive no separate utility from its status component implies that $\lambda = 0$. The alternative hypothesis that individuals care about the pure status signaling component of a product, on the other hand, would imply that $\lambda > 0$.

Our first experiment is designed to test these competing hypotheses. In the following sections, we will return to and extend this simple theoretical framework in order to generate

testable predictions and elucidate the experimental design.

2.3.2 Experimental Design

Set-up and Experimental Protocol

The sample for this experiment consisted of 1,260 customers identified by the bank.[17] The customers on this list were randomly drawn from the set of current gold card holders with a credit limit of at least Rp 20 million (US$1,420), who were current on their credit card payments, and were not bank employees. Essentially, these were customers whom – for the purpose of our relatively small experiment – the bank was willing to offer an upgrade to the platinum card, even though they may not have normally qualified for it. Customers in this sample were then assigned to one of the treatment conditions described below. Treatment status was assigned randomly at the individual level, stratifying on income (below Rp 300 million per year, between Rp 300 million and Rp 500 million, or above Rp 500 million) and on customers' current annual card fee (equal to Rp 240,000 or waived). Appendix Table B.1 displays sample characteristics for all experiments. In the sample for our first experiment, 24% of participants are female, and the average age is 47 years.

To implement the experiment, the bank made marketing calls to customers in this sample. In the calls, all customers were offered a paid upgrade to the benefits, services and credit limit available to the bank's platinum card holders. However, customers were randomly and individually assigned to one of two treatment arms, described in greater detail in Section

---

[17]This experiment was registered in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0000828.

2.3.2 below, which varied the details of the script, as well as the characteristics of the product that was being offered. Customers in a treatment group were offered an upgrade to an actual platinum card, while customers in a control group were informed that they were offered an upgrade to all the benefits and services usually reserved to the platinum card, but as an add-on to their current gold card.

In order to minimize any effects that might arise from the offer's impact on participants' beliefs about themselves (i.e., their self-image), customers were told that they had been randomly selected to receive this offer. In both treatment conditions, customers were informed that the upgrade was available for a price of Rp 360,000 (US\$ 26), in addition to the customer's current annual fee.[18]

The experiment was conducted over the course of one week. Each day, four callers made phone calls to a randomly assigned list of credit card customers from the sample.[19] The order of client names on each caller's list was randomized, and callers made phone calls in the order provided on the list. Each client received the offer only once, but up to three call attempts were made if a client could not be reached or was busy at the time of a previous attempt. However, no additional calls were made once any part of the offer had been revealed to a respondent. All calls were recorded and checked to ensure adherence to the script. The full scripts for all experiments are available in the Supplementary Appendix.

---

[18]Customers who already pay an annual fee of Rp 240,000 thus will have to pay a total of Rp 600,000 to obtain these services (the same annual fee as that of a platinum card), while customers who have their annual fee waived will start to pay Rp 360,000 a year if they want the benefits upgrade.

[19]In total nine phone callers worked on this marketing experiment, rotating over different days.

Experimental Treatments

The treatments in this experiment were designed to hold the instrumental benefits of the offer constant, while varying the status component of the product by randomizing the appearance of the card (gold or platinum) customers would receive upon accepting the offer.

Credit card customers assigned to a treatment group –the *platinum upgrade* treatment condition– were offered an upgrade to an actual platinum card, while customers assigned to a control group –the *benefits upgrade* treatment– were offered these services as an add-on to their current gold card. Hence, customers assigned to the *platinum upgrade* treatment were offered the benefits upgrade along with the bank's regular platinum card, using the following script:

> *You have been randomly chosen to receive an upgrade to our platinum [name of card] card. With this upgrade, you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. [...] To make all the extra benefits available, we will have to send you a new [name of card] card. The card you will receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it!*

while customers in the *benefits upgrade* treatment were offered the same upgrade as an add-on to a card that looks identical to the credit card they currently hold, using the script:

> *You have been randomly chosen to receive an upgrade on your gold [name of card] card. With this upgrade, you will get the same services, benefits, credit*

limit, terms and conditions offered to platinum [name of card] card cardholders.

[...] To make all the extra benefits available, we will have to send you a new gold

[name of card] card. It looks just like the one you already own, but includes all

the benefits and services of our platinum [name of card] card.

Hence, all customers are offered an upgrade to the same instrumental benefits. They are also informed that only 10% of customers normally qualify for these benefits, in order to hold equal beliefs about the exclusivity of the benefits. All customers who accept the offer are sent a new card in the mail, to hold hassle costs equal across the two arms. The only difference is the physical appearance of the new card they receive. One group receives the conspicuously labeled platinum card, while the other does not.

One might be concerned that telling customers they were randomly chosen to receive the upgrade offer is unnatural. This is certainly not how the bank usually markets platinum credit cards. We thus also implemented a mild variation of the platinum script, the *platinum upgrade merit* condition, in which customers were informed that they had been selected as a result of being among the bank's top customers. Both statements are true, since customers were randomly selected from a relatively high-income sub-population of the bank's gold card customers. Customers in the *platinum upgrade merit* were read the same script as described above, but with one twist: instead of being told they were randomly chosen, they were told that *"As one of our top customers, you have been chosen to receive an upgrade to our platinum [name of card] card."* This treatment might be perceived as more natural, but potentially boosts customers' self-image by providing them positive information about their own relative standing. As it turns out, the luck versus merit variations of the treatment have

73

no differential effect on take-up, so we pool them in future results.

Testable Predictions

We use the simple theoretical framework described above to derive two testable predictions regarding the existence and demand for status goods.

Demand for status. We assume that the good is available at price p, which enters linearly into the buyer's utility function, and has some inherent instrumental value, from which individuals derive utility b_i. In the benefits upgrade condition, customers are not given the option to purchase the status component of the good, so that $a_i^{platinum} = 0$, always. Hence, a customer will accept the offer if $b_i - p > 0$. That is, if the purely instrumental value of the benefits upgrade is greater than its cost. In the platinum condition, on the other hand, customers have the option of purchasing the status component of the product, and $a_i^{platinum} = 1$ if and only if the customer accepts the offer. The customer will accept the offer if $b_i - p + S^{platinum}(1) - S^{platinum}(0) > 0$. That is, if the utility from the instrumental and status benefits of the upgrade is greater than its cost.

Prediction 1. *If customers care about social image and the platinum card is a status good (that is, $S^{platinum}(1) > S^{platinum}(0)$), then take-up of the upgrade offer in the platinum upgrade condition will be higher than in the benefits upgrade condition.*

Hence, if the share of customers demanding the platinum upgrade is higher than the share of customers demanding the benefits upgrade when it comes with a gold card, we will have established that customers derive utility from the status associated with the appearance of the platinum card.

Income and the demand for status. Since our experiment considers a marginal income-signaling decision, it is worth noting that individuals might have other chances to signal their income, independent of their decision in the experiment. It therefore seems reasonable to assume that the marginal gain in social image from the status good is smaller for higher-income individuals.[20] This will be true, for example, in a model in which wealthy individuals have access to a larger set of status goods, individual i owns multiple status goods, and is perceived as wealthy if at least one of these status goods is observed by others.

Prediction 2. *The difference in take-up rates between the platinum upgrade and benefits upgrade conditions is smaller among individuals with higher incomes than among those with lower incomes.*

2.3.3 Main Results: The Demand for Status

Treatment Effects

Main result. We begin by comparing take-up of the control and treatment offers in Figure 2.2. At the same price, the take-up rate for the benefits upgrade offer is 13.7%, compared to 21% for the actual platinum card. The 7.3 percentage point difference between the two treatment effects is statistically significant at the 5% level (p-value=0.029).[21] Table 2.1 presents OLS regressions. Column (1) replicates the findings from Figure 2.2. In column (2), we add caller fixed effects and baseline covariates. The results are unchanged, consistent with successful

---

[20]In the terms of our theoretical framework, this implies that the perceived probability of being perceived as the high income type is $\Pr_{-i}(y_i = h \mid a_i = 1, y_i = h) - \Pr_{-i}(y_i = h \mid a_i = 0, y_i = h) < \Pr_{-i}(y_i = h \mid a_i = 1, y_i = l) - \Pr_{-i}(y_i = h \mid a_i = 0, y_i = l)$.

[21]The p-values for all experimental results are based on permutation tests. This ensures that our inferences are valid in finite samples.

randomization across treatment conditions. Since both groups were offered exactly the same financial benefits and customer service, we interpret the difference in demand for the two products as evidence for a demand for status.

We next compare take-up rates in the *platinum upgrade* and *platinum upgrade merit* treatment conditions in Figure 2.2. The take-up increases only marginally from 21% to 23% in the *platinum upgrade merit* relative to the *platinum upgrade* condition (p-value=0.539). On the one hand, this provides reassuring evidence that being informed that they were "randomly chosen" to receive the platinum offer was not off-putting or perceived as particularly unnatural by customers. We hesitate to conclude, however, that self-image or identity play no role in the demand for status goods. Instead, we consider it likely that the merit script simply failed to move self-image or identity substantially. Since there is no significant difference in take-up rates between these two conditions, we pool these two groups in the following analysis to increase precision. When we pool the two platinum card treatments in Table 2.1, column (3), take-up in the *platinum pooled* condition is 22% as compared to 13.7% in the *benefits upgrade* condition, and the difference is statistically significant at the 1% level (p-value=0.005).

Price variation. In order to price the status value of the platinum card, we compare the increase in take-up from offering the platinum card (relative to the benefits upgrade) with the effect of price discount on the benefits upgrade offer. While we did not employ randomized price variation due to limited sample size, we obtained *within person* price variation in the benefits upgrade group. Specifically, the bank made a second call to customers who had declined the first offer, and offered them a discount of Rp 90,000 per year (approximately

$6). This 25% discount increased demand for the benefits upgrade by only 3.7 percentage points, less than half the effect of offering the platinum card itself, as presented in Figure 2.3. If anything, we suspect this overstates the effect of the price discount, since being asked a second time might induce some consumers to accept the offer even in the absence of a price cut.[22] A simple calibration exercise (see Appendix B) matching take-up of the platinum, benefits upgrade and discount treatments suggests that the average consumer values the status aspect of the card by Rp. 218,000 ($15.5) per year. While interpreting this magnitude, it is important to note that the platinum card provides limited natural opportunities to signal status: one must be making a purchase in a social context, at an establishment which accepts credit cards, with others present for the card to be noticed. In the next section, we combine this estimate with details on the usage of the card, to infer how much status value the card provides per use.

Heterogeneity. We next estimate the heterogeneity of treatment effects by income.[23] In Figure 2.4, we present take-up rates for the *benefits upgrade* and *platinum pooled* conditions, separately for customers with incomes below and above Rp 300 million. We find evidence that demand for status is higher for lower-income customers. The difference in take-up

---

[22]The bank was able to reach 70% of the customers who had declined the first offer after hearing the price details, and 9.6% of them accepted the second offer. We assume that (i) customers' decisions in the second call would have been the same as in the first call if they were offered the benefits upgrade with a 25% price discount, (ii) customers that we did not reach again would behave similarly to those that were reached, (iii) the 13.7% of customers who accepted the offer at full price would have also accepted the offer at a lower price, and (iv) the 48% of customers who were reached and refused the offer after hearing about the product but before hearing about the price would have also refused the offer at a discount. This yields a take-up rate for the benefit upgrade with discount of 17%. The p-value of two-sided test that the effect of platinum is the same as the effect of a 25% discount is 0.12. The p-value of a one-sided test against the alternative hypothesis that the effect of platinum is higher is 0.06. These p-values are calculated using a bootstrap procedure.

[23]Recall that we stratified the randomization by income, using the income groups $y_i <$Rp 300 million, Rp 300 million $\geq y_i <$ Rp 500 million, and $y_i \geq$ Rp 500 million.

rates between platinum and benefits upgrade conditions is 10.5 percentage points for lower-income customers (p-value=0.003), while the same difference for higher-income customers is only 3.1 percentage points (p-value=0.392).[24] We find similar heterogeneity results when we control for caller fixed effects and baseline covariates (Table 2.1, column (4)) and when we consider Rp 500 million as a threshold to define higher- and lower-income individuals (Table 2.1, column (5)). This is consistent with the hypothesis that the marginal gain in social image from owning the status good is decreasing in income. In Table B.3, we also provide suggestive evidence that the demand for status is higher in cities and larger metros, consistent with Veblen's hypothesis that conspicuous consumption will be higher in areas with greater anonymity and income mobility.

Alternative Channels and Interpretations

In this subsection, we consider a number of confounding factors that could explain our results and discuss which of these alternative channels can be ruled out.

We first consider reasons unrelated to status signaling which might make the benefits upgrade treatment unattractive relative to the platinum card offers. First, customers might not have believed that the terms and conditions – such as the credit limits and customer service – in the *benefits upgrade* condition would in fact be identical to the platinum card, despite the fact that the bank explicitly stated in the offer that they would. Second, customers might have been offended that they were offered the instrumental benefits of the platinum card, but not the actual platinum card.

---

[24]The p-value of a test that the effect of the platinum offer is the same for lower and higher-income groups is 0.284.

To test for these concerns, we conducted a follow-up survey with customers in the *benefits upgrade* condition who had turned down the offer. The interviewer first asked customers an open-ended question about why they rejected the offer. Next, respondents were prompted with a list of potential reasons, including (1) beliefs about the benefits and services relative to the platinum card, (2) the usefulness of the benefits, (3) the annual fee, and (4) reactions to being offered a benefits upgrade instead of being offered the platinum card itself. Only 1% of the respondents stated that they had doubts that the quality of the benefits and services would be identical to the platinum card. None of the respondents reported being offended by not being offered the platinum card. Among the stated reasons for not accepting the offer, 67% of respondents answered that the annual fee was too high, and 68% said they did not use their existing card enough to justify paying for an upgrade. None of the respondents reported being concerned that the benefits package would come to differ from the platinum card benefits in the future. Taken together, these results suggest that the benefits upgrade offer was found to be believable, and the striking difference in take-up between the instrumental benefits and the platinum card is not explained by customer suspicion, confusion or any offense from not being offered the platinum card.

We next turn to alternative interpretations of the heterogeneity of treatment effects by income. One clear concern is differential selection into the experimental sample itself. Our sample was drawn from existing gold card customers. Could it be that the higher-income customers in our sample were individuals with a particularly low demand for status goods, who had previously declined the platinum card? In fact, our sample consists only of customers who had previously never received an offer of a platinum offer. In addition, we can

focus on customers with income between Rp 300 million and Rp 500 million. These customers were not eligible for the platinum card when they opened their account (and therefore could not have turned down a platinum offer). When we exclude customers with income above Rp 500 million in Table 2.1, column (6), we find a very similar result, suggesting that our income-heterogeneity results are unlikely to be driven by selection. Instead, consistent with our prior, lower-income customers appear to value the status signal more at the margin.

Finally, while we focus on the interpretation that higher-income individuals may have alternative ways to signal that they are high income than lower-income individuals, there could be alternative reasons why the treatment effect is higher for lower-income customers. It could be that higher-income individuals simply care less about social status. However, our second experiment, in which we offer the highest-income customers an opportunity to further differentiate themselves from the premium cardholder population, helps rule out this alternative explanation (Section 2.5).

2.4 Status Signaling in Credit Card Transaction Data

The results of our first experiment show that customers exhibit substantial demand for the pure status component of the credit card, beyond any instrumental benefits that the card provides. We suspect that individuals use the card to signal their income in order to build social status. In this section, we use detailed historical transaction data for a large sample of credit card customers to examine whether the usage of platinum cards in everyday life is consistent with social signaling motivations. To do so, we proceed in two

steps. We first identify certain transactions, such as spending in restaurants and bars, as *'social expenditures'*, in which the credit card is likely to be visible to one's peers. We then examine whether platinum cardholders are more likely than gold cardholders to use their card in such social contexts.

2.4.1 Data and Empirical Strategy

We use credit card transaction data for a sample of 2,492 customers with active credit cards who opened their accounts between January 2014 and August 2015. Within this sample, we focus on customers with credit limits of Rp 20 million, Rp 30 million, Rp 40 million, and Rp 50 million, respectively. The credit limit for each customer is assigned based on a combination of the customer's income and credit history, and there are multiple credit limits within each tier of the card. With few exceptions, gold card holders have a credit limit of Rp 20 million or Rp 30 million, while the vast majority of platinum card holders have a credit limit of Rp 40 million or Rp 50 million.

For the customers in our sample, we observe all transactions between January 2014 and August 2015, along with detailed information on the transaction amount, type and location. Using this information, we categorize transactions as either visible, online, or retail. We define visible transactions as transactions made in restaurants, cafes, and bars (89%), in membership clubs (2%), movie theaters (2%), and other amusement and recreational services (7%). The idea is to identify uses of the credit card which are likely to be observed by one's peers, such as friends, family or business associates, to whom one might wish to signal high income. The opposite type of transaction would be an online purchase, where no one other

than the cardholder observes the card being used. We identify online transactions by looking for internet-related terms, such as "www", ".com", "e-store", in the text description that comes with each transaction.[25] The third category we consider consists of retail transactions where the card may be visible to a salesperson, but that do not occur in an explicitly social setting. These transactions comprise purchases in supermarkets, grocery and convenience stores (30%), department stores (10%), service stations (7%), clothing stores (6%), and other merchants, such as pharmacies, etc. (47%).

Note that in there is no experimental variation in platinum card ownership in this sample, so that we must address the likely omitted variable bias introduced by simply comparing gold and platinum card holders.[26] Our approach is to compare the share of different types of transactions for customers with Rp 40 million credit limit (the lowest-income platinum card holders) with customers with a Rp 30 million credit limit (the highest-income gold card holders). We then exploit the existence of different credit limits within each tier of the card. Specifically, we contrast differences in card usage at the Rp 30 versus Rp 40 million credit limit (where the credit card tier *and* the credit limit change) with differences in card usage around the Rp 30 and Rp 20 million credit limit (where the card tier is held constant at gold, but the credit limit changes) and the Rp 40 and Rp 50 million credit limits (where the card tier is held constant at platinum, but the credit limit changes). We can therefore identify differences in transaction patterns due to a different tier of the card from changes

---

[25]We exclude all the purchases from airlines, since the bank offers special travel promotions to platinum cardholders.

[26]We were not able to separately acquire the transactions data for the experimental sample from the partner bank. In addition, given the moderate take-up of the cards, it is unlikely that this sample would provide sufficient statistical power to allow us to detect changes in transaction patterns.

in transaction patterns due to a different credit limit.

2.4.2 Results

Main Result: Visible Transactions

Figure 2.5 displays the raw shares of visible transactions for customers with different credit limits. Column 1 of Table 2.2 presents these results in regression format. The highest credit gold card customers (Rp 30 million credit limit) have 12% of their transactions in the visible category. This share increases by 6.1 percentage points considering the lowest credit platinum customers (Rp 40 million credit limit). There is no significant change in the share of online transactions (Table 2.2, column 3), and a significant decrease in the proportion of retail transactions (Table 2.2, column 5).

In contrast, there is no significant difference in the shares of visible, online and retail transactions between customers with Rp 30 versus Rp 20 million credit limits (both gold card holders) and between customers with Rp 50 versus Rp 40 million credit limit (both platinum card holders). These results suggest that the difference in consumption patterns between customers with Rp 40 million and Rp 30 million credit limit is not simply related to a credit limit increase.[27] The same pattern remains once we control for customers' observable characteristics, such as income, age, gender, and religion (Table 2.2, columns 2, 4, and 6).[28]

---

[27]The p-value of a test that the difference in the share of visible transactions for customers with credit limits of Rp 40 million and Rp 30 million is the same as that between customers with Rp 30 million and Rp 20 million is less than 0.01. The p-value of a test that the difference in the share of visible transactions between customers with Rp 40 million and 30 million credit limit is the same as that for customers with Rp 50 million and Rp 40 million is 0.09.

[28]We also consider an alternative regression model in which we instrument platinum card with a dummy equals to one if credit limit is greater or equal to Rp 40 million and control for credit limit linearly. This

Interpretation: A Costly Signal

Changes in consumption versus changes in modes of payment. Do these changes reflect actual differences in consumption, or customers switching to using the platinum card instead of using cash or other credit cards? Note that both possibilities are consistent with the use of the card for status-seeking behavior. To shed light on this question, we conduct a retrospective consumption survey with 362 customers randomly drawn from this sample, and find only a small (and statistically insignificant) increase in the number of restaurant meals in the last month. Owning a platinum card thus does not actually make customers go to restaurants more. Nor do platinum card owners appear to be differentially selected for their interest in restaurant visits. Yet, they pay quite differently for these restaurant expenditures. Is this costly signaling behavior, or are there other reasons to use platinum cards in restaurants over other means of payment?

Opportunity cost of card usage. The platinum card we study offers discounts on some luxury brands like Armani and Gucci, but does not offer cash back or discounts in restaurants. The increase in the share of visible transactions is thus not driven by simple price effects. In fact, a survey with these customers reveals that 48% of platinum customers with the Rp 40 million credit limit own other credit cards that *do* offer cash back.[29] Platinum card holders therefore appear willing to pay a cost to show off their platinum cards, forgoing cash back

---

model estimates the effect of holding a platinum card on consumption patterns controlling for the effect of credit limit, taking into account that a few customers with credit limit lower than Rp 40 million hold a platinum card. The coefficient for the dummy equal to one if credit limit is greater or equal to Rp 40 million in the first-stage regression is equal to 0.98. Results using this alternative model are also consistent with a change in consumption patterns for platinum card holders, as presented in Appendix Table B.4.

[29]This share is only 39% for gold customers with Rp 30 million credit limit (the p-value of the difference is 0.0676).

from other cards.

Note that we cannot claim to have identified the *causal* effect of owning a platinum card on transaction patterns. Our results are consistent with differential selection into the cards: those who have a higher demand for social status (although not, apparently, a higher demand for restaurants per se) might have been more likely to accept the platinum card offer. In either case, the results are consistent with customers using the platinum card to signal status.

The Value of Social Status

Based on our calibration exercise in Section 2.3.3, the average consumer values the status aspect of the card by Rp. 218,000 ($15.5) per year. Assume that the utility customers derive from the social aspect of the card comes from the signal of higher income the card sends others when they use it in a social situation. The survey described in Section 2 confirms that people who observe the card correctly infer that platinum cardholders have higher income. Suppose their posterior upon observing a card is equal to the average annual income associated with the card type: Rp. 522 million for platinum cards, and Rp. 215 million for gold cards, a difference of approximately Rp. 300 million. The average platinum card customer uses the card 4.8 times per year in social or visible situations. Then, consumers on average must value being seen as having Rp. 300 million higher income by Rp. 218,000/4.8 (=$3.2) per interaction.

Of course, the above calculation relies on various assumptions. For example, the number of "views" of the card may be higher than we observe, if customers sometimes show the card

to others without actually using it. This would cause us to overestimate the status value *per* transaction. On the other hand, we are likely underestimating how much consumers value being seen as higher income, per dollar of higher income. One's peers likely have somewhat informed priors about one's income, and thus the extent of updating of beliefs from noticing the card is likely smaller than Rp 300 million. Customers would thus be willing to pay $3.2 for a much smaller boost in social image.

The platinum card provides limited opportunities to signal status, and is ultimately visible to a small if important set of peers. If consumers value income signaling similarly in other interactions, then products that convey status to a larger audience, such as expensive cars and clothing, might command a substantial premium. A car that signals a similar Rp. 300 million income difference to a thousand people might have a status value of $3,200. The economic importance of status considerations in consumption could thus be quite large.

2.5 Experiment 2: *Positional Externality*

Intuitively, the signaling value of a status good depends on the type of customers who are expected to own it. To earn status, one wants to display goods that are known to be owned by 'high types', and inaccessible to 'low types'.[30] This implies that when individuals with comparatively lower social status gain access to a status good, the signaling value of the good diminishes, imposing a negative 'positional' externality on high-status owners of the status good. This, in turn, should induce the earliest adopters to demand a more exclusive

---

[30]In our setting, 'type' is synonymous with income. However, there are of course also status goods that are not allocated based on income, such as membership in prestigious clubs or professional organizations.

status good, a dynamic captured in models of fashion cycles (see Pesendorfer 1995).

In this section, we describe an experiment with credit card customers that tests for positional externalities in the consumption of a status good. Conceptually, our experiment relies on two steps: First, we truthfully inform a random subset of platinum cardholders about a reduction in the income eligibility threshold for the platinum card from Y_1 to $Y_2 < Y_1$, which should reduce the perceived income signaling value of the platinum card. Second, we estimate the impact that such a reduction has on the demand for a new status good – one with unchanged signaling value, that is, an income eligibility cutoff held constant at Y_1.

The design of our experiment takes advantage of a recent change in the credit card's income eligibility requirements. A few months prior to this experiment, the bank had reduced the income threshold necessary to qualify for a platinum credit card from Rp 500 million (US$ 35,500) to Rp 300 million (US$ 21,300). Our research design relies on *existing* platinum card customers, who joined under the old income criterion, being unaware of the recent change. At the same time, the bank was considering introducing a new credit card tier above platinum –the 'diamond card'– reserved for its highest-income customers.

As part of the bank's market research surrounding the new product, we conducted a take-up experiment offering the diamond card as a paid upgrade to a sample of existing platinum card customers. The experimental treatments varied whether these customers were additionally informed that the income threshold for their current credit card (the platinum card) had been recently reduced. We show that demand for the more exclusive status good, the diamond card, is causally higher when customers are informed about the new income requirements for the platinum card. We interpret this as evidence of a positional externality

imposed by new lower-income customers, who are now able to acquire the platinum card, thus reducing the income signal the card provides.

2.5.1   Set-up and Experimental Protocol

The experiment was conducted with a sample of credit card customers who had been identified by the bank as being eligible for an upgrade to the diamond card, once the new card would become available. All 180 clients in this sample were customers who, at the time of the experiment, had a platinum card, and an annual income of at least Rp 500 million (US$ 35,500).

The bank made calls to eligible customers, again following a set of standardized scripts. All customers were informed that the bank was considering launching a new credit card, reserved for its top customers. The caller explained that the diamond card would have *the exact same* services, benefits, credit limit, and additional services available on the platinum card, but would differ from the platinum card in color and design. Customers were then informed of the annual fee for the new card, and were asked if they would be interested in the diamond card once it became available. To introduce a cost to answering 'yes', we added a fee for receiving the formal upgrade invitation upon launch. The calls thus held the instrumental characteristics of the two cards identical, while experimentally manipulating perceptions of the income signal associated with the platinum card.

The calls were conducted following a procedure similar to that of our main experiment. Callers were assigned a list of randomized phone numbers and were instructed to follow the order of clients on the list. Call attempts were made on different days of the week. Each

client received the offer only once, but up to three call attempts were made if a client could not be reached at the phone number provided by the bank. However, no further call attempts were made in cases where a respondent had been reached and any part of the offer had been disclosed. Calls were recorded to evaluate adherence to the experimental scripts, and no substantive deviations were discovered.

2.5.2 Experimental Treatments

We implement two treatment conditions. In both arms, customers were first informed that the bank is considering introducing a new credit card. This was explained using the following script:

> *I am calling from [name of bank] and would like to ask you a question related to your [name of card] credit card. [...] We'd like to hear the opinion of our customers before deciding whether to launch a new credit card. The new card we are considering will be called the diamond [name of card] card. The diamond card will have exactly the same credit limit, benefits, services, and terms as the platinum [name of card] card, which you presently own. The only difference is that the diamond card will come in a new and different design and color from the platinum card you currently have.*

Customers assigned to the *positional externality control* group received *only* this product description, while customers assigned to the *positional externality treatment* group were additionally informed that the bank had recently relaxed the eligibility criteria for the platinum

89

card, so that more customers with lower average incomes are now eligible for the platinum card:

> *Everyone knows that nowadays banks have started giving platinum cards to nearly anyone. Even at [name of bank], we have recently reduced the income eligibility criteria for the Platinum card to 300 Million Rp, so now many customers with a lower income than yours will get the platinum card. However, these lower income customers can not apply for a diamond card.*

All customers were then asked whether they would upgrade to the new diamond card at an annual fee of Rs 650,000 (US\$ 46) – Rs 50,000 more than the fee associated with the platinum card. To add real (albeit modest) stakes to the sign-up decision, customers were also asked whether they were willing to be charged Rp 10,000 (approximately US\$ 1) to receive a formal offer once the card was launched. In practice, all customers who indicated that they would sign up for the card agreed to to pay this fee, suggesting the stated preference was not simply cheap talk.

2.5.3 Testable Predictions

If demand for the status aspect of the cards arises from income signaling motives, informing customers that individuals with lower income can now access the platinum card will weaken the associated income signal, reducing demand for the platinum card relative to the diamond card (holding instrumental benefits fixed).

In terms of our theoretical framework, with social signaling utility S, this can be expressed

as follows. Let the social image or income signal associated with ownership of the diamond card be denoted $S^{diamond}(1)$. The perceived social image associated with the platinum card will depend upon whether customers are informed about the reduced income threshold or not, denoted as $S^{platinum}_{info}(1)$ and $S^{platinum}_{no-info}(1)$, where we assume it to be public knowledge that the diamond card is associated with a higher income criterion than the platinum card.

Since the instrumental benefits b are explicitly held equal across the platinum and diamond card offers, the decision of an existing platinum cardholder to take up the diamond card will simply depend upon whether the perceived gain in social image exceeds the difference in price: customers who are informed about the changes in the eligibility criteria will take up if $S^{diamond}(1) - S^{platinum}_{info}(1) > p$ while customers who are not informed about these changes will take up if $S^{diamond}(1) - S^{platinum}_{no\ info}(1) > p$. Positional externalities due to income signaling motivations imply that $S^{platinum}_{info}(1) < S^{platinum}_{no\ info}(1)$. That is, in the presence of positional externalities, the social image associated with the card is reduced when lower-income individuals have access to the card, which leads to the following testable prediction.[31]

Prediction 3. *If positional externalities are present and the platinum credit card is a status good, the share of customers demanding an upgrade to the diamond card will be higher in the treatment group, where customers are informed that the income eligibility criterion for the platinum card has been reduced, than in the control group where no such information is provided.*

---

[31]For simplicity, we neglect the case where the customer would prefer to cancel the platinum card, due to its lower income signal, but would not consider it worthwhile to obtain the more expensive diamond card.

2.5.4 Results: Positional Externality

We begin by comparing raw take-up rates of the control and treatment groups from the positional externalities experiment in Figure 2.6. Demand for the diamond card increases by almost 19 percentage points, from 21.5% to 40% (p-value=0.069), when customers are informed that the platinum card is now available to a larger group of customers. Table 2.1, column (1) reports the corresponding OLS regression results. Table 2.1, column (2) shows that the results are nearly unchanged when we include baseline covariates. Just as predicted by our stylized theoretical framework and models of fashion cycles in consumption, we show that the (relative) demand for a status good depends upon who else consumes it. If lower-status consumers gain access to the good, they dilute the associated status symbol, causing higher-status consumers to flee the product in favor of more exclusive and expensive products.

It is worth noting that we find demand for the upgrade to the new status good despite the fact that customers were explicitly informed that the instrumental benefits of the platinum and diamond cards are *identical*. Bagwell and Bernheim (1996) suggest that, in many settings, the instrumental benefits that are usually bundled with the social signaling component of a status good might provide an important 'functional alibi' for purchasing a status good. Our results suggest that such a functional alibi may not always be necessary, at least when it comes to justifying the purchase to the marketer and to oneself. Another surprising aspect of our results is the relatively high baseline take-up (21%) of the diamond card in the *no-info* condition. This could be explained by the higher price of the diamond card implying higher status, even with the same income criterion. In addition, some customers might have

already been aware of the recently lowered criterion for the platinum card, implying that our information treatment was unnecessary for some customers.

Beyond providing evidence of positional externalities in the consumption of status goods, this exercise also serve as a robustness check that reinforces the conclusions of our first experiment. Note that customers in the positional externality treatment and control groups received the exact same offer, and calls differed only by whether customers were additionally informed about recent changes in the platinum card income requirements. Moreover, the scripts used in the positional externalities experiment explicitly state that the only difference between the platinum and diamond cards (aside from the income qualification criteria) is their design. Unlike in the first experiment, we thus avoid the possibility of offending participants in the control by denying them access to the status good, and still find significant demand for the status component of the card. The consistency of the results between the two experiments also makes it less likely that the results of the first experiment are explained by skepticism about the instrumental benefits in the control group.

2.6   Experiment 3: *Self and Social Image*

Thus far, we have provided evidence that social-image motives play an important role in the demand for a status good. But why do consumers value social image? One reason could be purely instrumental: having better social image might lead to greater professional opportunities or improved marriage-market prospects. Alternatively, social image could be valued for purely hedonic reasons: a person might care innately what others think of them.

While we suspect that both aspects of social image are relevant in practice, in this section we attempt to better understand the psychological – *and thus arguably hedonic* – determinants of the demand for status.

In particular, we test whether a person's self-image –how they view themselves– affects their demand for status goods. To do so, we conduct two experiments in different settings. In both, we experimentally manipulate self-esteem –one important dimension of self-image– and estimate its effect on the demand for a status good relative to a control product. If self-esteem affects the demand for status goods, this suggests that at least part of the demand for status is due to hedonic or psychological motives. In addition, since status goods themselves are a means to earn social image, the *direction* of the effect sheds light on whether social and self-image are substitutes or complements. This is an entirely open question in the literature on self and social signaling in economic behavior (Andreoni and Bernheim 2009a; Bénabou and Tirole 2006), with implications for policy tools aimed at strengthening or weakening the power of social norms.[32]

2.6.1   Self-Esteem Intervention: *Credit Card Customers*

Set-up and Experimental Protocol

The first self-esteem experiment uses a sample of 203 current gold card customers who had been identified by the bank as being eligible for an upgrade to the platinum card. These are

---

[32]On the one hand, it could be that having higher self-image increases one's demand for social admiration. Picture Muhammad Ali knocking out Sonny Liston, and then asking journalists to answer the question,"Who's the greatest?" (Hauser (1992), Pg. 79). Alternatively, it could be that individuals with high self-image feel little need to impress others, making self and social image substitutes.

customers who, at the time of the experiment, had a credit limit of at least Rp 20 million (US\$ 1,420), were current on their credit card payments, and were not employees of the bank.

These customers are assigned to one of four treatment conditions in a 2x2 cross-randomized design. The first randomization in this design determined whether customers were assigned to complete a self-affirmation intervention, taken from the psychology literature, designed to boost one's self-esteem, or a placebo exercise. The second randomization determined whether customers in the sample would then receive an offer to upgrade to the benefits of the platinum card as an add-on to their current gold card, or an offer to upgrade to the actual platinum card (the same offers as in the *benefits upgrade* and *platinum upgrade* treatment conditions in the first experiment). We include the benefits upgrade offer as one of the treatment arms in our design to rule out that the self-esteem intervention also increases demand for a good that does not confer social status.

The main outcome of interest in this experiment is whether receiving the self-esteem intervention affects take-up of the visible status good. If self and social image are complements, demand for the platinum upgrade should be higher among customers who receive the self-esteem intervention. If, on the other hand, self and social image are substitutes, demand for the platinum upgrade should be lower among customers who receive the self-esteem intervention.

Experimental Treatments

The self-affirmation exercise used in this experiment is adapted from the psychology literature (Steele 1988, Cohen et al. 2009, Hall et al. 2013). The exercise involves asking the respondent to reflect on a recent experience or achievement that made them feel proud. We show that this treatment delivers a boost to one's self-esteem, as measured using standard tests such as the Rosenberg (1965) scale. Following this literature, customers assigned to the *self-image treatment* group were asked to complete the following task before receiving an upgrade offer:

> *At [name of bank], we think it's important to understand our customers really well. So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please describe a specific incident in your life, something you did or achieved, that made you feel successful or proud of yourself? It could be from any aspect of your life, whether family related, education, or professional.*

Customers assigned to the *self-image control* group completed a placebo exercise, which asked participants to describe their media preferences and did not contain any statements or questions that might affect the respondent's self-image:

> *At [name of bank], we think it's important to understand our customers really well. So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please tell me which are your favorite TV channels and why? This would be a great help to us in understanding our clients media preferences.*

After completing one of these tasks, all customers received either an offer to upgrade to the platinum card, or an offer to upgrade to the platinum benefits package as an add-on to their current credit card. These offers were made using the same protocol and experimental scripts as in the first experiment, described in Section 2.3.2.

2.6.2 Self-Esteem Intervention: *mTurk Experiment*

Set-up and Experimental Protocol

The above experiment has a number of limitations. First, the bank permitted only a small sample of customers for this final experiment, thus limiting statistical power. Second, a phone call from ones bank is likely not the ideal format for the self-affirmation treatment, and some customers understandably prefer to skip the question (although they are still included in the results). Thus, we designed an additional experiment using the online crowdsourcing platform *mTurk*. This allows us to test for the substitutability of self and social image motivations in an alternative setting and population, where we can implement a parallel experimental design, but have tighter experimental control, and more statistical power (due to larger sample sizes and likely a stronger first stage).

The sample for the online experiment consists of 405 individuals who signed up to complete an incentivized task on the online platform *mTurk*. In the first part of the experiment, participants were randomly and individually assigned to one of two tasks, which mirrored the self-esteem intervention in the previous experiment. All participants received the same briefing and instructions which, in this case, were communicated using the online platform's interface, rather than a phone call.

In the second part of the experiment, all participants were then asked to make incentivized binary choices between the two types of gift certificates of different amounts, one for a standard brand and one for a luxury brand. This version of the standard *Becker-DeGroot-Marschak* multiple price list procedure allows us to elicit a truthful measure of the differential willingness to pay for a luxury brand gift card, compared to a non-luxury brand gift card.

The willingness to pay for the luxury gift card is our main outcome of interest in this experiment: If self and social image are complements, the self-esteem intervention should increase the willingness to pay for the luxury brand gift card. If, however, self and social image motives are substitutes, one would expect that the self-affirmation intervention reduces the demand for the luxury gift card.

Experimental Treatments

The experimental treatments are designed to mirror those in the credit card self-esteem experiment. Participants assigned to the *online self-esteem treatment* group were asked to write a paragraph about a recent experience or achievement that made them proud, using the following instructions:

> *Can you please describe an event that made you feel successful or proud of yourself? It could be from any aspect of your life, whether personal, social or family related, educational, or professional. Please be as specific as possible, and include as many details as possible. You should use all of the blank space below.*

Participants in the *online self-esteem control* group were asked to complete a placebo task analogous to that in the previous self-affirmation experiment:

Can you please tell the title and summarize the story of the last movie you have seen? Please be as specific as possible, and include as many details as possible. You should use all of the blank space below.

After completing one of these tasks, we measured participants' self-esteem, using the standard Rosenberg (1965) scale.[33] This allows us to verify that the treatment increases self-esteem as intended. The questionnaire consisted of a series of statements, such as *"On the whole, I am satisfied with myself"*, and asks respondents whether they strongly agree, agree, disagree, or strongly disagree with the statement. As reported below, we detect a meaningful and statistically significant increase in self-esteem as a result of the treatment.

Next, all participants were informed that they qualify to participate in a lottery in which they can win either a $500 gift certificate for a standard brand (*Old Navy*) or a $400 ($450, $500, $550 $600) gift certificate for a luxury brand (*Armani*). Participants were asked to make incentivized binary choices between the two types of gift certificates at different monetary values. The elicited willingness to pay for the different types of gift cards is the main outcome of interest which we use to test the complementarity of self and social image motivations in the demand for status goods.

Finally, participants were asked to rank the values they consider important in life (Steele and Liu, 1983), to test whether the self-affirmation treatment causes participants to reevaluate the importance of different aspects of their life, such as family, religion, work or financial success. We detected no such effects, suggesting that any impacts of self-affirmation on

---

[33]The survey instrument used is available in the Supplementary Appendix

consumption were not driven by changes in values.[34]

2.6.3   Testable Predictions

To derive testable predictions, we extend our standard framework to allow for an interaction between self and social image. Specifically, we extend the agent's utility to include the self-image term ω_i:

$$I_i(a_i, \omega_i) = \lambda_i(\omega_i)\mathrm{Pr}_{-i}\left(y_i = h \mid a_i\right) \tag{2.2}$$

This approach allows the agent's valuation of social image to depend upon their self-image through the weight $\lambda_i(\omega_i)$. We make the simplifying assumption that self-image is unidimensional, while in reality, people might have higher self-image with respect to some aspects of their lives than others. In our experiments, we additionally assume that a boost to one's self-esteem is synonymous with an improvement in self-image.

In both self-affirmation experiments, we are interested in testing the sign of $\lambda_i'(\omega_i)$. A positive (negative) derivative would be evidence that self and social image are complements (substitutes). A priori, we do not expect the self-affirmation intervention to have any systematic effect on the purchase of non-status goods such as the benefits upgrade, since they do not involve social image.

---

[34]We asked subjects to rank eight aspects (family, friends, leisure time, financial success, health, politics, work, and religion) from most important to less important. We test for the null hypothesis of no effect of the self-affirmation treatment for each of these aspects. Since the outcome variable is ordinal (a rank from 1 to 8), we use a permutation test based on Volfovsky et al. (2015). The p-value of a joint test of no effect of the self-affirmation treatment for all aspects is 0.62. Neither does any individual aspect show significant effects.

Prediction 4. *The self-affirmation intervention will reduce the demand for the platinum card if self-image and social image are substitutes ($\lambda'_i(\omega_i) < 0$). The self-affirmation intervention will increase demand for the platinum card if self-image and social image are complements ($\lambda'_i(\omega_i) > 0$). The self-affirmation intervention will have no systematic effect on demand for non-status goods such as the benefits upgrade.*

2.6.4 Results: Self and Social Image

Treatment Effects: *Credit Card Customers*

Figure 2.7 presents the raw take-up rates by treatment, separately for the status good (platinum card) and placebo good (benefits upgrade) offers. The take-up rate for the benefits upgrade does not respond to the self-affirmation treatment, although limited precision means we cannot rule out moderate effects. In contrast, the self-affirmation treatment reduces take-up of the platinum card by approximately 15 percentage points (from 33% to 18%). Although this difference is economically large, it is not statistically significant (permutation test p-value=0.192). Table 2.4 reports these results in regression format, including caller fixed effects and baseline covariates.

Overall, these results provide suggestive evidence that self and social image are substitutes, rather than complements. However, due to the limited sample size available for this experiment, these results are not conclusive. To provide additional evidence on the relationship between self and social image motivations, we therefore turn to the results of the complementary experiment, which we implemented using a separate population on the online platform *mTurk*.

Treatment Effects: *mTurk Experiment*

We present the results of the *mTurk* experiment in Table 2.5. In Table 2.5, column (1), we first report the effect of the self-esteem treatment on subjects' self-esteem, as measured using the Rosenberg (1965) scale. The results confirm that the self-esteem treatment was indeed successful at delivering a boost to participants' self-esteem. On average, participants in the self-image treatment group scored 1.22 points, or 0.17 standard deviations, higher on the self-esteem measure than participants in the control group (statistically significant at 10%).

In Table 2.5, columns (2) to (6), we report the effects of the self-esteem treatment on demand for the luxury brand gift certificate. We find that the self-esteem treatment has a negative impact on the proportion of subjects who prefer the luxury brand for all values (the difference is statistically significant for 3 out 5 prices). Figure 2.8 presents the cumulative distribution for the willingness to pay for the *Armani* gift card relative to the *Old Navy* gift card for both groups, which confirms our result that the self-affirmation treatment has a negative effect on the willingness to pay for the *Armani* gift card. Adding baseline covariates again yields very similar results (Table 2.5, panel ii).

Discussion and Interpretation

This section has provided evidence that higher self-esteem causally reduces the demand for status goods. Our interpretation of this result is that a higher *self* image reduces individuals' demand for *social* image. That is, self and social image are substitutes. To our knowledge, this is the first evidence on the relationship between self and social image. It predicts that social signaling behavior will be particularly strong among those with low self-esteem, and

that such individuals may thus be more likely to conform to social norms. When these norms are judged by policy makers to be 'negative', such as social stigma from studying hard in low-income minority schools (Bursztyn and Jensen, 2015a), policy tools to build self-esteem or a sense of self-worth might be effective in weakening the power of the social norm, as in Cohen et al. (2009). Conversely, higher self-esteem might reduce compliance with 'positive' social norms, such as those encouraging charitable donations (DellaVigna et al., 2012a) or voting (DellaVigna et al., 2017a).

An additional mechanism may be at play in these experiments. In particular, models of identity imply that individuals prefer to take actions consistent with their self-image or identity (see, for example, Akerlof and Kranton 2000b). Under such theories, high-income individuals might purchase status goods simply because it is consistent with their high self-image. Boosting their self-esteem might further increase their demand for such status goods. Yet, we observe a reduction in demand for status goods from boosting self-esteem, suggesting that any such effect in our experiment is overpowered by the strong substitutability of self and social image.

2.7 Conclusion

This paper provides the first field experimental evidence of the existence of status goods. In particular, we show that the status aspect of premium credit cards – due to their potential to signal income – is an important driver of the demand for the product, over and above its instrumental benefits. Our experiments also identify a positional externality associated

with the consumption of these status goods, thus confirming a key prediction of theories of status goods. We also show that higher self-esteem causally reduces demand for status goods, suggesting that self and social image are substitutes in the context we study, at least at the margin.

We believe this work can be usefully extended in several directions. First, more work on the overall economic importance and welfare consequences of status goods would be valuable. Second, understanding reference groups is a promising avenue: whom do individuals want to impress, and whom do they compare themselves to? Third, while we provide evidence that self and social image are substitutes in our context, at least in the short run, it will be important to understand whether this is true in other contexts and along other dimensions of image. Finally, we believe that understanding the effect of self-esteem on economic choices is a promising avenue for future work, especially in settings where self-esteem may be particularly low, such as in populations facing poverty, low social status, and negative stereotypes.

2.8 Figures and Tables

Figure 2.1: **The Credit Cards**

Notes: The figure shows the design of the *platinum*, *gold* and *basic* credit cards used in the experiments (from left to right).

Figure 2.2: **Experiment 1: Demand for Status**
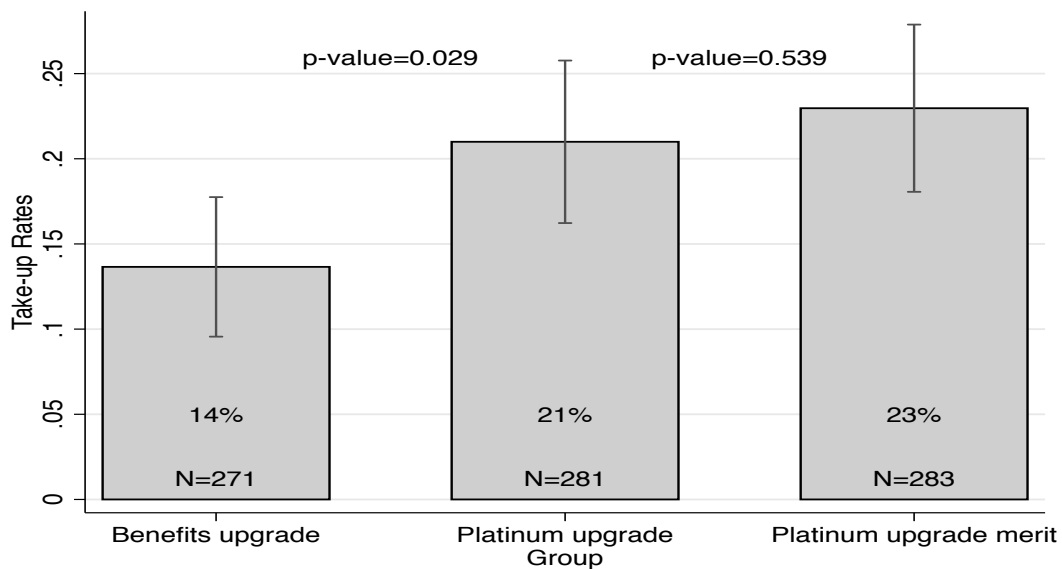
Notes: This figure presents the mean (and 95% confidence interval) of take-up rates for the *benefits upgrade*, *platinum upgrade*, and *platinum upgrade merit* groups. The p-values are based on permutation tests.

Figure 2.3: **Benchmarking**
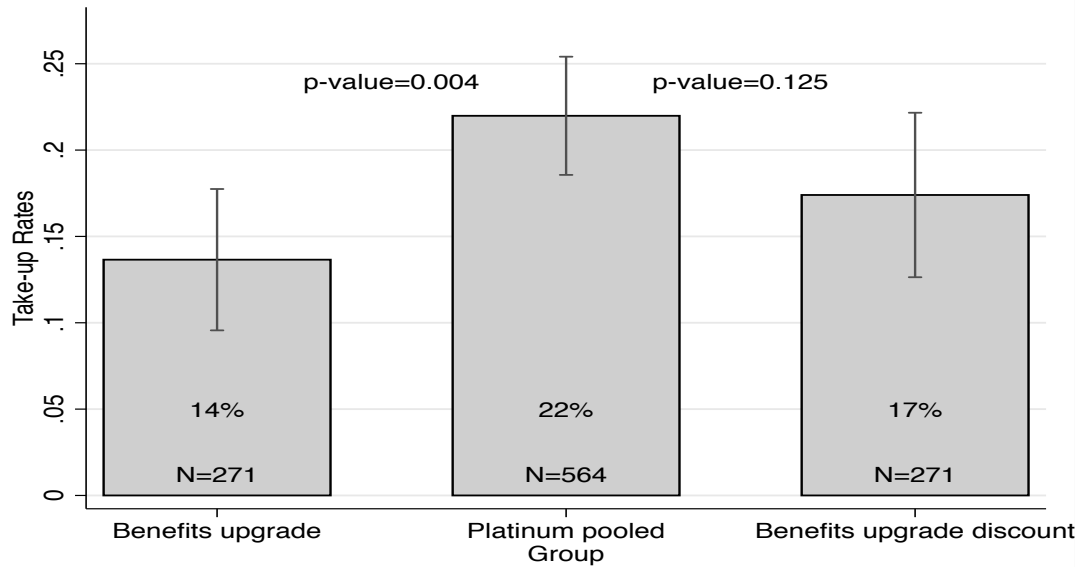
Notes: This figure presents the mean (and 95% confidence interval) of take-up rates for the *benefits upgrade* and *platinum pooled* groups. We also present the take-up rate for the benefits upgrade with a 25% discount in the annual fee. This take-up rate is based on the *benefits upgrade* sample. As explained in Section 2.3.3, we called again customers that declined the benefits upgrade offer in the first call after hearing the price details of the offer, and offered them the benefits upgrade at a 25% price discount. For this case, we calculate the standard error using bootstrap. For the *benefits upgrade* versus *platinum pooled* comparison, the p-value is based on a permutation test.

Figure 2.4: **Experiment 1: Demand for Status - Income Heterogeneity**
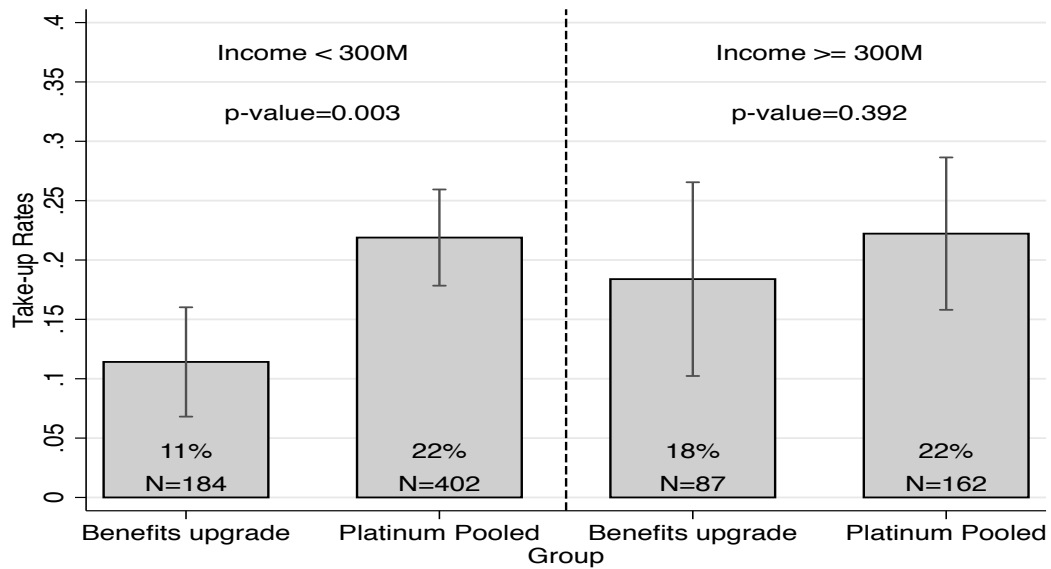
Notes: This figure presents the mean (and 95% confidence interval) of take-up rates for the *benefits upgrade* and *platinum pooled* groups separately for customers with income lower than Rp 300 million and customers with income greater or equal than Rp 300 million. The p-values are based on permutation tests.

Figure 2.5: **Transaction data: Share of Visible Transactions**
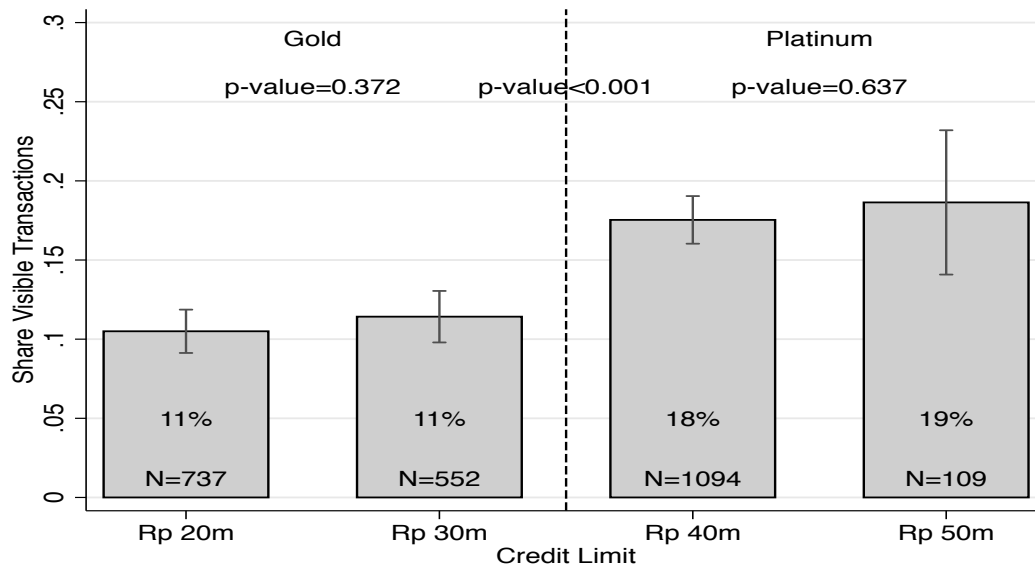
Notes: This figure presents the share of visible transactions (and 95% confidence intervals) for customers with different credit card limits.

Figure 2.6: **Experiment 2: Positional Externalities**
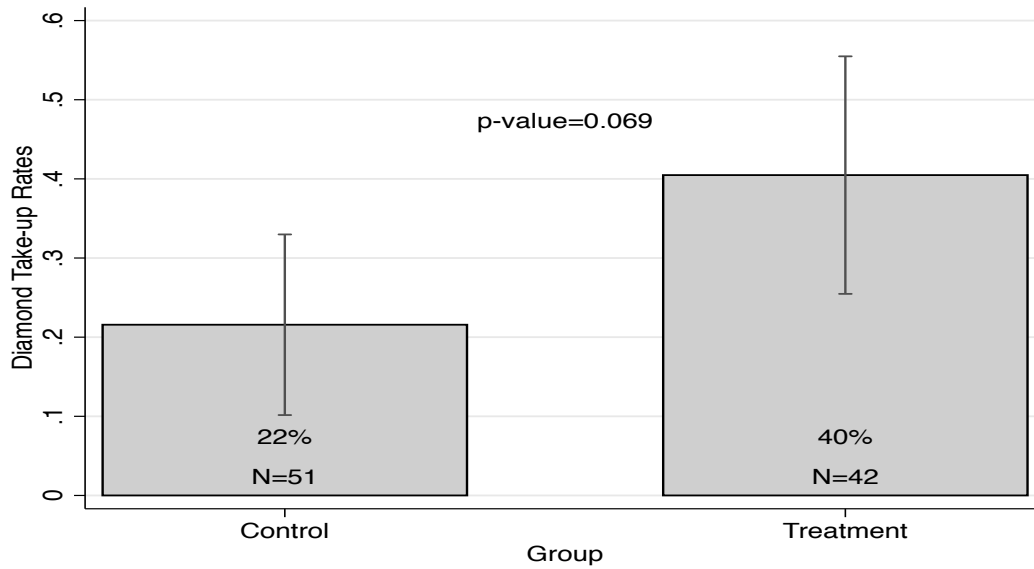
Notes: This figure presents the mean (and 95% confidence interval) of take-up rates for the control and treatment groups. The p-value is based on a permutation test.

Figure 2.7: **Experiment 3: Self and Social Image**
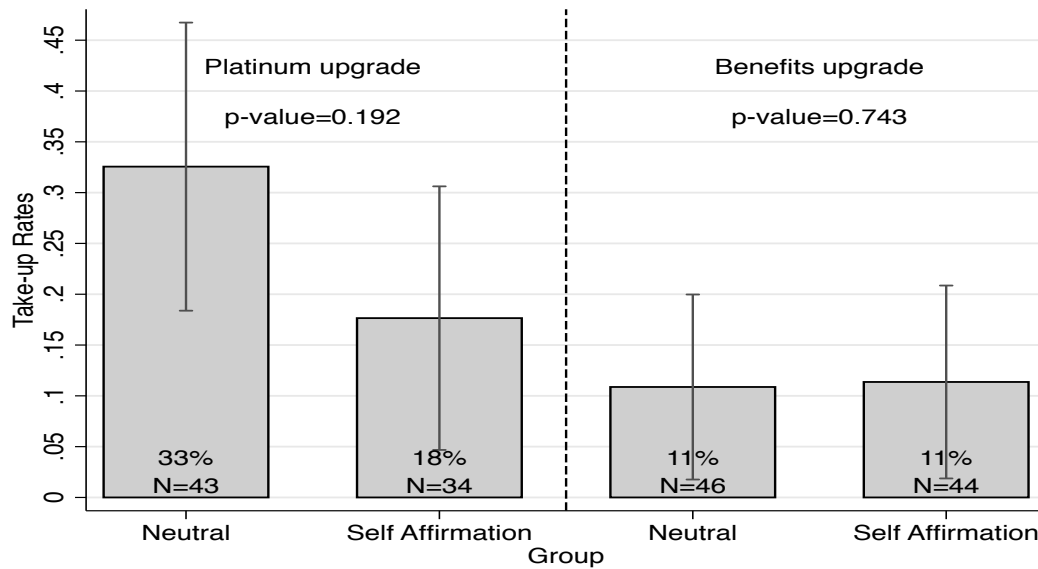
Notes: This figure presents the mean (and 95% confidence interval) of take-up rates for the control and self-affirmation groups, separately for the platinum upgrade and for the benefits upgrade offers. The p-values are based on permutation tests.

Figure 2.8: **mTurk Experiment: Self and Social Image**
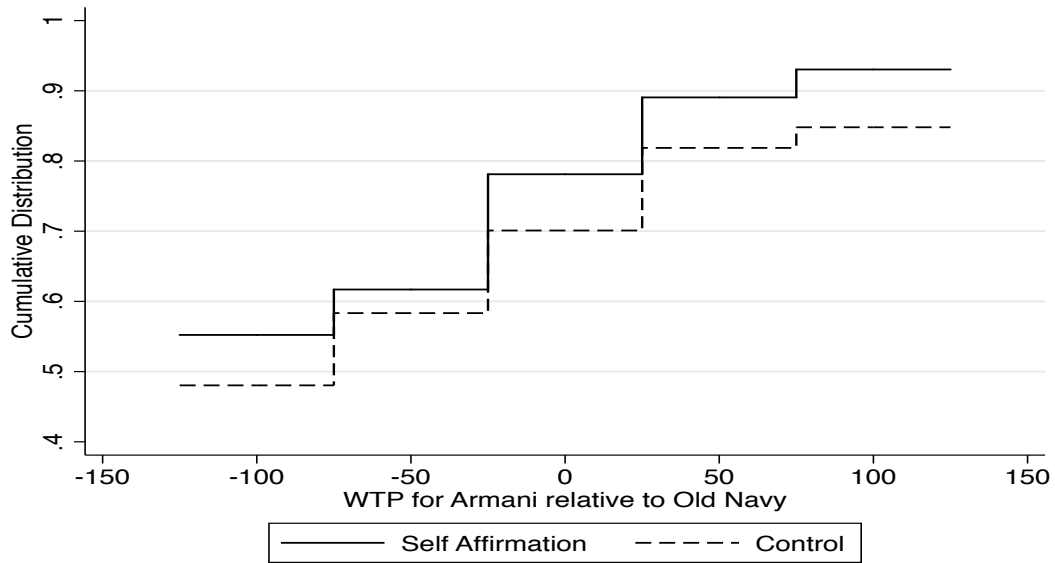
Notes: cumulative distribution of the willingness to pay to receive a luxury brand (*Armani*) gift card instead of a standard brand (*Old Navy*) gift card for the control and the self-affirmation groups.

Table 2.1: **Demand for Status (Experiment 1)**

	(1)	(2)	(3)	(4)	(5)	(6)
Platinum	0.073**	0.072**	0.082***			
	[0.032]	[0.032]	[0.027]			
	(0.029)	(0.033)	(0.005)			
Platinum × 1$\{y_i < \text{cutoff}\}$ (a)				0.105***	0.094***	0.105***
				[0.032]	[0.028]	[0.032]
				(0.004)	(0.001)	(0.003)
Platinum × 1$\{y_i < \text{cutoff}\}$ (a)				0.031	0.013	0.035
				[0.049]	[0.084]	[0.048]
				(0.543)	(0.887)	(0.534)
p-value ((a) = (b))				0.222	0.364	0.280
Mean	0.137	0.137	0.137	0.137	0.137	0.105
(benefits upgrade)	[0.021]	[0.021]	[0.021]	[0.021]	[0.021]	[0.020]
Income Cutoff	-	-	-	300m	500m	300m
Sample	*Platinum upgrade* and *benefits upgrade* conditions	*Platinum upgrade* and *benefits upgrade* conditions	All	All	All	Exclude income ≥ 500m
Controls	No	Yes	Yes	Yes	Yes	Yes
Sample size	552	552	835	835	835	704
R2	0.009	0.086	0.070	0.071	0.071	0.043

Notes: Column 1 presents the results of a regression of a dummy variable equal to one if the client accepted the offer on a dummy for platinum treatment using customers in the *platinum upgrade* and *benefits upgrade* conditions. The regression presented in column 2 includes strata dummies, credit limit, female, muslim, and Jakarta as covariates. The regression presented in column 3 pools customers in the *platinum upgrade* and *platinum upgrade merit* conditions as the platinum group. The regressions presented in columns 4 and 5 include interactions of the platinum treatment dummy with a dummy if income is lower than the cutoff and another dummy if income is higher or equal than the cutoff. In column 4 the cutoff is defined as 300M Rp while in column 5 it is defined as 500M Rp. The regression presented in column 6 replicates column 4 but excludes clients with income greater or equal than 500M. Robust standard errors in brackets. Permutation test p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.2: **Effects of Platinum Card on Credit Card Usage (Transaction Data)**

	(1)	(2)	(3)	(4)	(5)	(6)
Gold (30M CL) - Gold (20M CL)	0.009 [0.011]	0.008 [0.011]	-0.010 [0.009]	-0.009 [0.009]	0.011 [0.018]	0.012 [0.018]
Platinum (40M CL) - Gold (30M CL)	0.061*** [0.011]	0.053*** [0.012]	-0.005 [0.007]	0.000 [0.008]	-0.090*** [0.017]	-0.095*** [0.018]
Platinum (50M CL) - Platinum (40M CL)	0.011 [0.024]	0.015 [0.025]	0.009 [0.013]	0.007 [0.013]	-0.023 [0.033]	-0.017 [0.032]
Mean (Gold (CL 20M))	0.105 [0.007]		0.054 [0.006]		0.673 [0.012]	
Controls	No	Yes	No	Yes	No	Yes
Number of clients:						
Gold (20M CL)	737	737	737	737	737	737
Gold (30M CL)	552	552	552	552	552	552
Platinum (40M CL)	1094	1094	1094	1094	1094	1094
Platinum (50M CL)	109	109	109	109	109	109
p-value (a)=(b)	0.008	0.020	0.708	0.549	0.002	0.001
p-value (a)=(c)	0.946	0.779	0.223	0.321	0.363	0.440
p-value (b)=(c)	0.085	0.195	0.391	0.665	0.094	0.049

Notes: Column 1 reports raw comparisons of share of visible transactions for clients with different credit limits/type of card. Column 2 reports comparisons controlling for income, female dummy, muslim dummy, Jakarta dummy, and age. Columns 3 and 4 report results for online transactions, while columns 5 and 6 report results for share of retail transactions. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.3: **Positional Externalities (Experiment 2)**

	(1)	(2)
Information treatment	0.189**	0.206**
	[0.096]	[0.097]
	(0.069)	(0.039)
Mean (no information)	0.216	0.216
	[0.058]	[0.058]
Controls	No	Yes
Sample size	93	93
R2	0.042	0.143

Notes: Column 1 presents the results of a regression of a dummy variable equal to one if the client accepted to get on the invite list for the diamond card on a dummy for information treatment. The regression presented in column 2 includes income, credit limit, female, muslim, and Jakarta as covariates. Robust standard errors in brackets. Permutation test p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.4: **Self and Social Image - Credit Card (Experiment 3)**

	Platinum upgrade		Benefits upgrade	
	(1)	(2)	(3)	(4)
Self-Affirmation	-0.1491	-0.1548	0.0049	-0.0350
	[0.0981]	[0.1060]	[0.0670]	[0.0615]
	(0.192)	(0.142)	(0.743)	(0.604)
Mean (neutral)	0.326		0.109	
	[0.072]		[0.046]	
Controls	No	Yes	No	Yes
N	77	76	90	90
R2	0.0285	0.1811	0.0001	0.2380

Notes: Column 1 presents the results of a regression of a dummy variable equal to one if the client accepted the platinum upgrade offer on a dummy for self-affirmation treatment. The regression presented in column 2 includes income, credit limit, female, muslim, and Jakarta as covariates. The regressions presented in columns 3 and 4 present results using a dummy variable equal to one if the client accepted the benefits upgrade offer. Robust standard errors in brackets. Permutation test p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.5: **Self and Social Image - *Armani* Gift Cards (mTurk Experiment)**

| | Rosenberg Self-Esteem Score | Prefer \$ ___ *Armani* to \$500 *Old Navy* | | | | |
		400	450	500	550	600
	(1)	(2)	(3)	(4)	(5)	(6)
		Panel i: without controls				
Self-Affirmation	1.2214*	-0.0823***	-0.0719**	-0.0801*	-0.0336	-0.0718
	[0.7023]	[0.0310]	[0.0349]	[0.0434]	[0.0488]	[0.0497]
	(0.080)	(0.010)	(0.048)	(0.052)	(0.551)	(0.141)
Mean (neutral)	19.8333	0.1520	0.1814	0.2990	0.4167	0.5196
	[0.5076]	[0.0252]	[0.0270]	[0.0321]	[0.0346]	[0.0351]
Sample size	405	405	405	405	405	405
		Panel ii: with controls				
Self-Affirmation	1.2318*	-0.0829***	-0.0728**	-0.0805*	-0.0319	-0.0680
	[0.6890]	[0.0309]	[0.0349]	[0.0430]	[0.0469]	[0.0489]
	(0.090)	(0.008)	(0.041)	(0.061)	(0.509)	(0.174)
Mean (neutral)	19.8333	0.1520	0.1814	0.2990	0.4167	0.5196
	[0.5076]	[0.0252]	[0.0270]	[0.0321]	[0.0346]	[0.0351]
Sample size	405	405	405	405	405	405

Notes: Column 1 presents results of a regression of Rosenberg self-esteem Score on a dummy for self-affirmation treatment. Columns 2 to 6 present results of a regression of a dummy equal to one if the subject chose the *Armani* rather than the *Old Navy* gift card on a dummy for self-affirmation treatment for the corresponding offer. Panel i presents regressions without additional controls, while Panel ii presents results including race, gender, age, marital status, education and income as covariates. Robust standard errors in brackets. Permutation test p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

117

Chapter 3

From Extreme to Mainstream: How

Social Norms Unravel

3.1 Introduction

Social norms are an important element of any society: some behaviors and opinions are socially desirable while others are stigmatized. There is growing evidence that individuals care to a large extent about how they are perceived by others and that such concerns might affect important decisions in a variety of settings, from charitable donations (DellaVigna et al., 2012b), to schooling choices (Bursztyn and Jensen, 2015b), to political behavior (DellaVigna et al., 2017b; Enikolopov et al., 2017b; Perez-Truglia and Cruces, Forthcomingb). Moreover, these social image concerns matter both in interactions with other people from the same social group (Bursztyn and Jensen, 2015b) and in interactions with strangers, such as surveyors and solicitors (DellaVigna et al., 2012b, 2017b). In particular, even individuals with very strongly-held political views might avoid publicly expressing them if they believe their opinion is not popular in their social environment (Bursztyn et al., 2016).

A recent literature has documented the persistence of cultural traits and norms over long periods of time (Voigtländer and Voth, 2012; Fernández, 2007; Giuliano, 2007; Algan and Cahuc, 2010; Alesina et al., 2013). However, little is known about what factors might lead long-standing social norms to *change*, or even more so, to change *quickly*. In this paper, we argue that aggregators of private opinions in a society, such as elections, might lead to updates in individuals' perceptions of what people around them think, and thus induce fast changes in the social acceptability of holding certain opinions and in the likelihood that these opinions are publicly expressed.

Consider the support for the communist regime in the Soviet Union in the late 1980s. Kuran (1991) argues that many individuals opposed the regime but believed others supported it. In that environment, a referendum on the regime would have quickly updated people's opinions about the views of others. Incorrect beliefs about the opinions of others are not restricted to totalitarian regimes, where expressing personal views is often risky. In fact, as we argue below, if most individuals assume that a specific opinion is stigmatized, the stigma might be sustained in equilibrium.[1] In this paper, we formalize this idea and test it using

---

[1] This phenomenon is known in social psychology as "pluralistic ignorance" (Katz and Allport, 1931), where privately most people reject a view, but incorrectly believe that most other people accept it, and therefore end up acting accordingly. For example, in 1968 most white Americans substantially over-estimated the

experiments with real stakes. In our main experiment, we show that Donald Trump's rise in popularity and eventual victory during the 2016 U.S. presidential campaign causally increased individuals' perception of the social acceptability of holding strong anti-immigration (or xenophobic) views and their willingness to publicly express them.

We present a simple model of strategic communication between individuals who can hold one of the two mutually exclusive convictions (e.g., xenophobic and tolerant). Each *sender* delivers a message (states his view) to a *recipient*, who in turn uses this message to make a Bayesian update on the sender's true conviction (e.g., whether or not the sender is xenophobic). In our setting, as in DellaVigna et al. (2017b), lying (or expressing a conviction that is not own's own) is costly, so messages have some credibility; at the same time, citizens can pay this cost and lie, so communication need not be truthful. Importantly, the sender has social image concerns: he values other people's perception of his type. Thus, for example, if he thinks most other people are tolerant, he might want to hide his xenophobic views and pay the cost of lying; conversely, if he believes that most other people are xenophobic, he would not only feel free to state such a view if he shares it, but may express xenophobia even if he is tolerant. Thus, senders have incentives to 'pander to the majority'. In the extreme cases we may have a 'political correctness' equilibrium, wherein all senders state the same view regardless of their true conviction, thus preventing learning about true convictions. Receivers are Bayesian, and their perception of a sender of a message depends on the strategy that senders use. In the example above, if xenophobic individuals say that they are tolerant because they believe that most are tolerant, then anyone who expresses xenophobic opinions must indeed be xenophobic. At the same time, if most receivers are believed to be xenophobic so that even some tolerant people express such views, then a person stating xenophobic views is not necessarily a xenophobe, and thus will be judged less harshly by tolerant people.

Our model predicts, in particular, that a positive update in beliefs about the share of xenophobes in a society makes xenophobic messages more likely. Our framework thus illustrates how a shock to the informa-

---

support among other whites for racial segregation (O'Gorman, 1975). A related concept is that of "preference falsification" (Kuran, 1995): people's stated, public preferences are influenced by social acceptability, and might be different from their true, private preference.

tion possessed by individuals, such as an election where this particular issue (tolerance vs. xenophobia) is salient, can rapidly change the social norm in communication behavior: what was unacceptable and rarely, if ever, spoken, could become acceptable and normal in a matter of weeks, if not days.

Quite interestingly, the logic of the paper suggests that information aggregation can make an 'extreme' topic 'mainstream', but not the other way around: such a shock cannot make a mainstream topic extreme. This is easy to see: if a topic is mainstream and socially acceptable, individuals know how widely and by whom each opinion is shared, in which case information aggregation is unlikely to reveal new information. Of course, if underlying fundamentals change – for example, more people become radicalized, or people become more judgmental of others – then a certain viewpoint could move from mainstream to extreme. However, since this involves changes in fundamentals, this is likely to be a slow and gradual process.

We test the model on both the senders' and recipients' sides using real-stakes experiments on Amazon Mechanical Turk (MTurk) based on two recent votes. On the sender's side, we focus on the 2016 U.S. presidential election. Throughout his campaign, Donald Trump proposed, among other things, the construction of a wall separating the US and Mexico and a ban on Muslims from entering the U.S. His popularity might thus send an informative signal to the population about the number of people who sympathize with these proposals and thus about those who hold xenophobic views. As the experimental evidence we provide suggests, this signal is likely leading to a positive update in beliefs about the share of xenophobes in the country. As a result, Donald Trump's electoral success potentially caused a shift in social norms regarding expressed views on immigrants.

More specifically, in the two weeks before the election which took place on November 8, 2016, participants were offered a bonus cash reward if they authorized the researchers to make a donation to a strongly anti-immigration organization on their behalf. Accepting the offer is therefore a *profitable* xenophobic action. At baseline, participants who randomly expected their decision to be potentially observed by and discussed with a surveyor in a future interaction (the "public" condition) were significantly more likely to forgo the donation bonus payment than those who expected their choice to be entirely anonymous (the "private" condition). This suggests the presence of social stigma associated with the action. Before making the donation decision,

121

a random subset of participants received information that positively updated their perceptions of Trump's popularity in their home state on the eve of the election. We first show that this information positively updates their beliefs about the local popularity of holding xenophobic views. We then show that, for these participants, the wedge in the likelihood of undertaking the xenophobic action in private and public disappeared. This difference with respect to the baseline condition was driven entirely by an increase in the donation rate in the public condition, with no change in the private condition.

We also exploited the "natural experiment" of Trump's unexpected victory as an alternative "treatment" that could generate increases in the willingness to publicly express xenophobic views. We replicate the experimental intervention shortly after the election, restricting the design to the baseline condition with no initial information on Trump's popularity. We find that after the election, the wedge between private and public donation rates disappeared, even in the absence of the experimental information intervention. Again, this difference was entirely driven by an increase in the public donation rate; the private donation rate remained unchanged from the pre-election survey. Our results suggest that Donald Trump's rise in popularity did not casually make these participants more xenophobic, but instead made the already more-intolerant ones more comfortable about publicly expressing their views. We also discuss suggestive evidence of the precise mechanisms driving our findings.

A second, auxiliary experiment studies the receiver's side. We exploit the general lack of knowledge among MTurk users in Democrat-leaning U.S. states about the 2009 Swiss referendum that banned the construction of minarets in that country. We consider this an ideal setting for two reasons. First, the topic of the referendum was closely aligned to that of our main experiment. Second, the fact that most participants were unaware not only of the ban, but also of the fact that the majority of the Swiss population supported it, allows us to manipulate their beliefs about the public support for the ban.

After stating their personal views on whether the construction of minarets should be banned, participants played a dictator game with a real, anonymous Swiss person. We randomize what information was given about the Swiss person to evaluate via revealed preference how different information affected the participant's opinion of that person. We focus on participants who reported being *against* banning the construction of

122

minarets. In the first treatment group, individuals who were randomly told that the Swiss person is in favor of the ban significantly reduced their donation levels in the dictator game, when compared to a control group where that piece of information was not provided. In the second treatment group, participants were additionally informed that 57% of Swiss respondents also supported the ban, which positively updated participants' priors about the share of Swiss individuals in support of the ban. Note that this was the actual support level for the ban in the 2009 referendum. Thus, when informed that the Swiss person is surrounded by other people who support the ban, participants judged that person less negatively for *expressing* that view, understanding that he might have done so to pander to the majority: donation amounts went up significantly compared to the first treatment group.

A potential alternative interpretation is that participants judge the Swiss person less negatively for *holding* his opinion when the majority of Swiss support the ban, *regardless of whether his support was expressed in public*. For example, it could be that participants feel that they cannot blame a person for privately holding a view if that person is surrounded by many other people who also hold that view and who could also have influenced the private opinions of that person. To directly rule out this alternative interpretation and confirm that our findings are driven by the anticipation by participants that the Swiss person might have had strategic reasons to publicly express an intolerant view, we also ran an additional experiment emphasizing that the Swiss person's opinion was expressed *anonymously*. In that treatment, informing participants of a higher level of public support for the ban in Switzerland does *not* increase donation amounts.

Our results suggest that Trump's rise in popularity and eventual electoral victory might have already casually changed social norms regarding the expression of xenophobic views in the U.S. Though we detect no changes in *privately-held* views, we believe the findings on public expression are of great policy relevance. For example, increases in public expression of anti-foreign sentiment might also lead to more frequent acts of hate crime against foreign minorities. Indeed, recent reports indicate a steep increase in the number of hate crimes against these groups throughout the campaign period, and especially after Donald Trump's election. For example, a recent report by the Center for the Study of Hate and Extremism indicates that across eight

major metropolitan areas in the U.S., the number of hate crimes increased by more than 20% in 2016 (Center for the Study of Hate and Extremism, 2017). This increase is significantly larger in both absolute and relative terms than any year-to-year increase in hate crimes in these cities since 2010. Increases in public expression of hate against some minorities might also spur increases in expression of hate against other minorities. For example, the same report shows that hate crimes against blacks and Jews in New York City also increased substantially between March 2016 and March 2017. More common public expression of certain views might also facilitate coordination for large-scale actions, such as demonstrations and movements, and recent work has provided evidence that such demonstrations and movements might affect many important outcomes, from election results (Madestam et al., 2013) to the stock market valuation of different firms (Acemoglu et al., Forthcoming). In addition, reductions in the stigma associated with holding previously-extreme views might lead to shifts in the language used in and reported by the popular media, and might also reduce the stigma associated with consuming and discussing certain news sources on the far side of the political spectrum. For example, the news website *Breitbart* has more than doubled its share of general news audience between the end of 2014 and July 2016, reaching 9% of the market – 18 million visitors – that month.[2] Increases in public expression of such views can thus lead to increases in individuals' overall exposure to them, and more exposure might eventually lead to changes in privately-held views, via persuasion or simple conformism.

Our results contribute to a growing literature that examines the impacts of political institutions on social norms and culture more generally. This literature typically studies the long-run impact of political institutions (e.g., Lowes et al., Forthcoming); here we show that changes on the political side can lead to fast reversals in social norms. Our paper also adds to a recent theoretical literature on social norms (e.g., Benabou and Tirole, 2011; Acemoglu and Jackson, 2017) by studying how new information may lead to the unraveling of such norms. Our findings speak to a cross-disciplinary literature on the consequences of political action, both theoretical (e.g., Lohmann, 1993) and empirical (e.g., Madestam et al., 2013). Methodologically, this paper also relates to a literature on the measurement of sensitive attitudes, which includes approaches such as the "randomized response technique" (Warner, 1965), the "list experiment" (Raghavarao and Federer, 1979),

---

[2]See https://fivethirtyeight.com/features/trump-made-breitbart-great-again/.

the "endorsement experiment" (Sniderman and Piazza, 1993), and the "lost letter technique" (Milgram, 1977). In our study, the private donation decision provides a measure based on revealed preference, and where concerns about social desirability are minimized due to anonymity. Conceptually, our main experiment is novel since it exploits both experimental and natural occurring variation as two alternative approaches. Randomized informational interventions are often subject to the criticism of not being entirely natural, while before/after designs based on natural variation suffer from lack of true randomization. In our setting, the two approaches yielded very similar results, which provides extra assurance of the validity of our tests of the model.

Theoretically, the two most important precursors to our paper are Bernheim (1994) and Morris (2001). Bernheim (1994) studies the behavior of individuals with social concerns and predicts the emergence of social norms. People with similar preferences will adhere strictly to such norms, while people with extreme tastes will choose not to do so. The possibility of complete pooling and the resulting non-transmission of information is suggested in Morris (2001), where an advisor who is afraid of being perceived as biased ultimately avoids giving advice in all states of the world. Our model can be viewed as combining and simplifying the two approaches to get a model with straightforward comparative statics results that highlight the role of individual beliefs and information shocks. One can loosely view the model as adapting the binary message structure of Morris (2001) to the social communication setting of Bernheim (1994).

Our work also relates to existing papers studying the economic consequences of conformity. Prendergast (1993) identifies rational incentives for managers to conform to supervisors' opinions in order to appear competent, which in turn hampers information transmission. Andreoni James and Siegenthaler (2017) study 'conformity traps,' situations where groups of individuals fail to coordinate on a beneficial action due to individual incentives to conform to the predominant and inefficient behavior. In a laboratory experiment they find, in particular, that opinion polls can facilitate changes of norms that benefit the group. Their setting, however, is one of full information, and thus opinion polls facilitate switching from one equilibrium to another. Our model has incomplete information and features a unique equilibrium, and elections can change the beliefs about the distribution of other people's opinions (though we do not take a position on

whether overcoming conformity is necessarily socially beneficial). In a different setting, Kets and Sandroni (2016) study the trade-off between the performance of conforming versus diverse groups of individuals.

The remainder of this paper proceeds as follows. We introduce a simple framework formalizing our argument in Section 2. In Section 3, we present the design and results from the main experiment studying the 2016 U.S. presidential elections. In Section 4, we present the design and results from the auxiliary dictator game experiment. Section 5 concludes.

3.2   Theoretical Framework

3.2.1   A model of communication

There is a continuum of citizens.[3] Citizens can hold one of two mutually exclusive convictions, A or B; we sometimes call it their type and write $t_i \in \{A, B\}$ for citizen i. (One can think of these as 'beliefs' in colloquial sense, but we reserve the word 'beliefs' for the game-theoretical notion.) The share of citizens holding conviction A is p, so $\Pr(t_i = A) = p$. We do not assume that p is known, and instead allow citizens to hold an incorrect belief about the share of citizens with conviction A, which we denote by q. To avoid dealing with higher-order beliefs, we assume that q is common knowledge.

The citizens are paired with one another; within each pair, there is a sender i and recipient j. Sender i sends message m_i, and we assume that the message space is binary, so message $m_i \in \{A, B\}$. Sending a message m_i may involve the following costs for citizen i. First, sending a message $m_i \neq t_i$ incurs a cost of lying l_i; this cost is assumed to be distributed uniformly on $[0, 1]$ (this can be thought of as normalization), and independently of t_i (i.e., citizens holding either conviction are equally averse to lying).[4] Second, the sender enjoys a benefit proportional (with intensity factor denoted by a) to his belief that the recipient

---

[3]We could instead assume communication between two individuals. The results would not change at all.

[4]Thus, in our model, communication is not 'cheap talk', as in Crawford and Sobel (1982). See Kartik et al. (2007) and Kartik (2009) for models of strategic communication with lying costs in an uninformed principal – informed agent framework. The cost of lying could also be interpreted as the (lost) intrinsic utility obtained from expressing one's true own conviction.

approves his type. Consequently, the expected utility of a citizen i with two-dimensional type (t_i, l_i) from sending message m_i to recipient j is given by

$$U_i(m_i) = -l_i \mathbf{I}_{\{m_i \neq t_i\}} + aq \Pr_j(t_i = A \mid m_i) + a(1-q)(1 - \Pr_j(t_i = A \mid m_i)), \qquad (3.1)$$

where $\Pr_j(t_i = A \mid m_i)$ is the recipient's posterior that the sender is type A conditional on the message he sent, m_i.

In this game, we are interested in Perfect Bayesian Equilibria, which furthermore satisfy the D1 criterion (Cho and Kreps, 1987).

3.2.2 Analysis

We start our analysis with the case $q > \frac{1}{2}$, which means that types $t_i = A$ are perceived to be more numerous. Then senders with the same type $t_i = A$ communicate truthfully. Indeed, by sending message $m_i = B$ they would incur the cost of lying and furthermore make the receiver's opinion about them worse, on average.[5] For senders of type $t_i = B$, there is a cutoff z, so that those with $l_i < z$ send message $m_i = A$ (i.e., they lie if it is not too costly) and those with $l_i > z$ send $m_i = B$. As a result, recipient j who got message $m_i = A$ believes that the sender is $t_i = A$ with probability

$$\Pr_j(t_i = A \mid m_i = A) = \frac{q}{q + (1-q)z}$$

and a recipient who got message $m_i = B$ believes that the sender is type $t_i = B$ for sure, so

$$\Pr_j(t_i = A \mid m_i = B) = 0.$$

---

[5]This argument implicitly uses the fact that types $t = A$ are relatively more likely to send message $m = A$ than types $t = B$. We formally prove that this holds in every equilibrium in the Supplemental Appendix. Intuitively, this follows from the presence of costs of lying, which make the two messages asymmetric.

This implies that for a sender with type $(t_i = B, l_i)$, sending message $m_i = A$ results in expected utility

$$U_i\left(m_i = A\right) = aq\frac{q}{q + (1-q)\,z} + a\left(1-q\right)\frac{(1-q)\,z}{q + (1-q)\,z} - l_i,$$

whereas sending message $m_i = B$ results in expected utility

$$U_i\left(m_i = B\right) = a\left(1-q\right).$$

(Because only recipients of type B will approve of him). The sender with type $(t_i = B, l_i = z)$ is indifferent if and only if

$$aq\frac{q}{q + (1-q)\,z} + a\left(1-q\right)\frac{(1-q)\,z}{q + (1-q)\,z} - z = a\left(1-q\right),$$

or, equivalently,

$$-\left(1-q\right) z^2 - qz + aq\left(2q - 1\right) = 0.$$

The left-hand side is positive at $z = 0$ and equals $aq\left(2q - 1\right) - 1$ at $z = 1$. Thus, if $aq\left(2q - 1\right) \geq 1$, then (almost) everyone sends message $m_i = A$. If $aq\left(2q - 1\right) < 1$, then there is a unique interior solution, given by

$$z = \sqrt{\frac{1}{4}\left(\frac{q}{1-q}\right)^2 + a\frac{(2q-1)\,q}{1-q}} - \frac{q}{2\left(1-q\right)}. \tag{3.2}$$

If $q < \frac{1}{2}$, the analysis is similar; the difference is that all types with $t_i = B$ communicate truthfully, whereas at least some types with $t_i = A$ (those with lowest cost of lying) send message $m_i = B$. More precisely, if $a\left(1-q\right)\left(1 - 2q\right) \geq 1$, then everyone sends $m_i = B$, and otherwise the cutoff of citizens with $t_i = A$ is given by

$$\tilde{z} = \sqrt{\frac{1}{4}\left(\frac{1-q}{q}\right)^2 + a\left(1 - 2q\right)\frac{1-q}{q}} - \frac{1-q}{2q}. \tag{3.3}$$

Lastly, if $q = \frac{1}{2}$, then everyone communicates truthfully. We summarize these results in the following proposition.

Proposition 1. *Denote* $v = \frac{2}{\sqrt{a^2 + 8a} + a}$. *There is a unique equilibrium, taking the following form.*

(i) If $q \leq \frac{1}{2} - v$, then every sender sends message $m_i = B$.

(ii) If $q \in \left(\frac{1}{2} - v, \frac{1}{2}\right)$ then citizens with conviction B send message B for sure, while citizens with conviction A send message B iff their costs of lying $l_i < \tilde{z}$, where $\tilde{z} \in (0,1)$ is given by (3.3), and send message A otherwise.

(iii) If $q = \frac{1}{2}$, then every sender sends message $m_i = t_i$, i.e., communicates truthfully.

(iv) If $q \in \left(\frac{1}{2}, \frac{1}{2} + v\right)$, then citizens with conviction A send message A for sure, while citizens with conviction B send message A iff their costs of lying $l_i < z$, where $z \in (0,1)$ is given by (3.2), and send message B otherwise.

(v) If $q \geq \frac{1}{2} + v$, then every sender sends message $m_i = A$

Thus, if $q = \frac{1}{2}$ (case *iii*), then everyone would communicate truthfully regardless of a. If a is high enough ($a > 1$) and q is close to 0 or 1 (cases *i* and *v*), then the equilibrium takes the 'political correctness' form, where all senders communicate the message that a majority of receivers would prefer. The requirement $a > 1$ reflects that the sender's concern about the receiver's opinions has to be strong enough to overcome even the highest cost of lying. For intermediate values of q, citizens who have a low cost of lying pander to the majority, but those who have high cost of lying do not.

Even if the equilibrium is not fully of the 'political correctness' form, communication depends on the beliefs that citizens have as well as their sensitivity to the opinion of others. The cutoffs (3.2) and (3.3) are both increasing in a, meaning that if senders are more concerned with the receiver's opinion, they are more likely to send a message preferred by a majority of receivers. Furthermore, (3.2) is increasing in q, provided that $q > \frac{1}{2}$. This happens for two reasons. First, a higher q makes it more likely that the recipient is of type A, which makes the sender even more willing to send message A. Second, and more subtly, a higher q increases the receiver's prior that the sender is type A, which means that the drop in his opinion about the sender who discloses his conviction B is now higher, and the sender's chance to 'pass' as type A if he sends message A is also higher. The cutoff (3.3) is decreasing in q for similar reasons, which again implies that the share of senders who send A is increasing in q even if $q < \frac{1}{2}$. Of course, a higher objective number of citizens of type A, p, also increases the share of senders sending A: while this does not change the strategy

of an individual sender, this is true since citizens with conviction A are relatively more likely to send A.

We summarize these comparative statics results in the following proposition.

Proposition 2. *Suppose that $q \in \left(\frac{1}{2} - v, \frac{1}{2} + v\right)$, so in the unique equilibrium both messages are sent with positive probabilities. Then:*

(i) An increase in p, the objective share of citizens with conviction A, leads to more senders sending message A;

(ii) An increase in q, the citizens' belief about the share of those with conviction A, also leads to more senders sending message A.

(iii) An increase in a leads to more senders sending message A if $q > \frac{1}{2}$ and to more senders sending message B if $q < \frac{1}{2}$. There is no effect if $q = \frac{1}{2}$.

More generally, for $q < \frac{1}{2} - v$ or $q > \frac{1}{2} + v$, these results hold in a weak sense.

Proposition 2 describes the effect of parameters on the behavior of senders. We are also interested in their effect on the receiver's beliefs about the sender following the receipt of either message. To understand these effects, suppose that $q > \frac{1}{2}$. Then after receiving message B, the receiver is certain that the sender was of type B, so $\Pr_j(t_i = A \mid m_i = B) = 0$. At the same time, as argued above, after receiving message A, the receiver's posterior probability that the sender is of type A equals $\Pr_j(t_i = A \mid m_i = A) = \frac{q}{q+(1-q)z}$. Since z is increasing in a, this posterior belief is decreasing in a. This is natural: if a is higher, more citizens of type B misrepresent themselves by sending message $m_i = A$, thus making at less likely that a random sender of message A is truly type A. Now consider the comparative statics with respect to q. Here, there are two effects. First, the cutoff z is increasing in q, thus making it harder to be sure that a sender with message A is indeed type A. But, on the other hand, a higher q implies an overall higher share of citizens with type A (as perceived by the receiver), which implies a higher prior they have on sender being type A, which, all things equal, would imply a higher posterior. These two effects contribute to a non-monotone overall effect: the posterior $\Pr_j(t_i = A \mid m_i = A)$ is highest (equal to 1) at $q = \frac{1}{2}$ and $q = 1$, and otherwise it is U-shaped, with the minimum attained at $q = \sqrt{\frac{1}{2}} \approx 0.71$. We thus have the following result.

Proposition 3. *Suppose again that both messages are sent with positive probabilities. Then the posterior of a receiver j that the sender who sent message $m_i = A$ is indeed $t_i = A$, $\Pr_j(t_i = A \mid m_i = A)$, is constant and equal to 1 for $q < \frac{1}{2}$, decreases in q when $q \in \left(\frac{1}{2}, \sqrt{\frac{1}{2}}\right)$ and increases back to 1 when $q \in \left(\sqrt{\frac{1}{2}}, 1\right)$. This posterior always exceeds $\frac{1}{2}$ and is decreasing in a if $q > \frac{1}{2}$. It does not depend on the objective share of citizens with type A, p.*

In other words, the receiver's posterior that the sender who sent message $m_i = A$ is indeed type $t_i = A$ evolves as follows. If conviction A is shared only by a minority ($q < \frac{1}{2}$), then message $m_i = A$ is taken at face value, as only $t_i = A$ would send it. As the share of citizens with conviction A becomes a majority, receivers would recognize that this message could be sent for strategic reason and this would decrease $\Pr_j(t_i = A \mid m_i = A)$ down from 1. However, at some point (specifically, at $q = \sqrt{\frac{1}{2}}$), they would recognize that there are many citizens with type $t_i = A$, and the posterior would increase again, up to 1 as q tends to 1). Of course, a similar result applies with respect to $\Pr_j(t_i = A \mid m_i = B)$, which we omit to save space.

3.2.3   Dynamics

Let us now consider a simple extension of the previous model that will enable us to study how certain shocks to the underlying parameters affect communication strategies within the society, as well as social learning. There may be multiple reasons for shocks: arrival of new information that makes some people change their opinions, a speech by an influential politician or celebrity, or others. Here, we consider elections as a particular case of a shock that aggregates convictions and allows citizens to learn about members of their society that they do not directly communicate with.[6]

Specifically, we consider a three-period model. In each of the periods, each citizen communicates with a continuum of other citizens, and we assume that in these communications he is both the sender and the receiver. Informally, we think about citizens adopting a certain 'persona' for the whole period, and if what

---

[6]Political campaigns that precede elections may, of course, persuade people to change convictions about issues or make some issues more salient than others. This may also change communication strategies in the society. Here, we only focus on the moment of elections as the source of the information shock.

they learned during that period about the rest of the society suggests it would be better to act differently, they can send different messages in the subsequent periods. Formally, we assume that each citizen is matched with three disjoint continua of other citizens, and in each of the periods, he sends a message to all other citizens matched to him in that period and receives messages from all of them.[7] Before period 1, the share of citizens having conviction A, i.e., p, is realized, and Nature instead sends a public noisy signal to all citizens, so citizens start period 1 with a common belief that the share of citizens with conviction A is q. They communicate in period 1, update on p and communicate in period 2. Between periods 2 and 3, Nature reveals p (e.g., through elections), after which the citizens have a final round of communication. Such timing allows us to compare the social norms learned through strategic communication (period 2) and the social norms learned through public information aggregation (period 3). We do not introduce any discounting, but since each agent is infinitesimal and does not expect to change others' beliefs for the subsequent periods, each agent would communicate to maximize his static payoffs with or without discounting.

We have the following result.

Proposition 4. *Let again* $v = \frac{2}{\sqrt{a^2 + 8a} + a}$. *Then:*

(i) If the initial signal q satisfies $q \in \left(\frac{1}{2} - v, \frac{1}{2} + v\right)$, then each citizen sends the same message in periods 2 and 3. In period 1, some act differently, unless $p = q$.

(ii) If the initial signal q satisfies $q \notin \left(\frac{1}{2} - v, \frac{1}{2} + v\right)$, then each citizen sends the same message in periods 1 and 2. They also act the same way in period 3 if $p, q \leq \frac{1}{2} - v$ or $p, q \geq \frac{1}{2} + v$, and differently otherwise.

This result is mainly of interest if the original belief was wrong. If q is relatively close to $\frac{1}{2}$, then both messages are sent with positive probabilities. Each citizen is then perfectly able to deduce p from what they observe in period 1. Indeed, suppose, for the sake of the argument, that $q \geq \frac{1}{2}$. In this case, the share of messages A equals $p + (1 - p)z$, where z is given by (3.2). For the assumed values of q, $z < 1$, and therefore p can be perfectly learned. Once this is the case, there is no learning between periods 2 and 3, and therefore

---

[7]For example, suppose that citizens are randomly allocated in a unit cube with coordinates (x, y, z). In period 1, citizen (x, y, z) communicates with citizens of the form $(*, y, z)$, in period 2, he does so with $(x, *, z)$, and in period 3, he communicates with those of the form $(x, y, *)$. Within a period, all citizens communicate simultaneously, so the message that a citizen sends is not affected by what he receives in that period.

citizens act identically, knowing that the share of citizens with conviction A is p and is common knowledge.

The pattern is very different if q is not close to $\frac{1}{2}$ (and $a > 1$). In this case, all citizens, regardless of their conviction, send the same message in period 1 in equilibrium (A if q is close to 1 and B if q is close to 0). As a result, what each citizen observes does not depend on p, and therefore no citizen makes any inference about p. The second period is then identical to the first one. However, the true value of p is revealed before period 3, and this can potentially lead to a different behavior. Namely, if $q > \frac{1}{2} + v$, so all citizens were sending A in periods 1 and 2, then this will only continue if $p > \frac{1}{2} + v$ as well, while if $p \le \frac{1}{2} + v$, then some (or, in the extreme case $p \le \frac{1}{2} - v$, all) citizens will start sending message B.

We thus observe the following. If the original citizens' beliefs about their fellow citizens are sufficiently moderate, the society is able to learn the true parameter through communication, and elections do not reveal new information and therefore do not lead to a change in behavior. However, if the original beliefs led the society to a 'political correctness' equilibrium, then there is no learning in communication. In this case, elections can reveal new and surprising information, and thus can lead to a discontinuous change in beliefs and behaviors.

3.2.4 Predictions

The model implies the following testable predictions.

1. An increase in the citizens' prior belief about the share of citizens having a certain conviction also increases the likelihood that the corresponding message is sent.

2. The receiver's posterior that the sender who sent a message indeed has the corresponding conviction depends on this prior nonmonotonically. Namely, it equals 1 for priors on $\left[0, \frac{1}{2}\right] \cup \{1\}$, and it is U-shaped on $\left(\frac{1}{2}, 1\right)$, while always exceeding $\frac{1}{2}$.

3. A shock that aggregates preferences in the society does not change individual behavior if the society was not in a 'political correctness' equilibrium. If it was, then a shock may change behavior.

3.3   Main Experiments: U.S. Presidential Elections

We developed two experiments with workers from the online platform MTurk. A number of recent papers in economics have used the same platform to conduct surveys or experiments (e.g., Kuziemko et al., 2015; Elias et al., 2016). The platform draws workers from very diverse backgrounds, though it is not representative of the U.S. population as a whole.

3.3.1   Experimental Design

Intervention Before the Election

During the two weeks prior to the presidential election, we recruited participants (N = 458) from the eight states in which the expected probability of Donald Trump's victory at the state level was 100%, according to the website *Predictwise*: Alabama, Arkansas, Idaho, Nebraska, Oklahoma, Mississippi, West Virginia, and Wyoming. MTurk workers with at least 80% approval rate could see our request, which was described as a "5 minute survey" with a reward of $0.50. Each worker could participate in the survey only once. Workers who clicked on the request were displayed detailed instructions about the task, and given access to links to the study information sheet and the actual survey. The survey was conducted on the online platform *Qualtrics.*

After answering a number of demographic questions, half of the participants were randomly informed about the 100% local odds from the website (*information* condition) while the other half were not informed (*control* condition).[8] Though restricting to these states might affect the external validity of the findings, it also allows us not to worry about the role of heterogeneous priors (and updates) in response to an informational treatment: the 100% forecast ensured that for this half of the sample, the direction of the update in Trump's perceived local popularity is either zero or positive, but never negative.[9]

---

[8]See the survey scripts in the Supplemental Appendix.

[9]Eliciting priors in the control group to assess the direction of the update would have been challenging since the forecast information was available online. Therefore, asking the question before the donation decision could have undone the treatment. Answers to the question if asked after the donation decision could have been affected by the decision itself and by the private/public condition later assigned to the participant.

Participants were then all asked to predict to share of individuals in their state that agree with a relatively strong anti-immigration statement:[10]

> Both legal and illegal immigration should be drastically reduced because immigrants undermine
>
> American culture and do not respect American values.

This provides a measure of the perceived local popularity of anti-immigrant sentiment.

In the next part of the intervention, we measured the perceived social acceptability of strong anti-immigrant sentiment using a donation experiment with real stakes. Participants were first told that they would be given the possibility to make a donation to a randomly drawn organization that could either be anti- or pro-immigration, to ensure that participants would not associate the experimenters with a specific political view. To maximize power and avoid direct deception, the randomization was such that more than 90% of participants (N=428) would get assigned the organization we were interested in: the *Federation of American Immigration Reform*.[11] To make sure participants were aware of the organization's very strong anti-immigration stance, a few more details about the organization and its founder were provided in the experiment:

> The Federation for American Immigration Reform (FAIR) is an **immigration-reduction or-**
>
> **ganization** of concerned individuals who believe that immigration laws must be reformed,
>
> and seeks to reduce overall immigration (both legal and illegal) into the United States. The

---

[10]Here we describe the protocol of the experiment as it was registered on the AEA RCT registry with number AEARCTR-0001752. As described in the registry, we planned to reach 400 individuals by November 7 – the day before the election and thus conceptually the last day in which the survey could be done (also pre-registered as the trial end date). In the piloting phase (on October 26 and 27,) we were able to recruit 184 participants. We thus expected not to have any issue recruiting 400 more subjects in the eight days between October 31 and November 7, given that in two days we were able to survey nearly half of that sample size. However, only 274 MTurk workers selected themselves into the study during the registered trial dates. The number of active MTurk workers in these states is lower than we had originally expected (and to our knowledge no estimates of the MTurk population in those states exist), which made it difficult to recruit enough participants before the election. In order to reach the desired (and registered) sample size, we decided to include individuals who participated in the pilot experiment conducted before the registration with nearly identical versions of the protocol. In particular, both the wording of the informational treatment and the wording of the donation decision were completely unchanged. If we restrict the analysis to the 274 subjects who followed the registered protocol, results are directionally similar, as discussed in Subsection 3.3.3.

[11]The pro-immigration organization was the National Immigration Forum.

founder of FAIR is John Tanton, author of 'The Immigration Invasion' who wrote "I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."

Participants were then asked if they would like to authorize the researchers donate \$1 to that organization on their behalf. The money would not come from the subject's \$0.50 payment for participation in the study. Moreover, the participant would also be paid an *extra* \$1 (or about 1/6 of an hourly wage on MTurk) if he/she authorized the donation. Rejecting the donation would not affect the monetary payoffs to the participants in any way other than through the loss of this extra amount.

In addition to the original randomization of informing subjects about Trump's probability of victory in the participant's state, we introduced a second layer of cross-randomization at the donation stage. Half of the participants were assured that their donation authorization would be kept completely anonymous, and that no one, not even the researchers would be able to match their decision to their name: we refer to this condition as the *Private* condition. Specifically, participants were told:

> Note: just like any other answer to this survey, also **your donation decision will be completely anonymous**. No one, not even the researchers, will be able to match your decision to your name.

The other half of the subjects were instead informed, right before the donation question was displayed to them, that they might be personally contacted by the research team to verify their answers to the questions in the remaining part of the survey: this is what we refer to as the *Public* condition.

> Important: in order to ensure the quality of the data collected, a member of the research team **might personally contact you** to verify your answers to the next question and the following ones.

Names and contact information were not collected during the intervention, since the practice is not allowed on MTurk. As a result, it was not possible to credibly lead participants to believe that their decision would be observed by other individuals, for example, from their state. However, on MTurk it is possible to

136

contact participants individually on the platform via their worker ID. We were therefore able to minimize the use of deception since the decision was anonymous yet researchers could still potentially contact participants (moreover, participants in the public condition might have believed that they would be asked for personal information in case they were contacted later on). As mentioned before, social acceptability with respect to surveyors and solicitors is also informative to the study of social pressure and social image concerns, as examined in DellaVigna et al. (2012b) and DellaVigna et al. (2017b).

Intervention After the Election

We also exploited the natural experiment of Trump's unexpected victory as an alternative "treatment" potentially leading to an increase in the social acceptability of holding xenophobic views. We repeated the experimental intervention in the same states during the first week after the election, restricting the design to the control condition with no additional initial information on Trump's popularity. We recruited both subjects who had participated before the election (N = 168; 166 of them assigned to the anti-immigration organization) and new participants (N = 218; 215 assigned to that organization). Based on naturally occurring variation, we can assess the impact of Trump's electoral victory on the perceived social acceptability of xenophobia.

3.3.2   Linking the Experiment to the Theoretical Framework

In what follows, we assume that A is the xenophobic conviction, while B is the opposite (tolerant) one. In the main experiment, the communication is between the subject (the *sender*) and the researcher (the *receiver*). Moreover, q here corresponds to the sender's beliefs about the share of xenophobes in the country. We consider two treatments that test the effect of manipulating q on the likelihood that a xenophobic message is sent, thus testing the comparative static result associated with q from Proposition 2:

- Treatment 1: the researchers communicate to the subject that Trump is winning for sure in their state.

- Treatment 2: Trump wins the election.

Note that though both treatments increase q, their effect might operate through different channels. In particular, since Treatment 1 provides information about the area where the sender lives, the effect is likely operating through the sender updating his beliefs about what the receiver thinks of him in the absence or in the presence of a xenophobic message being communicated. On the other hand, since Treatment 2 provides information at the country level, it might also lead to an update in the sender's priors on whether the receiver is a xenophobe. In our discussion of results, we provide suggestive evidence consistent with these hypotheses.

3.3.3 Main Results

Appendix Table C.1 provides evidence that individual characteristics are balanced across all four pre-election experimental conditions, confirming that the randomization was successful. The first four columns of Figure 3.1 display our main findings from the pre-election experiment. In the control condition before the election, we observe a large and statistically significant wedge between donation rates in private and in public: a drop from 54% in private to 34% in public (the p-value of a t test of equality is 0.002). Among individuals in the information condition, we observe no difference in private and public donation rates, which are 47% and 46%, respectively (p-value=0.839). Moreover, we find no significant difference in private donation rates between the information and control conditions (p-value=0.280), suggesting that the information is not increasing privately-held xenophobia. The increase in public donation rates between the two conditions is statistically significant (p-value=0.089), as is the difference in differences between donation rates in private across conditions and donation rates in public across conditions (p-value=0.050). These results indicate that the information provided causally increased the social acceptability of the action to the point of eliminating the original social stigma associated with it. The first two columns of Table 3.1 display the difference in differences results in regression format and show that our results are unchanged when individual covariates are included. The table also displays p-values from permutation tests, showing that our findings are robust to that inference method.

As an additional way of examining the effect of Trump's increased popularity on public expression of the xenophobic action, we compare the private and public donation rates in the control condition before

and after the election. In the last two columns of Figure 3.1, we analyze the actions of respondents who participated in both waves of the experiment. Though we focus on a subset of the original participants, we find no evidence of selective attrition, and the samples in the different conditions (before and after) are again well balanced (see Appendix Table C.2). In private, we again observe no increase in donation rates (54% before the election and 49% after the election, p-value=0.440). In public, we observe a significant increase from 34% before the election to 48% after it (p-value=0.060). The difference in differences between donation rates in private before and after the election and donation rates in public before and after the election in the control condition is also statistically significant (p-value=0.062). It is worth emphasizing that the donation rates following the two different "treatments" (either experimental or natural) are extremely similar: 47% vs. 49% in private, and 46% vs. 48% in public. The last two columns of Table 3.1 display the results in regression format, and again confirm that the findings are robust to using permutation tests. Our results are also robust to different samples for the post-election experiment, such as also including new participants, as displayed in Appendix Table C.3.[12]

3.3.4   Discussion

Potential Mechanisms

In the experimental intervention before the election, the information provided to participants generated a positive (or null) update in their beliefs about Trump's local popularity. Although this might have also updated a participant's belief regarding whether the surveyor is a xenophobe, we believe that the main effect of the information shock was to update a participant's belief about the surveyor's priors about the share of

---

[12]If we restrict the analysis to the 274 subjects who followed the registered protocol results are directionally similar: raw donation rates are respectively, for the the private and public groups, 54% and 35% in the control condition before the election, 50% and 39% in the information condition before the election, and 39% and 45% in the control condition after the election. As with the full sample, the wedge between the public and private condition is significant in the control condition before the election (p-value 0.018), and not significant in the information condition (p-value 0.219), or after the election (p-value 0.607). The difference in differences between donation rates in private across conditions and donation rates in public across conditions is smaller than in the full sample and not significant before the election (8.4%, p-value 0.486), but large and significant after the election (24.9%, p-value 0.068).

xenophobes around the participant. While we don't have a direct measure of this belief about the surveyor's prior, we elicited the update in participants' beliefs about the share of xenophobes in their home state – a good proxy for the other variable. For the proxy to be valid, we need the update in participants' beliefs about the share of xenophobes in their state to go in the *same direction* as the update in their beliefs about the surveyor's beliefs about that same share. We believe this a reasonable assumption.

Figure 3.2 provides evidence that the information shock led to a positive update in beliefs about the opinions about other individuals in the same state. In the control condition before the election, the average guess was that 64% of other people in the participant's home state would agree with the xenophobic statement, while it was 68% in the information condition (p-value=0.062). This small increase in the average guess might not fully display the impact that the information intervention had on the *distribution* of beliefs about others. The distribution in the information condition first-order stochastically dominates the distribution in the control condition (a Kolmogorov-Smirnov test of equality of the two distributions yields a p-value of 0.072). For example, the percentage of participants who think that the share of those agreeing with the xenophobic statement in their state is above 90% increases substantially with the provision of information (from 9% to 17%, with a p-value of 0.018).

When we analyze the effects of Trump's actual victory on the social stigma associated with the donation decision, it is less clear *ex ante* that an update in respondents' beliefs about the surveyor's priors regarding the share of xenophobes in the respondents' state would the main driver of the findings. After the election, the surveyor is not directly providing any information that would necessarily lead to an unambiguously weakly positive update in perceptions of Trump's local popularity. For example, participants might now believe that Trump's actual margin of victory in their state might was smaller than what the surveyor had originally expected. Indeed, when comparing the pre-election control group with the group after the election, we find no shift in the distribution of beliefs about the share of xenophobes in participants' home states (see Appendix Figure C.1). However, there is an unambiguously weakly positive update in participants' perception of Trump's *national* popularity. According to our model, this will increase the probability that participants: (i) think that the surveyor is xenophobic, and (ii) think that the surveyor will judge them less

negatively even if the researcher is not xenophobic. Hence, although we do not provide direct evidence of these mechanisms, the nature of the intervention suggests that these two elements are the likely drivers of the findings in the comparison of pre- and post-election donation rates.

Further Interpreting the Findings

Although the effects we find are extremely large, it is important to note that they are coming from a marginal positive signal of Trump's local popularity starting from a situation where he was already believed to be extremely popular locally. Even in the control condition, the average individual believes that 64% of people in his/her state agree with the xenophobic statement, and 34% of participants make the donation in public. Though we cannot test this hypothesis, perhaps an increase in the perceived popularity of holding xenophobic views and a reduction of the related social stigma might have already taken place throughout the presidential campaign, before the experiment took place. It is also possible that the statement we chose might have been perceived as relatively mild toward the end of the campaign period, so the small update in beliefs about the views of others regarding the chosen statement could correspond to larger updates for a more extreme statement. Of course, we cannot estimate how large the effects of our experiment would have been had the baseline level of Trump's perceived local popularity been lower (as in the beginning of the campaign season) and thus the effect of the signal possibly stronger.

Finally, we can rule out that our main effect is coming from participants updating their beliefs about the instrumental benefits associated with donating to the organization (for example, because the organization is now more likely to fulfill its mission). Any effect explained by that factor would have showed up in the private donation rates too.

3.4 Auxiliary Experiment: Dictator Game

3.4.1 Experimental Design

Wave 1 – Non-Anonymous Behavior by the Swiss Player

In late February 2017, we recruited participants from the six states in which Hillary Clinton won the presidential election with the highest margin: California, Hawaii, Maryland, Massachusetts, New York, and Vermont. This was done to maximize the chances of recruiting subjects with liberal views, and in particular subjects with no anti-Muslim sentiment.[13]

First, after answering a number of demographic questions, all participants were explained that a minaret is a tower typically built adjacent to a mosque and traditionally used for the Muslim call to prayer. Second, they were asked whether they would support the introduction of a law prohibiting the building of minarets in their state. We focus on subjects who reported to be against the introduction of this law (N = 396), and we examine how they interact with a person who has opposite views. In order to do so, in the third part of the survey, participants were told they were matched with a subject from another survey and were asked to play a dictator game in which they could decide how to split $3 (half of an hourly wage on the platform) between themselves and the other participant. We randomly assigned our participants to three different groups and randomized the background information we gave to our participants about the person they were matched with. Participants in the control group were only told that the participant they were matched with was a 24-year-old male from Switzerland. Note that we used real 24-year-old male subjects from Switzerland recruited to take part of a short survey by a research assistant from the University of Zurich.

Participants in the *anti-minarets* group were additionally told that this person supports the prohibition of the building of minarets in Switzerland. Participants in the *anti-minarets, public support* group were

---

[13]As in the main experiment, Turk workers with at least 80% approval rate could see our request, which in this case was described as a "4-5 minutes short survey" with a reward of $0.50. Each worker could participate in the survey only once. Workers who clicked on the request were displayed detailed instructions about the task, and given access to links to the study information sheet and the actual survey. The survey was conducted on the online platform *Qualtrics*. The script used in this auxiliary experiment is displayed in the Supplemental Appendix.

instead told that "like 57.5% of Swiss respondents, the participant supports the prohibition of the building of minarets in Switzerland."

Wave 2 – Anonymous Behavior by the Swiss Player

If we find higher donations in the *Anti-minarets, public support group*, when compared to the *anti-minarets*, we can conclude that the participants may believe that the Swiss person has strategic reasons to state that he is anti-minarets, and for this reason judge him less for expressing that view. However, a potential alternative interpretation for this result is that participants might judge the Swiss person less negatively when the majority of Swiss support the ban, *regardless of whether his support was expressed in public.* For example, it could be that participants feel that they cannot blame a person for privately holding a view if that person is surrounded by many other people who also hold that view and who could also have influenced the private opinions of that person.

To explicitly rule out this possibility, immediately after we completed wave 1 we conducted a slightly modified version of our protocol on the days immediately following the first wave of the dictator game experiment. In the second wave, participants were informed about the fact that the 24-year-old male from Switzerland expressed his opinion in an *anonymous* survey. As in the original version, we recruited participants from the six states in which Hillary Clinton won the presidential election with the highest margin (California, Hawaii, Maryland, Massachusetts, New York, and Vermont). To make sure we could recruit enough participants, we also included six other states with a high share of votes for Clinton (Connecticut, Delaware, Illinois, New Jersey, Rhode Island, and Washington).[14]

The design of this experiment was almost identical to the original version. Once again, we focus on subjects who reported to be against the introduction of the ban (N = 427). The main difference with the original version, is that we emphasized how the Swiss participant expressed his opinion anonymously.

---

[14]As in the other two experiments, Turk workers with at least 80% approval rate could see our request, which in this case was described as a "4-5 minutes short survey" with a reward of $0.50. Each worker could participate in the survey only once, and only if she did not participate in our other experiment. Workers who clicked on the request were displayed detailed instructions about the task, and given access to links to the study information sheet and the actual survey. The survey was conducted on the online platform *Qualtrics*. The script used in this second wave of the auxiliary experiment is displayed in the Supplemental Appendix.

Both in the control and in the treatment conditions, instead of writing as before that "we matched you with a participant from another survey", in this version we wrote "we matched you with a participant from another anonymous survey." In our treatment groups we emphasized once again that the survey the Swiss participants participated in was anonymous: "In our anonymous survey, like the one you just completed, he said he supports the prohibition of the building of minarets in Switzerland." We call this first treatment group the *anonymous anti-minarets* group. Finally, instead of writing "like 57.5% of Swiss respondents, the participant supports the prohibition of the building of minarets in Switzerland" in this case we wrote "According to numbers from 2009, 57.5% of Swiss respondents are in favor of prohibiting the building of minarets." We call this second treatment group the *anonymous anti-minarets, public support* group.[15]

Elicitation of Beliefs

At the end of the intervention, subjects in the control group were also asked about their beliefs regarding the share of the Swiss population who supports banning the construction of minarets, and whether they believed the ban is legal in Switzerland. In the first wave we did not collect this information for individuals in the *anti-minarets* and *anti-minarets public support* groups. To check whether their beliefs about the share of the Swiss population supporting the ban are not changed by the treatments, we included both the control group and the treatment groups in the second wave. The share thinking that a majority of Swiss support the ban is almost identical in the control group and the *anti-minarets* group (respectively 20 and 25%, with a p-value for the test of equality of 0.301), but increases to 63% in the *anti-minarets public support* group (p-values of the test of equality are 0.000 against both groups). The median belief about the share of

---

[15]Our design also included a fourth group (N=136 in wave 1, and N=139 in wave 2), where participants were instead told: "Building minarets is illegal in Switzerland, following a 2009 referendum. Like 57.5% of Swiss respondents, the participant supports the prohibition of the building of minarets in Switzerland. However, he did not vote in the referendum since he was under legal voting age" in wave 1, and "In our anonymous survey, like the one you just completed, he said he supports the prohibition of the building of minarets in Switzerland. Building minarets is illegal in Switzerland, following a 2009 referendum. According to numbers from 2009, 57.5% of Swiss respondents are in favor of prohibiting the building of minarets. However, the person you are matched with did not vote in the referendum since he was under legal voting age" in wave 2. This treatment was intended to test whether providing information that a view is not only held by a majority but is also *official* would additionally change the donation rates. We found no effect of this additional treatment relative to the second treatment group. We report these results in Appendix Figures C.2 and C.3.

144

the Swiss population supporting the ban is 30% in both control and *anti-minarets* groups, and 55% in the *anti-minarets public support* group. This confirms that our experimental manipulation indeed shifted beliefs about the level of popular support for the ban in Switzerland.[16]

Participants across conditions were also asked whether they believed the construction of minarets is legal in Switzerland: across the three groups, the majority reported to think that constructing minarets was legal (88% in the control group, 77% in the *anti-minarets* group, and 74% in the *anti-minarets public support* group).[17] We can thus rule out that the effects are affected by the fact that the ban is actually legal, and can thus isolate the role of public opinion on participants' judgment of the Swiss player.

3.4.2   Linking the Experiment to the Theoretical Framework

We have the following predictions, which help us test Proposition 3.

- *Anti-minarets* group: The Swiss person is revealed to be anti-minarets. The participants' prior that Swiss people are on average anti-minarets is relatively low. This has two implications. First, before the information about the Swiss person is communicated, the perceived probability that he is anti-minarets is low. Second, and more importantly, he is not thought to have strategic reasons to pretend to be anti-minarets, because this view is not thought to be popular among citizens in Switzerland (the subjects believe that the equilibrium that this Swiss person faces would imply that he will only say $m_i = A$ if he indeed has type $t_i = A$). Thus, the posterior that this person is anti-minarets increases all the way to 1. This lowers donations when compared to the control group.

- *Anti-minarets, public support* group: The Swiss person is revealed to be anti-minarets, and also we also reveal that 57% of Swiss respondents support banning minarets. Now, compared to the previous

---

[16]Here we report the numbers from the second wave of the experiment, since the first wave only asked beliefs in the control group. The numbers for this group are very similar across waves; for example, in the first wave, 17% of control group participants believe a majority of Swiss people support the ban, compared to 20% in the second wave. The median belief is 30% for the control groups in both waves.

[17]While the beliefs are significantly different when comparing the control group with either of the two treatment groups (the p-values for the test of equality are 0.013 against the *anonymous anti-minarets* group and 0.002 against the *anonymous anti-minarets public support group*), there is no statistical difference between the two treatment groups (the p-value for the test of equality is 0.500).

experimental condition, the participants may believe that the Swiss person has strategic reasons to state that he is anti-minarets. This decreases $\Pr_j (t_i = A \mid m_i = A)$ from 1 to some value which, however, must be greater than $\frac{1}{2}$, as follows from Proposition 3. This implies that donations should be higher than in the previous treatment group, but not as high as in the control group.

The second wave of the experiment helps us further test our model and rule out the main alternative interpretation:

- *Anonymous anti-minarets* group: The Swiss person is revealed to be *privately* anti-minarets. Again, subjects' prior that Swiss people are on average anti-minarets is relatively low. The posterior that the Swiss player is anti-minarets is 1, so donations should be similar to the ones in the *Anti-minarets group*.

- *Anonymous anti-minarets, public support* group: The Swiss person is revealed to be *privately* anti-minarets, and also we also reveal that 57% of Swiss respondents support banning minarets. Unlike the non-anonymous version of this treatment, participants do not think the Swiss person has strategic reasons to state that he is anti-minarets, since the expression is anonymous. As a result, the posterior that the Swiss player is against anti-minarets is again 1 and the donation levels should be similar to those in the *Anti-minarets group*.

3.4.3 Results

Appendix Table C.4 provides a test of balance of individual characteristics across all six conditions from the two waves and separately for the three conditions of each wave. Although the two waves were not conducted simultaneously, the variables are fairly well-balanced when pooling the two waves. Only the share of white respondents is marginally unbalanced (p-value of 0.085). As a result, to simplify the exposition of our findings, below we pool the observations in the control groups from the two waves, and compare the raw outcome variables across the two waves. Appendix Figures C.4 and C.5 display the results separately for the two waves of the experiment and show that the numbers are almost identical for the two control groups.

We now turn to the main findings from the auxiliary experiment, displayed in Figure 3.3. Panel A

displays comparisons of average donations across groups. In the control condition, where participants are only told that they are matched with a 24-year-old male from Switzerland, we observe an average transfer to the Swiss participant of $1.03. The average transfer is substantially lower for subjects in the *anti-minarets* group, who are also told that this person supports the prohibition on building minarets in Switzerland: the average transfer for this group is $0.69. The effect of informing subject about the anti-Muslim views of the Swiss participant is statistically significant (p-value=0.000). However, the average transfer among subjects in the *anti-minarets, public support* group who are told that the majority of Swiss respondents are against minarets is $0.92, which is not statistically different from the average transfer in the control group (the p-value of the difference is 0.162) but is substantially higher than the average transfer in the *anti-minarets* group (p-value=0.013). The average donation in the *anonymous anti-minarets* group is identical to that in the *anti-minarets* group, at $0.69. The average donation in the *anonymous anti-minarets, public support* group is also very similar: $0.70. These two levels are significantly different from the average in the control group (p-value=0.000 in both cases). The average donation in the *anonymous anti-minarets, public support* group is also significantly lower than the one in the first-wave version of the treatment (p-value=0.014).

Panel B compares across conditions the share of participants who do *not* share any of their $3 endowment with the Swiss person. The percentage of participants deciding not to transfer anything to the Swiss respondent increases from 22% in the control group to 42% in the *anti-minarets* group (p-value 0.000), while only 27% of subjects in the *anti-minarets, public support* decide to keep all $3. This percentage is not statistically different from the one in the control group (p-value=0.370), but is substantially lower than the one for subjects in the *anti-minarets* group (p-value=0.013). Here again, the levels of the outcome variable in the two anonymous treatments are almost identical to the level in the *anti-minarets* group: 43% and 44%. Importantly, the share of participants not donating is significantly higher in the *anonymous anti-minarets, public support* group when compared to the non-anonymous version of the treatment (p-value=0.004).

Panel C displays a similar pattern for the median transfer, which is $1.50 in the control group, $0.30 in the *anti-minarets* group, and $1.20 in the *anti-minarets, public support* group. While the median transfer in the control group is higher than the median transfer in both other groups (p-value against the *anti-*

147

minarets is 0.000 and against the *anti-minarets public, support* group is 0.030), providing information about the fact that 57.5% of Swiss respondents support the prohibition on building minarets makes a statistically significant difference (p-value 0.000), when compared to the *anti-minarets* group. The median donation in the two anonymous groups are identical ($0.50). The median donation in the *anonymous anti-minarets* group is not significantly different from the median in the *anti-minarets* group (p-value 0.496). The median donation in the *anonymous anti-minarets, public support* group is significantly lower than the median in the *anti-minarets, public support* group (p=0.021).

Overall, across all three outcome variables, the results confirm the four predictions stemming from the model. Appendix Table C.5 displays the results in regression format and show that our results are unchanged when individual covariates are included.

3.5   Conclusion

We show that a positive update in people's beliefs about Donald Trump's popularity increases their willingness to publicly express xenophobic views. Building on a simple model of communication and on a second auxiliary experiment with a dictator game, we also provide suggestive evidence of the mechanisms behind the effects. By aggregating information about individuals' private views, elections can update positively people's perceptions about the share of people who support an opinion previously believed to be stigmatized. This may in turn change people's perceptions about the negative judgement they will face for expressing their opinion.

Though beyond the scope of our study, the potential interaction between private and public opinions might exacerbate the effects of events that trigger changes in the social acceptability of holding particular views. Moreover, higher electoral support for a candidate associated with certain views might shape individuals' perceptions of the social desirability of those views, regardless of whether or not the candidate wins the election.

Our findings shed light on the factors that can trigger rapid change in social norms, and in particular, norms against the expression of xenophobic views. Our results suggest that social norms regarding the

expression of such views in the U.S. might have already been causally changed by Trump's rise in popularity and eventual electoral victory. More broadly, the mechanisms we study in this paper might help explain the rise – and potential consequences – of other crucial recent events such as the Brexit vote in the UK, and more generally the rise in anti-immigrant and anti-minority sentiment in the developed world.

3.6 Figures and Tables

Figure 3.1: **Main Experiment: Donation Rates Before and After the Election**
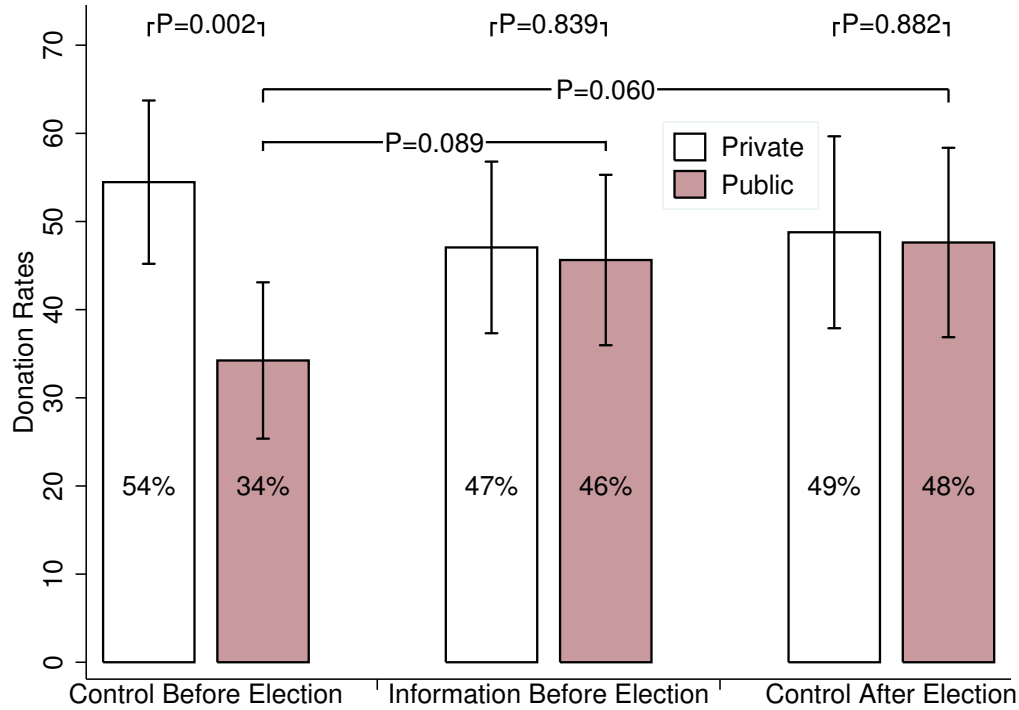
Notes: the two bars on the left display donation rates to the anti-immigration association for individuals in the private and public conditions in the control group before the election (full sample, respectively N=112 and N=111), the two central bars display those in the information group before the election (full sample, respectively N=102 and N=103), and the last two bars display those in the control group after the election (for individuals already surveyed before the election, respectively N=82 and N=84). Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Figure 3.2: **Main Experiment: Beliefs About Others**
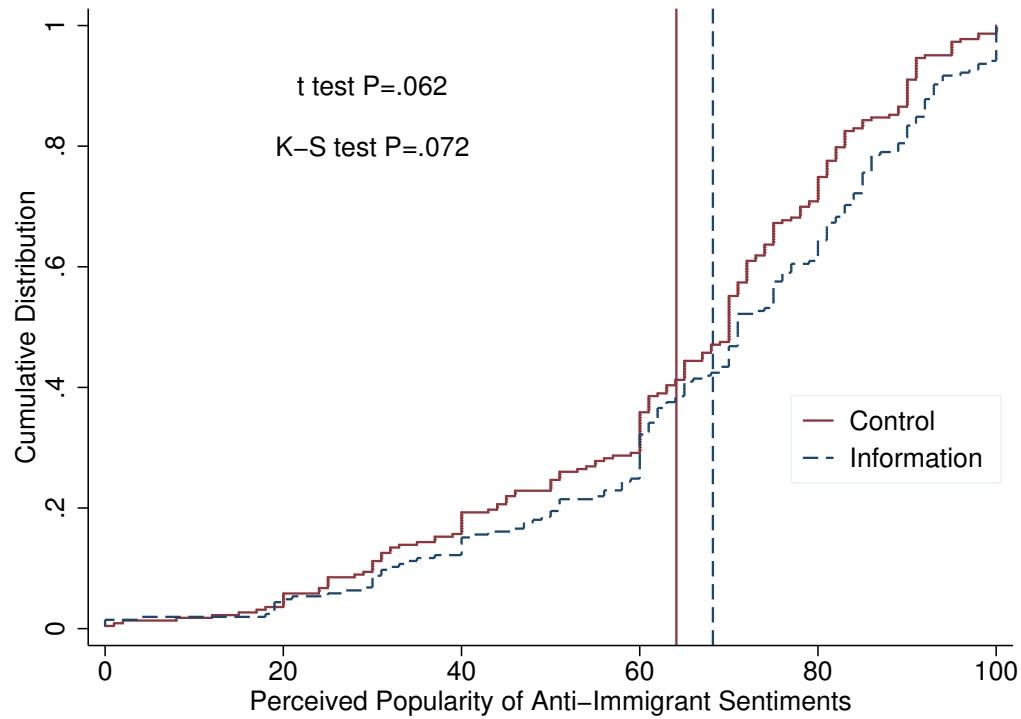
t test P=.062

K–S test P=.072

Notes: Empirical cumulative distributions of perceived popularity of anti-immigrant sentiments for individuals in the control and information conditions before the election (respectively N=223 and N=205). The two vertical lines display the means of the two distributions. K-S P is the p-value of a Kolmogorov-Smirnov test of equality of the two distributions, while t test P is the p-value of a test of equality of means.

Figure 3.3: **Auxiliary Experiments: Donation Rates**
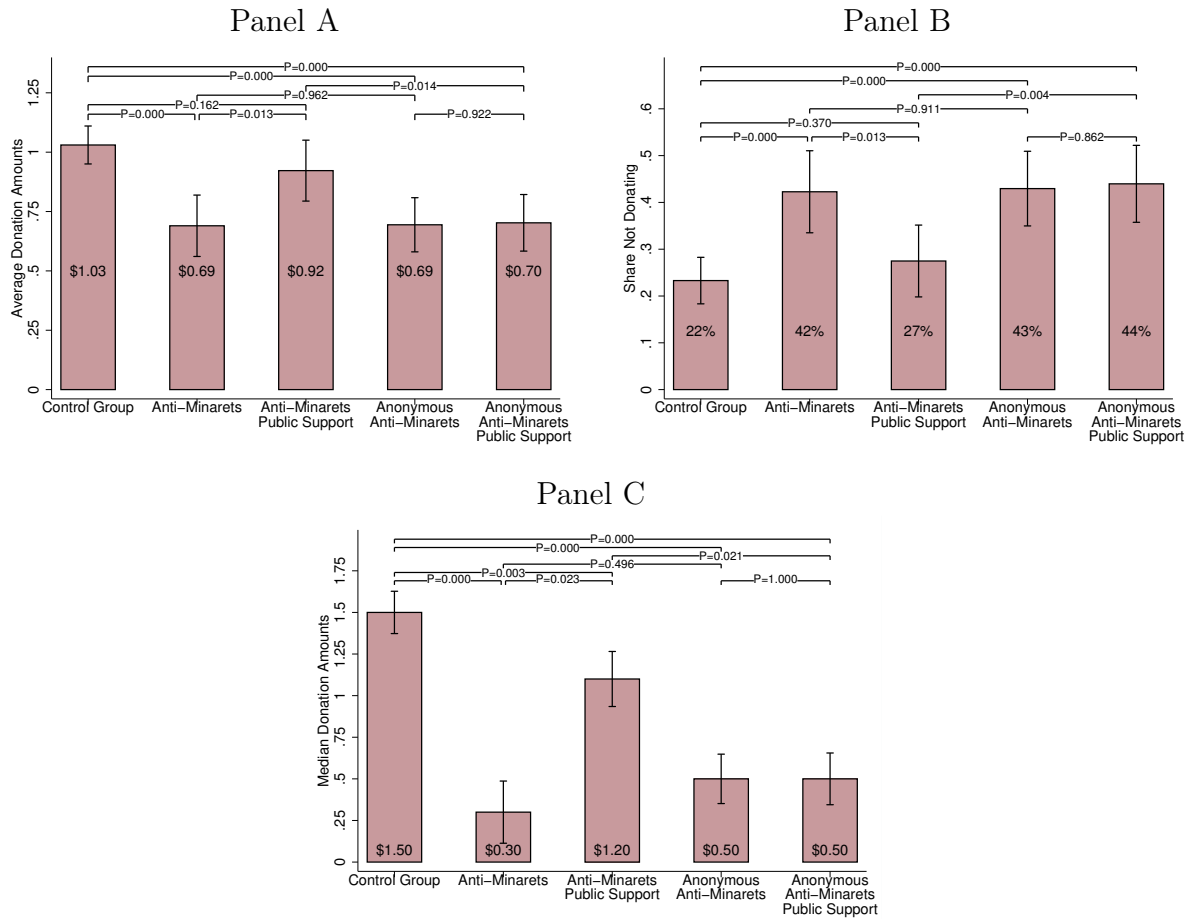
Notes: Panel A displays average donation amounts to the Swiss individual in the five experimental conditions: the control group (N = 279, pooling 142 observations from the first auxiliary experiment and 137 observations from the second anonymous version of the auxiliary experiment), the *anti-minarets* group (N=133), and the *anti-minarets public support* group (N=131), the *anonymous anti-minarets* group (N=149), and the *anonymous anti-minarets public support* group (N=141). Panel B displays the percent of subjects not making positive donations. Panel C displays median donation amounts. Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Table 3.1: **Main Experiment: Difference in Differences Regressions**

Dependent Variable	Dummy: individual donates to anti-immigrant organization			
	(1)	(2)	(3)	(4)
Public	-0.202***	-0.200***	-0.202***	-0.199***
	[0.065]	[0.066]	[0.065]	[0.065]
	(0.004)	(0.005)	(0.004)	(0.007)
Information	-0.074	-0.077	-0.074	-0.076
	[0.069]	[0.068]	[0.069]	[0.068]
	(0.277)	(0.266)	(0.277)	(0.281)
Public*Information	0.188*	0.178*	0.188*	0.178*
	[0.096]	[0.096]	[0.096]	[0.096]
	(0.045)	(0.062)	(0.045)	(0.062)
After Election			-0.057	-0.062
			[0.073]	[0.072]
			(0.380)	(0.304)
Public*After Election			0.191*	0.186*
			[0.102]	[0.101]
			(0.071)	(0.080)
Control Private Before Election			0.545	
Controls	No	Yes	No	Yes
N	428	428	594	594
R^2	0.022	0.033	0.017	0.034

Notes: Columns (1) and (2) includes the full pre-election sample. Columns (3) and (4) add the post-election sample of individuals already surveyed before the election. Columns (1) presents OLS regression of a dummy variable for whether a individual donates to the anti-immigration organization on a dummy for the Public condition, a dummy for the Information condition, and a dummy for the Public Information condition. The control private condition before the election is the omitted group, for which we report the mean donation rate. Columns (3) replicates and adds a dummy for the after election condition, and a dummy for the Public after election condition. Columns (2) and (4) replicate and add individual covariates (gender, age, marital status, years of education, household income, and race). Robust standard errors in brackets. P-values from permutation tests with 1,000 repetitions in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix A

Appendix "Moral Incentives in Credit Card Debt Repayment: Evidence from a Field Experiment"

Appendix Figures

Figure A.1: **Text Messages**
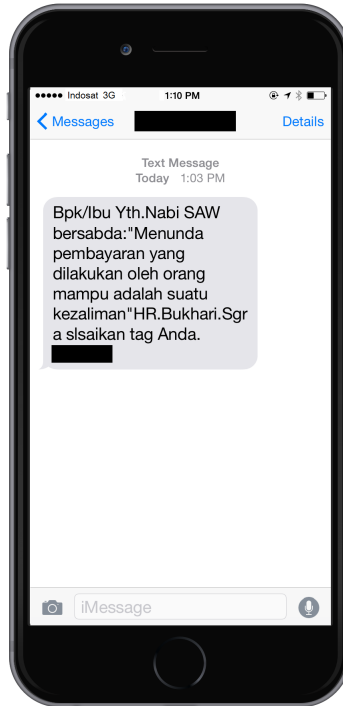
Notes: The figure shows the text message sent to experimental participants assigned to the "*moral incentive*" treatment condition.

Figure A.2: **Experimental Design**

156

Appendix Tables

Table A.1: **Sample Sizes by Wave**

	Treated	Control	Repeated	Excluded	Other Project	Total
	(1)	(2)	(3)	(4)	(5)	(6)
Wave I	2000	871	0	83	800	3754
Wave II	2000	985	0	1018	800	4803
Wave III	1000	965	0	1823	600	4388
Wave IV	1344	343	0	1652	0	3339
Wave V	1516	590	306	1075	0	3487
Wave VI	1448	366	592	1343	0	3749
Total	9308	4120	898	6994	2200	23520

Notes: Columns (1) and (2) show the number of customers randomized into treatment and control for the main experiment. Column (3) reports the number of customers who were randomized into treatment and contron for the follow-up experiment on the effect of repeated messages. Column (4) reports the number of customers excluded because they had previously received a text message treatment. Customers assigned to the control group in a previous month remained in the sample and could either be assigned to a treatment or be again in the control group. Column (5) reports the number of customers randomized into treatment for a different project. Column (6) reports the total number of late customers.

Table A.2: **Repeated Message Experiment: Balance and Treatment Cell Size**

	Panel A: Balance of Covariates			
	Full Sample	Repeated Moral Incentive	Control Group	*p*-value
	(1)	(2)	(3)	(4)
Age	42.29	42.43	42.15	0.653
	[9.375]	[9.375]	[9.384]	
Female	0.41	0.44	0.38	0.080
	[0.492]	[0.497]	[0.486]	
Muslim	0.90	0.91	0.89	0.321
	[0.296]	[0.282]	[0.309]	
Annual Income	126.72	124.07	129.35	0.702
(Rp, million)	[206.906]	[171.322]	[237.255]	
Credit Limit	13.10	13.38	12.82	0.368
(Rp, million)	[9.386]	[9.445]	[9.329]	
	Panel B: Treatment Cell Size			
Wave V	306	153	153	
Wave VI	592	295	297	
Total	898	448	450	

Notes: Panel A reports summary statistics for the follow-up experiment and presents a test of random assignment. Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample. Columns (2) and (3) report the mean level of each variable, with standard deviations in brackets, for the two experimental conditions. Column (4) reports the *p*-value of a test that means are the same in the two experimental conditions. Panel B reports treatment cell sizes by month.

Table A.3: **Heterogeneity of Treatment Effects**

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
			Dummy: customer repaid by the deadline			
Trait	Male	Age	Muslim	Local Religiosity	Debt to Income Ratio	Poor Credit History
Trait*Moral Incentive	-0.012	-0.013	0.034	0.047**	-0.040**	0.003
	[0.020]	[0.020]	[0.035]	[0.023]	[0.020]	[0.024]
Moral Incentive	0.057***	0.056***	0.018	0.038***	0.069***	0.049***
	[0.017]	[0.015]	[0.034]	[0.013]	[0.015]	[0.012]
Trait	0.012	0.005	0.010	0.054	0.023	-0.222***
	[0.015]	[0.021]	[0.026]	[0.060]	[0.014]	[0.017]
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Waves	All	All	All	All	All	All
N	8,730	8,730	8,730	8,730	8,730	8,730
R^2	0.050	0.050	0.050	0.050	0.050	0.050

Notes: The table shows heterogeneous treatment effects for the moral message (all versions). Each column shows results from a separate regression. The dependent variable in all regressions is a dummy for whether a customer has made the minimum payment by the deadline. The dependent variable is regressed on a dummy equal to one if a customer has received any version of the moral incentive treatment, the trait indicated at the top of the table and their interaction. The dummy for age is equal to one for customers with age larger than the median. The dummy for local religiosity is equal to one for customers living in provinces where the measure of local religiosity is higher than the median at the province level. The dummy on debt to income ratio is equal to one for customers with a debt to income ratio larger than the median. The dummy on poor credit history is equal to one for customer reported to the credit registry at least once in the previous six months. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.4: **First Three Waves Including Crowding-Out Experiment**

Dependent variable	Dummy: customer repaid within the deadline		
	(1)	(2)	(3)
Treatments: Main experiment			
Moral Incentive	0.060***	0.065***	0.063***
	[0.018]	[0.018]	[0.017]
Simple Reminder	0.006	0.011	0.010
	[0.018]	[0.018]	[0.017]
Religious Placebo	0.002	0.007	0.008
	[0.018]	[0.018]	[0.017]
Reputational Incentive	0.098***	0.103***	0.103***
	[0.014]	[0.014]	[0.014]
Treatments: Crowding-out experiment [multiple messages in one day]			
Moral Incentive	0.094***	0.099***	0.091***
+ Reputational Incentives	[0.018]	[0.018]	[0.018]
Simple Reminder	0.072***	0.077***	0.075***
+ Due Date Message	[0.018]	[0.018]	[0.017]
Mean Repayment Control Group	0.34	0.34	0.34
Month fixed effects	No	Yes	Yes
Controls	No	No	Yes
Waves	Waves I, II, and III	Waves I, II, and III	Waves I, II, and III
N	9821	9821	9821
R^2	0.008	0.018	0.076

Notes: Column (1), (2), and (3) restricts the sample to customers late in February, March and May 2015 and includes two groups with customers receiving multiple text messages on the same day ("Due Date Message and Simple Reminder", and "Moral and Reputational Incentives"), in addition to the control group and all other treatments run in those months. Column (1) presents OLS regression of a dummy variable for whether a customer repaid her credit card debt (made at least the minimum payment) within the deadline on treatment group dummies. The control is the omitted group, for which we report the mean repayment rate. Column (2) replicates and adds month fixed effects. Column (3) replicates and adds individual covariates (age, gender, Muslim dummy, province dummy, income, a dummy for being in the sample in a previous month, and a dummy for having been more than 30 days past due at least once in the previous 12 months). * significant at 10%; ** significant at 5%; *** significant at 1%.

Issues with the Crowding-Out Experiment: The setup of this experiment differs from the treatments in this paper in that recipients were sent *multiple messages on the same day*. The main treatment for the crowding-out experiment involved sending the moral *and* reputational messages to clients on the 16th day of the month. We find that the effect of receiving the two messages is similar to the effect of receiving the reputational message only. This is consistent with a case of strong crowding-out, but also with a ceiling effect. We are therefore not able to separate these two stories. Before running the intervention, we decided to include another "placebo" group for this second paper: in case the two messages had an effect over and above the effect of the reputational incentive alone, in principle, this could be due to an effect of a receiving any second message on the same day (in addition to the reputational message). Since, however, the moral message had no effect on top of the effect from the reputational message, *ex post* there was no real need for such placebo. In our design, we also had a placebo approach that was not ideal. The correct placebo would have been to send a neutral message in addition to the reputational message. Instead, the treatment implemented was to send two neutral messages on the same day, which complicates the interpretation. Aside from this design issue, there were also problems in the implementation of these treatments. Because the bank did not want to send two identical messages on the same day, one of the messages in the 'same day double reminder' group was a neutral reminder and the other one was the same message customers were used to receiving at the end of the billing cycle. As we later found out, this created confusion among customers who received both of these messages on the same day. Some customers erroneously believed that the bank had changed the billing cycle dates, or that they were at a later point in the billing cycle than was actually the case, since they had also received the standard end-of-billing cycle message. As a result, what we had intended as a placebo cannot really be interpreted as such. Since these treatments were part of a separate experiment and have a number of design and implementation issues, outlined above, these results were not part of our paper and we abandoned the idea of the crowding-out experiment altogether.

Table A.5: **First Time and Repeated Sample**

	Panel A: Balance of Covariates		
	(1)	(2)	(3)
	First Message Sample	Repeated Message Sample	p-value
Age	41.93	42.29	0.267
	[9.320]	[9.375]	
Female	0.39	0.41	0.382
	[0.489]	[0.492]	
Muslim	0.91	0.90	0.376
	[0.283]	[0.296]	
Income	151.52	126.72	0.013
(Rp, million)	[827.617]	[206.906]	
Credit Limit	13.64	13.10	0.092
(Rp, million)	[9.678]	[9.386]	
30DPD in last year	0.29	0.39	0.000
	[0.452]	[0.488]	
Late once before	0.10	1.00	0.000
	[0.298]	[0.000]	
	Panel B: Treatment Cell Size		
Wave I	2871	0	
Wave II	2985	0	
Wave III	1965	0	
Wave IV	1687	0	
Wave V	2106	306	
Wave VI	1814	592	
Total	13428	898	

Notes: Panel A reports summary statistics for the follow-up experiment and presents a test of random assignment. Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample. Columns (2) and (3) report the mean level of each variable, with standard deviations in brackets, for the two experimental conditions. Column (4) reports the p-value of a test that means are the same in the two experimental conditions. Panel B reports sample sizes by month.

Table A.6: **Text Messages**

	Bahasa Indonesia	English
Control: Basic Reminder	Bpk/Ibu Yth. Tag [name of the card] Anda tlh jth tempo. Utk kenyamanan & keleluasaan bertransaksi, segera lakukan pemby. Jk tlh membayar, abaikan SMS ini.[customer service number]	Dear Mr/Mrs. Your [name of the card] has reached the due date. Please make a payment at your earliest convenience. If you have already paid, ignore this text. Call [customer service number].
Moral Incentive [Religious]	Bpk/Ibu Yth.Nabi SAW bersabda:"Menunda pembayaran yang dilakukan oleh orang mampu adalah suatu kezaliman"HR.Bukhari.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. The Prophet (Peace and blessings be upon Him) says: "non-repayment of debts by someone who is able to repay is an injustice" (Imam al-Bukhari). Please repay your credit card balance at your earliest convenience. Call [customer service number].
Moral Incentive [Implicit]	Bpk/Ibu Yth.Menunda pembayaran yang dilakukan oleh orang mampu adalah suatu kezaliman.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. Non-repayment of debts by someone who is able to repay is an injustice. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Moral Incentive [Non-religious]	Bpk/Ibu Yth.Menunda pembayaran yang dilakukan oleh orang mampu adalah suatu ketidakadilan.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. Non-repayment of debts by someone who is able to repay is an injustice [non-arabic]. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Cash Rebate Incentive	Bpk/Ibu Yth.Bulan ini:slsaikan tag Anda utk mendapatkan hadiah uang tunai sebesar 50% dr pembayaran minimum pada tag berikutnya.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. This month, make your credit card payment to get a cash rebate equal to 50of your minimum payment on your next statement. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Credit Reputation Incentive I	Bpk/Ibu Yth.Ketrlmbtn pembyr dilaporkan k SistemInformasiDebitur BI,yg semua bank berkonsltasi&mengurangi kemampuan mendptkan krdt.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. Late payments are reported monthly to Bank Indonesia Sistem Informasi Debitur (SID), which all banks consult. This will diminish your ability to get credit in the future. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Credit Reputation Incentive II	Bpk/Ibu Yth.Ketrlmbtn pembyr dilaporkan k SistemInformasiDebitur BI,yg semua bank dapat berkonsultasi.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. Late payments are reported monthly to Bank Indonesia Sistem Informasi Debitur (SID), which all banks can consult. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Placebo: Simple Reminder	Bpk/Ibu Yth.Tagihan [name of the card] Anda jatuh tempo pada tanggal [due date] dan pmbayarn belum diterima.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. The due date of your [name of the card] bill was on [due date] and your payment has not been received yet. Please repay your credit card balance at your earliest convenience. Call [customer service number].
Placebo: Religious Message	Bpk/Ibu Yth.Nabi SAW bersabda:"Jika Allah menginginkan yg terbaik buat umatnya,IA melimpahkan padanya pengetahuan Kitab"HR.Bukhari.Sgra slsaikan tag Anda.[customer service number]	Dear Mr/Mrs. The Prophet (Peace and blessings be upon Him) says: "When Allah wishes good for someone, He bestows upon him the understanding of the Book" (Imam al-Bukhari). Please repay your credit card balance at your earliest convenience. Call [customer service number].

Survey Instruments

Religion and Religiosity

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I am calling from [bank name] and would like to ask a few questions to improve the services we offer with [name of the credit card]. This will take less than 5 minutes. Are you willing to participate?

1. Please rank the following in terms of importance in your life, from 1 (most important) to 4 (least important)

   - Family
   - Work
   - Friends
   - Religion

2. How important is religion in your life?
   Not important at all [1] [2] [3] [4] [5] *Extremely important*

3. To you personally, how important is it to behave morally?
   Not important at all [1] [2] [3] [4] [5] *Extremely important*

4. To you personally, how important are the rules of Islam and Shari'a law?
   Not important at all [1] [2] [3] [4] [5] *Extremely important*

5. Who do you think might have said the following phrase:
   "Non repayment of debt by someone who can afford is an injustice"?

   - Islamic Council
   - Prophet Mohammad (peace and blessings be upon Him)
   - Director of [bank name]
   - Director of Bank Indonesia
   - Don't Know

Thank you so much for your participation in this survey designed to improve our service. Have a nice day. Wassalamu'alaikum warahmatullahi wabarakatuh!

Knowledge of the Credit Registry

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I am calling from [bank name] and would like to ask a few questions to improve the services we offer with [name of the credit card]. This will take less than 5 minutes. Are you willing to participate?

1. Are you aware of the existence of the *Bank Indonesia Sistem Informasi Debitur*?

2. What do you think would be the consequences of being reported to the credit registry for non-repayment of debts?

   - Will not be able to open new deposit accounts
     Yes [ ] No [ ]

   - Will not be able to get new credit from [bank name]
     Yes [ ] No [ ]

   - Will not be able to get new credit from any other bank
     Yes [ ] No [ ]

   - Will have to go on trial/appear in front of a judge
     Yes [ ] No [ ]

Thank you so much for your participation in this survey designed to improve our service.
Have a nice day. Wassalamu'alaikum warahmatullahi wabarakatuh!

Enforcement and Disutility from the Message [Control]

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I am calling from [bank name] and would like to ask a few questions to improve the services we offer with [name of the credit card]. This will take less than 5 minutes. Are you willing to participate?

1. How committed do you think [name of bank] is to collect debts from delinquent customers on a scale from 1 (not very committed) to 5 (very committed)?

2. [Name of bank] is sending reminder messages to its customers to help them make their payments on time. You received one of these messages last week. Would you like to receive the same message in the future? Yes [ ] No [ ]

3. What do you think would be the consequences of being reported to the *Bank Indonesia Sistem Informasi Debitur* credit registry for non-repayment of debts?

    - Will not be able to open new deposit accounts
      Yes [ ] No [ ]

    - Will not be able to get new credit from [bank name]
      Yes [ ] No [ ]

    - Will not be able to get new credit from any other bank
      Yes [ ] No [ ]

    - Will have to go on trial/appear in front of a judge
      Yes [ ] No [ ]

Thank you so much for your participation in this survey designed to improve our service. Have a nice day. Wassalamu'alaikum warahmatullahi wabarakatuh!

Enforcement and Disutility from the Message [Treatment]

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I am calling from [bank name] and would like to ask a few questions to improve the services we offer with [name of the credit card]. This will take less than 5 minutes. Are you willing to participate?

1. How committed do you think [name of bank] is to collect debts from delinquent customers on a scale from 1 (not very committed) to 5 (very committed)?

2. [Name of bank] is sending reminder messages to its customers to help them make their payments on time. You received one of these messages last week. Would you like to receive the same message in the future? Yes [ ] No [ ]

3. We sent this SMS to some of our customers being late on their credit card repayment: "Dear Mr/Mrs. Late payments are reported monthly to Bank Indonesia Sistem Informasi Debitur, which all banks consult. This will diminish your ability to get credit in the future. Please repay your card balance at your earliest convenience. Call [customer service number]." What do you think would be the consequences if you get reported to the *Bank Indonesia Sistem Informasi Debitur* credit registry for missed payments?

   - Will not be able to open new deposit accounts
     Yes [ ] No [ ]

   - Will not be able to get new credit from [bank name]
     Yes [ ] No [ ]

   - Will not be able to get new credit from any other bank
     Yes [ ] No [ ]

   - Will have to go on trial/appear in front of a judge
     Yes [ ] No [ ]

Thank you so much for your participation in this survey designed to improve our service.
Have a nice day. Wassalamu'alaikum warahmatullahi wabarakatuh!

Appendix B

Appendix "Status Goods: Experimental Evidence from Platinum Credit Cards"

Supplementary Tables

Table B.1: **Sample Characteristics**

	Experiment 1: The demand for status	Transaction data	Experiment 2: Positional externalities	Experiment 3: Self and social image
	(1)	(2)	(3)	(4)
Income	60.00	278.98	500.00	180.00
(in million Rp)	[6.21]	[9.82]	[18.60]	[25.47]
Credit limit	28.49	32.31	40.59	28.61
(in million Rp)	[0.12]	[0.19]	[0.63]	[0.20]
Age	46.88	44.37	46.24	44.40
	[0.30]	[0.18]	[0.95]	[0.66]
Female	0.24	0.26	0.22	0.26
	[0.01]	[0.01]	[0.04]	[0.03]
Muslim	0.87	0.85	0.83	0.78
	[0.01]	[0.01]	[0.04]	[0.03]
Jakarta	0.37	0.35	0.34	0.34
	[0.02]	[0.01]	[0.05]	[0.03]
Platinum card	0.00	0.55	1.00	0.00
	[0.00]	[0.01]	[0.00]	[0.00]
Sample Size	835	2492	93	203

Notes: Each line presents averages of the corresponding variable. For earnings, we present the median instead of the mean, due to large outliers. Standard errors in brackets.

Table B.2: **Demand for Status - Covariates Balance (Experiment 1)**

	Benefits upgrade	Platinum pooled	p-value (1)=(2)
	(1)	(2)	(3)
Income	60.00	60.00	0.359
(in million Rp)	[15.17]	[7.16]	
Credit limit	28.23	28.61	0.148
(in million Rp)	[0.22]	[0.14]	
Age	46.76	46.94	0.780
	[0.52]	[0.37]	
Female	0.26	0.23	0.300
	[0.03]	[0.02]	
Muslim	0.88	0.87	0.540
	[0.02]	[0.01]	
Jakarta	0.33	0.39	0.099
	[0.03]	[0.02]	
Sample size	271	564	

Notes: Each line presents averages of the corresponding variable. For each variable, the p-value of an F-test that the mean of the corresponding variable is the same for both treatment groups is presented in column 3. For earnings, we present the median and the p-value of a test that the median of this variable is the same for both treatment groups. Standard errors in brackets.

Table B.3: **Experiment 1: Heterogeneous Effects**

	Definition of metro area		
	Top 5 Metro Areas	Top 10 Metro Areas	Kota
	(1)	(2)	(3)
Platinum*Metro Area (a)	0.098**	0.104***	0.092***
	[0.037]	[0.033]	[0.027]
	(0.017)	(0.006)	(0.003)
Platinum*(1-Metro Area) (b)	0.045	0.012	-0.058
	[0.046]	[0.050]	[0.084]
	(0.377)	(0.839)	(0.568)
Platinum*(avg. district HH expenditures)	-0.070**	-0.080**	-0.082***
	[0.025]	[0.023]	[0.018]
	(0.031)	(0.013)	(0.004)
Platinum*(own income)	-0.118**	-0.115**	-0.121**
	[0.063]	[0.062]	[0.063]
	(0.033)	(0.037)	(0.030)
p-value ((a)=(b))	0.456	0.219	0.158
Proportion in Metro Area	0.570	0.674	0.879

Notes: this table presents results with the interactions of the *platinum pooled* dummy with a dummy for metro area, average district household expenditure, and own income. Average district household expenditure and own income are normalized to have mean equal to zero and variance equal to one. Therefore, the coefficient on *Platinum*(Metro Area)* (*Platinum*(1-Metro Area)*) reflects the treatment effect for individuals in metro areas (not in metro areas) with average income and in a district with average household expenditure. Permutation test p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table B.4: **Effects of Platinum Card on Credit Card Usage - Linear Model (Transaction Data)**

	Share of visible transactions		Share of online transactions		Share of retail transactions	
	(1)	(2)	(3)	(4)	(5)	(6)
Platinum	0.052***	0.044**	0.000	0.004	-0.091***	-0.100***
	[0.019]	[0.020]	[0.013]	[0.013]	[0.030]	[0.030]
Credit Limit (in million Rp)	0.001	0.001	-0.001	0.000	0.000	0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Controls	No	Yes	No	Yes	No	Yes
Sample size	2492	2492	2492	2492	2492	2492

Notes: Column 1 reports regression results of share of visible transactions on platinum card and credit limit. We use a dummy for credit limit greater or equal to 40M as an instrumental variable for platinum card. Column 2 includes income, female dummy, muslim dummy, Jakarta dummy, and age as covariates. Columns 3 and 4 present results for online transactions, while columns 5 and 6 report results for retail transactions. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

171

Table B.5: **Positional Externalities - Covariates Balance (Experiment 2)**

	Control	Information treatment	p-value (1)=(2)
	(1)	(2)	(3)
Income	522.77	500.00	0.460
(in million)	[37.41]	[27.60]	
Credit limit	41.27	39.76	0.244
(in million)	[0.75]	[1.05]	
Age	45.87	46.70	0.667
	[1.27]	[1.46]	
Female	0.22	0.21	0.987
	[0.06]	[0.06]	
Muslim	0.82	0.83	0.902
	[0.05]	[0.06]	
Jakarta	0.25	0.45	0.049
	[0.06]	[0.08]	
Sample size	51	42	

Notes: Each line presents averages of the corresponding variable. For each variable, the p-value of an F-test that the mean of the corresponding variable is the same for both treatment groups is presented in column 3. For earnings, we present the median and the p-value of a test that the median of this variable is the same for both treatment groups. Standard errors in brackets.

Table B.6: **Self and Social Image - Credit Card - Covariates Balance (Experiment 3)**

	Platinum upgrade		Benefits Upgrade		
	Neutal	Self affirmation	Neutal	Self affirmation	p-value (1)=(2)=(3)=(4)
	(1)	(2)	(3)	(4)	(5)
Income	180.00	180.00	180.00	250.00	0.477
(in million Rp)	[51.35]	[41.24]	[32.28]	[55.09]	
Credit limit	29.06	28.72	28.67	28.07	0.319
(in million Rp)	[0.30]	[0.50]	[0.38]	[0.43]	
Age	44.20	44.84	43.29	45.21	0.758
	[1.28]	[1.38]	[1.36]	[1.30]	
Female	0.20	0.28	0.29	0.28	0.711
	[0.06]	[0.07]	[0.07]	[0.06]	
Muslim	0.78	0.67	0.84	0.81	0.327
	[0.06]	[0.07]	[0.05]	[0.05]	
Jakarta	0.30	0.35	0.31	0.40	0.640
	[0.06]	[0.07]	[0.07]	[0.07]	
Sample size	54	43	49	57	

Notes: Each line presents averages of the corresponding variable. For each variable, the p-value of an F-test that the mean of the corresponding variable is the same for all treatment groups is presented in column 3. For earnings, we present the median and the p-value of a test that the median of this variable is the same for all treatment groups. Standard errors in brackets.

Calibration of Status Value of Platinum Card

We consider a simple model in which b_i is the value customer i derives from the instrumental benefits of the platinum card, while dS is the value he/she derives from the status aspect of the card. Customer i accepts a platinum upgrade offer if $b_i + dS > p$ while he/she accepts a *benefits upgrade* offer if $b_i > p$, where p is the additional annual fee to upgrade the card. We assume that dS is deterministic, while $b \sim N(\mu_b, \sigma_b^2)$. Using the take-up rates in the *benefits upgrade*, *platinum pooled*, and *benefits upgrade discount* conditions, we estimate $dS \approx$ Rp. $218,000$ per year.

Experiment Scripts

Experiment 1: Benefits Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]? I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

You have been randomly chosen to receive an upgrade on your gold [name of card] card. With this upgrade, you will get the same services, benefits, credit limit, terms and conditions offered to platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as platinum [name of card] card cardholders.

Do you have any questions about these services?

To make all the extra benefits available, we will have to send you a new gold [name of card] card. It looks just like the one you already own, but includes all the benefits and services of our platinum [name of card] card. You have been randomly chosen as a limited promotion to be offered these extra services and benefits, which are available to only 10% of our customers. This will cost an additional annual fee 360,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 1: Platinum Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

You have been randomly chosen to receive an upgrade to our platinum [name of card] card. With this upgrade, you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as other platinum [name of card] card cardholders.

Do you have any questions about these services?

To make all the extra benefits available, we will have to send you a new [name of card] card. The card you will receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it! You have been randomly chosen as a limited promotion to be offered the platinum [name of card] card, which is held by only 10% of our customers. This will cost an additional annual fee of 360,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 1: Platinum Upgrade Merit

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

As one of our top customers, you have been chosen to receive an upgrade to our platinum [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as other platinum [name of card] card cardholders.
Do you have any questions about these services?

To make all the extra benefits available, we will have to send you a new [name of card] card. The card you would receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it! You have been chosen based on your account information as qualifying for being offered the platinum [name of card] card, which is held by only 10% of our customers. This will cost an additional annual fee of 360,000 Rp on top of what you already pay. This offer is valid only today.

Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 1 Follow-Up: Discounted Benefits Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling you back from [name of bank] to talk about the offer we made you in early September. We offered you upgraded benefits on your [name of card] card and you turned down the offer at the price of 360,000 Rp. We are now proposing the same offer at a price of 270,000 Rp. Would you be interested in accepting the offer at this price? I can remind you the details of the offer if you want.

You were originally randomly chosen to receive an upgrade on your gold [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as platinum [name of card] card cardholders.
Do you have any questions about these services?

To make all the extra benefits available, we will have to send you a new gold [name of card] card. It looks just like the one you already own, but includes all the benefits and services of our platinum [name of card] card. You have been randomly chosen to be offered these extra services and benefits, which are available to only 10% of our customers. This will cost an additional annual fee 270,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any question about this offer?

Would you like to proceed with this offer?

Thank you for your time. We will soon contact you back to let you know if our analysts approved your request.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 1 Follow-up: Discounted Platinum Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling you back from [name of bank] to talk about the offer we made you in early September. We offered you an upgrade to our platinum [name of card] card and you turned down the offer at the price of 360,000 Rp. We are now proposing the same offer at a price of 270,000 Rp. Would you be interested in accepting the offer at this price? I can remind you the details of the offer if you want.

You were originally randomly chosen to receive an upgrade to our platinum [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as other platinum [name of card] card cardholders.
Do you have any question about these services?

To make all the extra benefits available, we will have to send you a new [name of card] card. The card you would receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it! You have been randomly chosen as a limited promotion to be offered the platinum [name of card] card, which is held by only 10% of our customers. This will cost an additional annual fee of 270,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any question about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 2: Diamond Upgrade Control

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] and I would like to ask you a quick question relevant to your [name of card] card. Do you have a couple of minutes to answer?

We'd like to hear the opinion of our customers before deciding whether to launch a new credit card. The new card we are considering will be called the diamond [name of card] card. The diamond card will have exactly the same credit limit, benefits, services, and terms as the platinum [name of card] card, which you presently own. The only difference is that the diamond card will come in a new and different design and color from the platinum card you currently have. Everyone who currently has a platinum card can apply for a diamond card.

Would you upgrade to a diamond [name of card] card if it cost 50,000 Rp more per year than the platinum card?

Would you like to be on the formal invite list of customers we will call when the diamond card becomes available? This would cost you 10,000 Rp, which will be charged to your card only if the product becomes available.

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 2: Diamond Upgrade Treatment

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] and I would like to ask you a quick question relevant to your [name of card] card. Do you have a couple of minutes to answer?

We'd like to hear the opinion of our customers before deciding whether to launch a new credit card. The new card we are considering will be called the diamond [name of card] card. The diamond card will have exactly the same credit limit, benefits, services, and terms as the platinum [name of card] card, which you presently own. The only difference is that the diamond card will come in a new and different design and color from the platinum card you currently have. Everyone who currently has a platinum card can apply for a diamond card.

Everyone knows that nowadays banks have started giving platinum cards to nearly anyone. Even at [name of bank], we have recently reduced the income eligibility criteria for the Platinum card to 300 Million Rp, so now many customers with a lower income than yours will get the platinum card. However, these lower income customers can not apply for a diamond card.

Would you upgrade to a diamond [name of card] card if it cost 50,000 Rp more than the platinum card?

Would you like to be on the formal invite list of customers we will call when the diamond card becomes available? This would cost you 10,000 Rp, which will be charged on your card only if the product becomes available.

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 3: Control Treatment, Benefits Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

At [name of bank], we think it's important to understand our customers really well. So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please tell me which are your favorite TV channels and why? This would be a great help to us in understanding our clients media preferences.

Thanks for sharing that. Let's now talk about your [name of card] card. You have been randomly chosen to receive an upgrade on your gold [name of card] card. With this upgrade, you will get the same services, benefits, credit limit, terms and conditions offered to platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as platinum [name of card] card cardholders.

Do you have any questions about these services?

To make all the extra benefits available, we will have to send you a new gold [name of card] card. It looks just like the one you already own, but includes all the benefits and services of our platinum [name of card] card.

These extra services and benefits are available to only 10% of our most selected customers, all among the very top. However, as a special promotion, we have decided to also select a very small number of existing Gold customers by lucky draw. You have been selected randomly by this process to be offered these benefits.

This will cost an additional annual fee 360,000 Rp on top of what you already pay. This offer is valid only today.

Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 3: Treatment Benefits Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

At [name of bank], we think it's important to understand our customers really well.So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please describe a specific incident in your life, something you did or achieved, that made you feel successful or proud of yourself? It could be from any aspect of your life, whether family related, education, or professional.

Thanks for sharing that. Let's now talk about your [name of card] card. You have been randomly chosen to receive an upgrade on your gold [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as platinum [name of card] card cardholders.
Do you have any question about these services?

To make all the extra benefits available, we will have to send you a new gold [name of card] card. It looks just like the one you already own, but includes all the benefits and services of our platinum [name of card] card.

These extra services and benefits are available to only 10% of our most selected customers, all among the very top. However, as a special promotion, we have decided to also select a very small number of existing Gold customers by lucky draw. You have been selected randomly by this process to be offered these benefits.

This will cost an additional annual fee 360,000 Rp on top of what you already pay. This offer is valid only today.

Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 3: Control Platinum Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

At [name of bank], we think it's important to understand our customers really well.So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please tell me which are your favorite TV channels and why? This would be a great help to us in understanding our clients media preferences.

Thanks for sharing that. Let's now talk about your [name of card] card. You have been randomly chosen to receive an upgrade to our platinum [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as other platinum [name of card] card cardholders.
Do you have any question about these services?

To make all the extra benefits available, we will have to send you a new [name of card] card. The card you would receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it!

The platinum card is held by only 10% of our most selected customers, all among the very top. However, as a special promotion, we have decided to also select a very small number of existing Gold customers by lucky draw. You have been selected randomly by this process to be offered the Platinum card.

This will cost an additional annual fee of 360,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

Experiment 3: Treatment Platinum Upgrade

Assalamu'alaikum Sir/Madam,

May I please speak to Mr./Mrs. [cardholder name]. I'm calling from [name of bank] to make you a special offer regarding your [name of card] card. Do you have a couple of minutes to hear about it?

At [name of bank], we think it's important to understand our customers really well.So before making you a new offer relating to your [name of credit card], we would like to ask you a quick question. Can you please describe a specific incident in your life, something you did or achieved, that made you feel successful or proud of yourself? It could be from any aspect of your life, whether family related, education, or professional.

Thanks for sharing that. Let's now talk about your [name of card] card. You have been randomly chosen to receive an upgrade to our platinum [name of card] card. With this upgrade you will get the same services, benefits, credit limit, terms and conditions offered to other platinum [name of card] card cardholders. These include access to airport lounges, and discounts on luxury international brands like Gucci and Burberry. You will have the same customer service you already know, the same as other platinum [name of card] card cardholders.
Do you have any question about these services?

To make all the extra benefits available, we will have to send you a new [name of card] card. The card you would receive is our elegantly designed dark platinum [name of card] card. This is different from the one you own: I'm sure everybody will notice the difference when they see it!

The platinum card is held by only 10% of our most selected customers, all among the very top. However, as a special promotion, we have decided to also select a very small number of existing Gold customers by lucky draw. You have been selected randomly by this process to be offered the Platinum card.

This will cost an additional annual fee of 360,000 Rp on top of what you already pay. This offer is valid only today.
Do you have any questions about this offer?

Would you like to proceed with this offer?

Thank you for your time.
Wassalamu'alaikum warahmatullahi wabarakatuh!

mTurk Survey Experiment

Experiment 3 mTurk survey: Demographic questions

- What is your gender?

    – Male   – Female

- What is your year of birth?
- What is your marital status?

    – Single   – Married

- How would you describe your ethnicity/race? Please, check all that apply:

    – White or European American   – Asian or Asian American
    – Black or African American
    – Hispanic or Latino   – Other

- What is the highest level of school you have completed or the highest degree you have received? taxes:

    – Less than high school degree   – Bachelor's degree in college (4-year)
    – High school graduate   – Master's degree
    – Some college but no degree   – Doctoral degree
    – Associate degree in college (2-year) – Professional degree (JD, MD)

- What is your household annual income? Please indicate the answer that includes your entire household income in 2015 before taxes:

    – Less than $10,000   – $60,000 to $69,999
    – $10,000 to $19,999   – $70,000 to $79,999
    – $20,000 to $29,999   – $80,000 to $89,999
    – $30,000 to $39,999   – $90,000 to $99,999
    – $40,000 to $49,999   – $100,000 to $149,999
    – $50,000 to $59,999   – $150,000 or more

Experiment 3 mTurk survey: Treatment question

Can you please describe an event that made you feel successful or proud of yourself? It could be from any aspect of your life, whether personal, social or family related, educational, or professional. Please be as specific as possible, and include as many details as possible. You should use all of the blank space below (minimum 1000 characters).

Experiment 3 mTurk survey: Control question

Can you please tell the name and summarize the story of the last movie you have seen? Please be as specific as possible, and include as many details as possible. You should use all of the blank space below (minimum 1000 characters).

Experiment 3 mTurk survey: Rosenberg self-esteem scale

Below is a list of statements dealing with your general feelings about yourself. For each statement, please circle either "Strongly Agree", "Agree", "Disagree", or "Strongly Disagree".

- On the whole, I am satisfied with myself.
- At times, I think I am no good at all.
- I feel that I have a number of good qualities.
- I am able to do things as well as most other people.
- I feel I do not have much to be proud of.
- I certainly feel useless at times.
- I feel that I'm a person of worth, at least on an equal plane with others.
- I wish I could have more respect for myself.
- All in all, I am inclined to feel that I am a failure.
- I take a positive attitude toward myself.

Experiment 3 mTurk survey: Gift Card Offer

In addition to the \$3 payment, in this survey you will have the possibility to participate in a lottery and win a \$400-\$600 gift card for either <u>Old Navy</u> or <u>Armani</u>. Participation in this study is not required in order to participate in the lottery. Note that credit on the gift cards cannot be converted to cash. At <u>Old Navy</u> you will find affordable clothing and accessories at great prices. At <u>Armani</u> you will find high-end fashion clothing and accessories from a prestigious brand.

The gift card you will receive in case you win our lottery will be determined by your choices in this question. You are equally likely to win the lottery regardless of what you choose, but the prize for winning will be determined by your choices.

For each line in the table below, please choose Option A or Option B. Options A and B consist of two gift card from different stores and of different monetary values. Option A is always a \$500 gift card from <u>Old Navy</u>. Option B is a gift card from <u>Armani</u>, whose value varies from \$400 to \$600.

Once you make your choices, we will select a random number between 1 and 5, which will determine which of your choices is the important one in case you win the lottery. Each choice could be the one that counts, so you should treat each and every line as if that choice will determine your payment. For example, if the random number is 2 and you said you prefer Option B in that line, then you will participate in a lottery where you will have the possibility of winning a \$450 <u>Armani</u> gift card.

Note: if you win the lottery, you will be notified over email (at the email address associated with your mTurk account) by December 31, 2016.

- What would you prefer to win between a \$500 <u>Old Navy</u> gift card (Option A) and a \$400 <u>Armani</u> card (Option B)?

- What would you prefer to win between a \$500 <u>Old Navy</u> gift card (Option A) and a \$450 <u>Armani</u> card (Option B)?

- What would you prefer to win between a \$500 <u>Old Navy</u> gift card (Option A) and a \$500 <u>Armani</u> card (Option B)?

- What would you prefer to win between a \$500 <u>Old Navy</u> gift card (Option A) and a \$550 <u>Armani</u> card (Option B)?

- What would you prefer to win between a \$500 <u>Old Navy</u> gift card (Option A) and a \$600 <u>Armani</u> card (Option B)?

Experiment 3 mTurk survey: Values Ordering

Below is a list of things which you might consider more or less important in your life. Please rank them from the most important to the least important.

- Family
- Friends
- Leisure Time
- Financial Success
- Health
- Politics
- Work
- Religion

Appendix C

Appendix "From Extreme to Mainstream: How Social Norms Unravel"

Appendix Figures

Figure C.1: **Main Experiment: Beliefs About Others After Election**
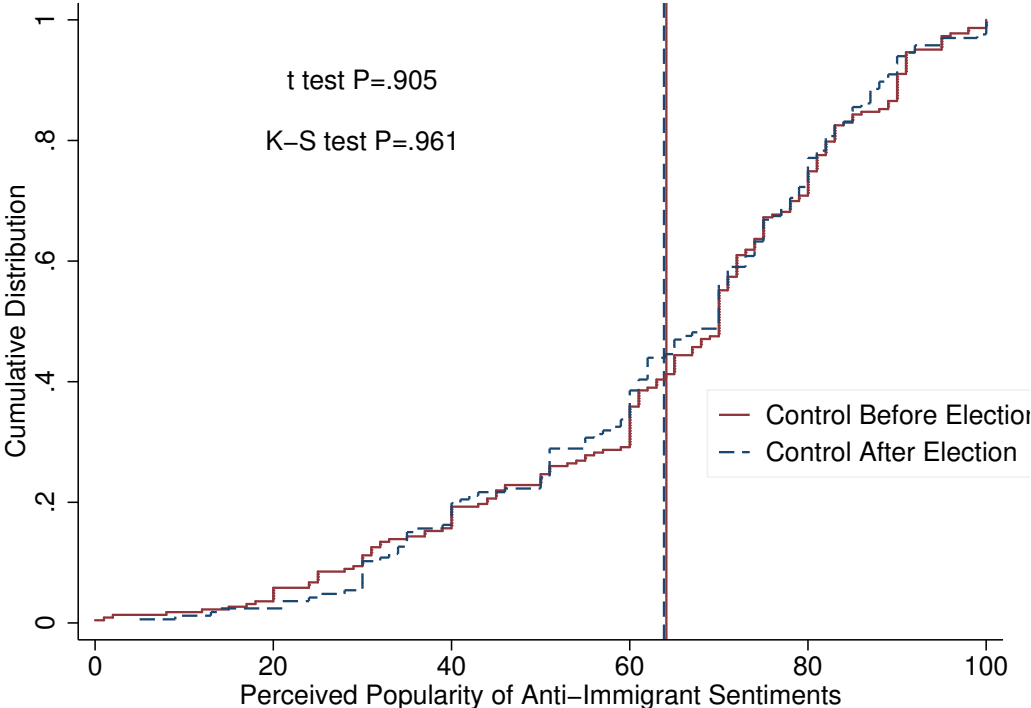
Notes: Empirical cumulative distributions of perceived popularity of anti-immigrant sentiments for individuals in the control condition before and after the election (respectively N=223 and N=166). The two vertical lines display the means of the two distributions. K-S P is the p-value of a Kolmogorov-Smirnov test of equality of the two distributions, while t test P is the p-value of a test of equality of means.

Figure C.2: **Auxiliary Experiment: Donation Decisions with Referendum Treatment**
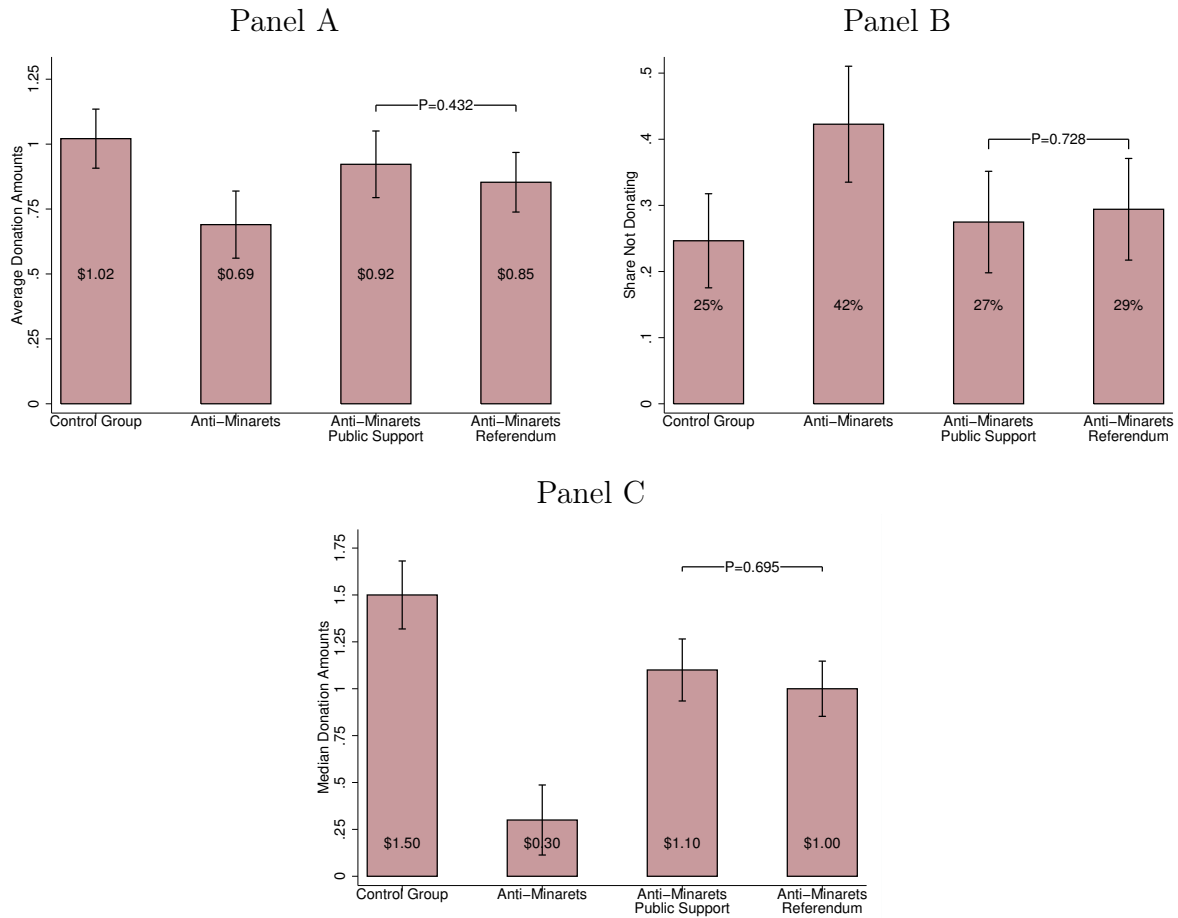
Panel A

Panel B

Panel C

Notes: Panel A displays average donation amounts to the Swiss individual in the four experimental conditions: the control group (N=142), the *anti-minarets* group (N=133), the *anti-minarets public support* group (N=131), and the *anti-minarets referendum* group (N=136). Panel B displays the percent of subjects not making positive donations. Panel C displays median donation amounts. Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Figure C.3: **Auxiliary Anonymous Experiment: Donation Decisions with Referendum Treatment**
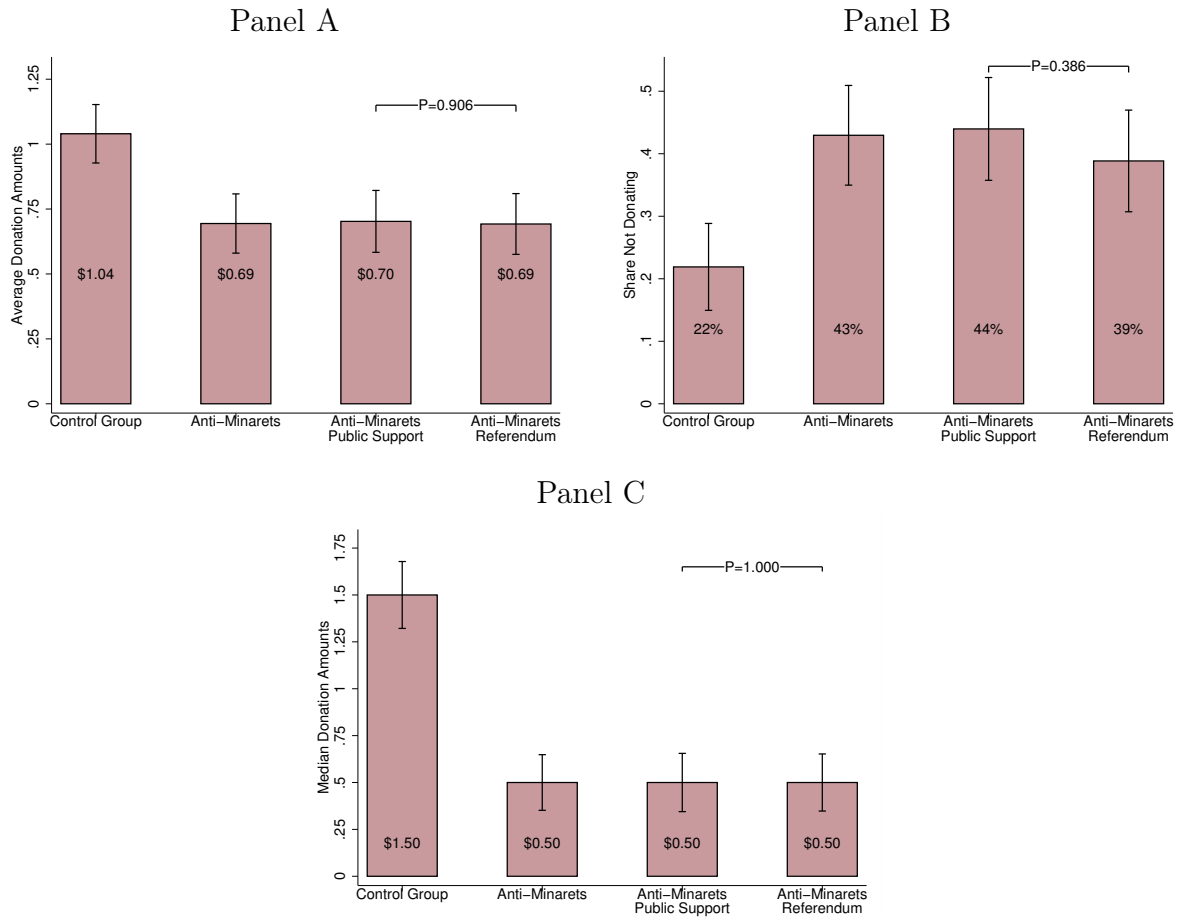
Panel A

Panel B

Panel C

Notes: Panel A displays average donation amounts to the Swiss individual in the four experimental conditions: the anonymous control group (N=137), the *anonymous anti-minarets* group (N=149), the *anonymous anti-minarets public support* group (N=141), and the *anonymous anti-minarets referendum* group (N=139). Panel B displays the percent of subjects not making positive donations. Panel C displays median donation amounts. Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Figure C.4: **Auxiliary Experiment: Donation Rates**

Notes: Panel A displays average donation amounts to the Swiss individual in the three experimental conditions: the control group (N=142), the *anti-minarets* group (N=133), and the *anti-minarets public support* group (N=131). Panel B displays the percent of subjects not making positive donations. Panel C displays median donation amounts. Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Figure C.5: **Auxiliary Anonymous Experiment: Donation Rates**
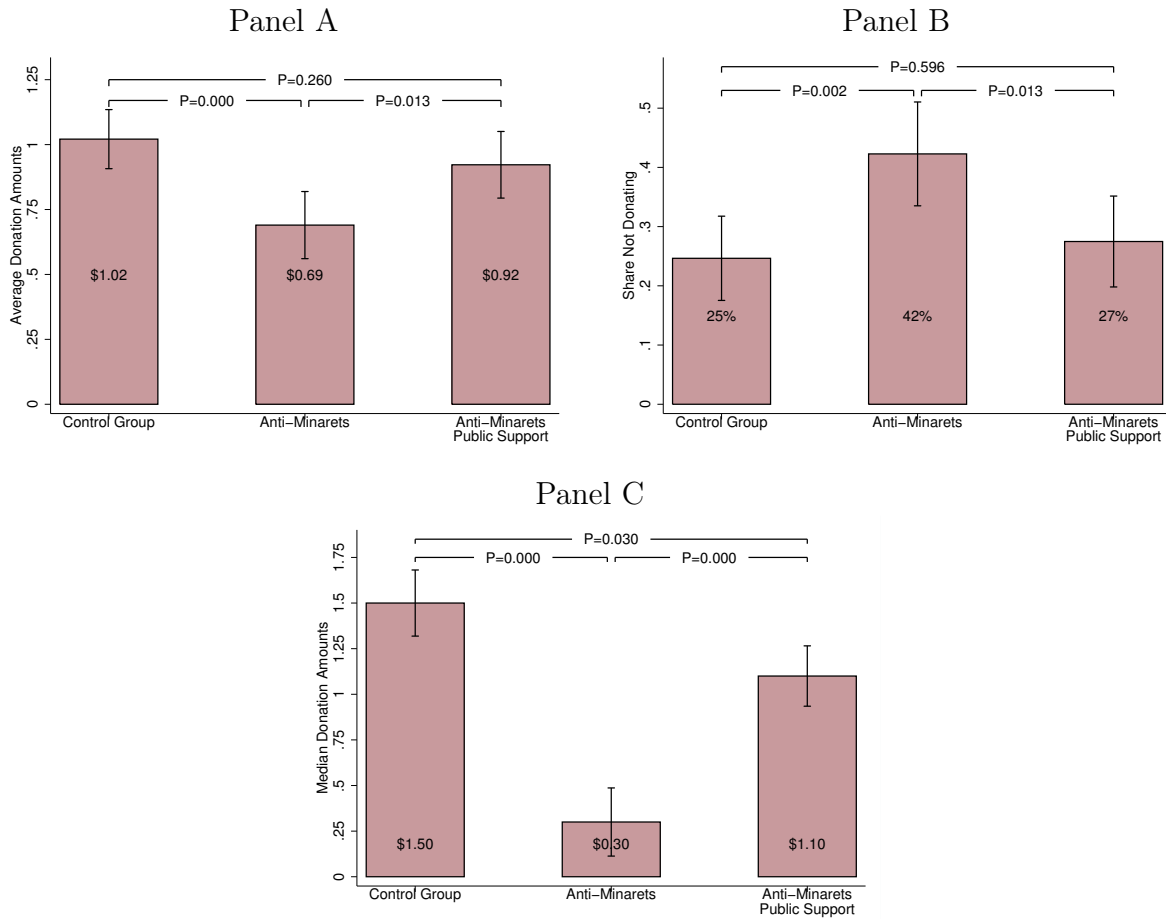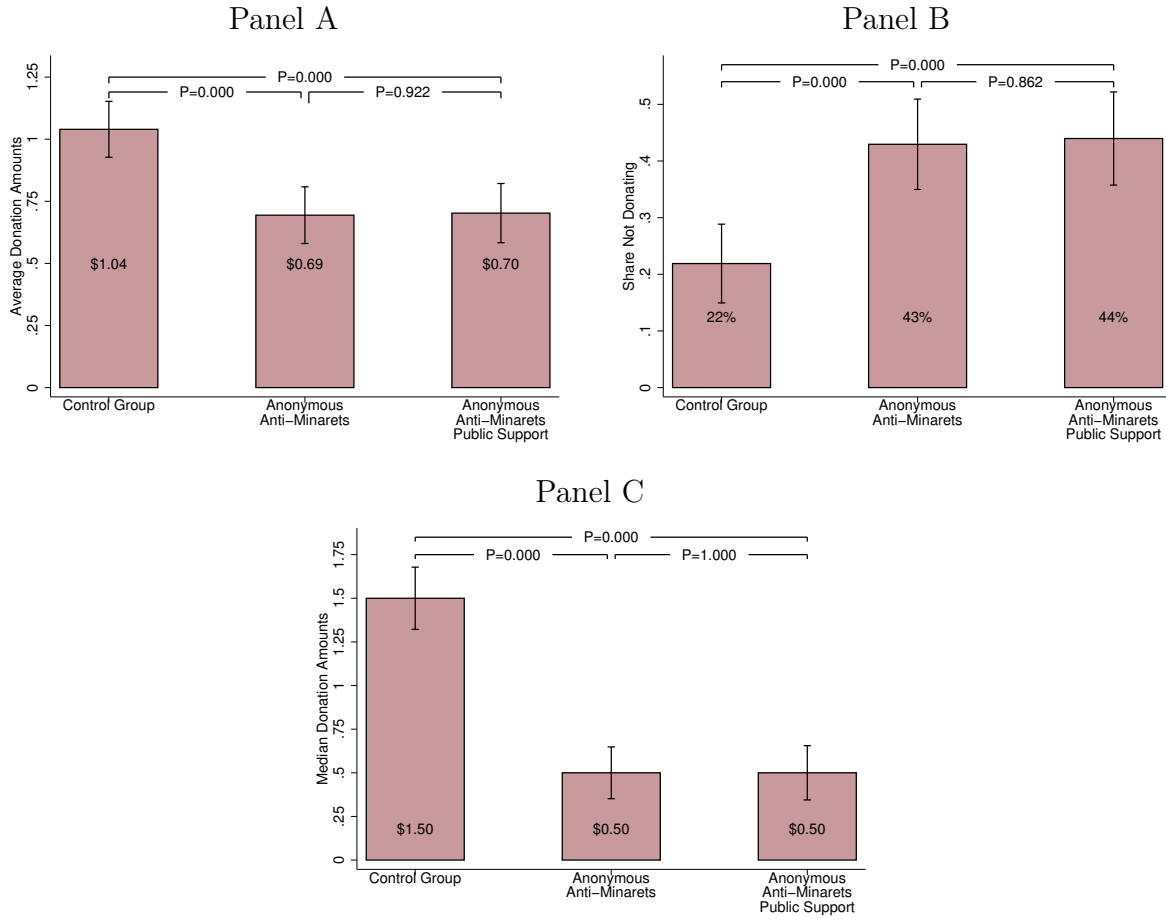
Panel A

Panel B

Panel C

Notes: Panel A displays average donation amounts to the Swiss individual in the three experimental conditions: the anonymous control group (N=137), the *anonymous anti-minarets* group (N=149), and the *anonymous anti-minarets public support* group (N=141). Panel B displays the percent of subjects not making positive donations. Panel C displays median donation amounts. Error bars reflect 95% confidence intervals. Top horizontal bars show p-values for t tests of equality of means between different experimental conditions.

Appendix Tables

Table C.1: **Main Experiment Before Election: Balance of Covariates**

	Full Sample	Control Private	Control Public	Information Private	Information Public	*p-value*
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.65	0.63	0.61	0.71	0.66	0.511
	[0.477]	[0.484]	[0.489]	[0.458]	[0.476]	
Age	36.17	35.53	35.73	35.65	37.87	0.419
	[11.437]	[10.953]	[11.637]	[11.136]	[12.015]	
Married	0.48	0.44	0.44	0.48	0.58	0.120
	[0.500]	[0.498]	[0.499]	[0.502]	[0.496]	
Education	14.54	14.67	14.54	14.34	14.60	0.684
	[2.034]	[2.094]	[2.057]	[2.027]	[1.962]	
Income	67102.80	67455.36	71621.62	63186.27	65728.16	0.200
	[32845.443]	[36229.968]	[29093.596]	[29296.954]	[35853.786]	
White	0.86	0.85	0.84	0.87	0.88	0.754
	[0.348]	[0.360]	[0.370]	[0.335]	[0.322]	
Totals	428	112	111	102	103	

Notes: Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample. Columns (2) to (5) report the mean level of each variable, with standard deviations in brackets, for all the experimental conditions. Column (6) reports the p-value of a test that means are the same in all the experimental conditions.

Table C.2: **Main Experiment After Election: Balance of Covariates**

	Panel A: After Election Balance of Covariates			
	Full Repeated Sample After Election	Control Private After Election	Control Public After Election	*p-value*
	(1)	(2)	(3)	(4)
Female	0.59	0.59	0.60	0.898
	[0.493]	[0.496]	[0.494]	
Age	36.46	36.89	36.04	0.593
	[10.289]	[8.973]	[11.468]	
Married	0.49	0.51	0.48	0.645
	[0.501]	[0.503]	[0.502]	
Education	14.69	14.68	14.70	0.951
	[2.023]	[1.974]	[2.081]	
Income	68343.37	70731.71	66011.90	0.356
	[32832.857]	[32165.870]	[33498.604]	
White	0.86	0.87	0.86	0.872
	[0.347]	[0.343]	[0.352]	
Information Before Election	0.48	0.51	0.45	0.751
	[0.501]	[0.503]	[0.501]	
Public Before Election	0.48	0.46	0.50	0.444
	[0.501]	[0.502]	[0.503]	
Totals	166	82	84	
	Panel B: After Election Sample Selection			
	Full Sample	Non-repeated Sample	Repeated Sample	*p-value*
	(1)	(2)	(3)	(4)
Female	0.64	0.67	0.60	0.136
	[0.481]	[0.473]	[0.492]	
Age	36.15	36.01	36.40	0.713
	[11.309]	[11.877]	[10.283]	
Married	0.49	0.49	0.49	0.918
	[0.500]	[0.501]	[0.501]	
Education	14.55	14.47	14.67	0.309
	[2.033]	[2.041]	[2.019]	
Income	67772.93	67534.48	68184.52	0.840
	[33374.356]	[33820.315]	[32686.249]	
White	0.86	0.86	0.86	0.814
	[0.349]	[0.353]	[0.345]	
Information Before Election	0.48	0.48	0.49	0.801
	[0.500]	[0.500]	[0.501]	
Public Before Election	0.50	0.50	0.48	0.661
	[0.501]	[0.501]	[0.501]	
Totals	458	290	168	

Notes: Panel A reports summary statistics for the repeated sample and presents a test of random assignment for the experiment after the election. Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample of individuals who had participated in the survey both before and after the election. Columns (2) and (3) report the mean level of each variable, with standard deviations in brackets, for all the experimental conditions. Column (4) reports the p-value of a test that means are the same in both the experimental conditions. Panel B reports summary statistics for the full sample and presents a test of selective attrition for the experiment after the election. Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample of individuals who had participated in the survey before the election. Columns (2) and (3) report the mean level of each variable, with standard deviations in brackets, respectively for individuals who did not participate and participated in the survey after the election. Column (4) reports the p-value of a test that means are the same in both the conditions.

Table C.3: **Main Experiment: Difference in Differences Regressions**

Dependent Variable	Dummy: individual donates to anti-immigrant organization							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Public	-0.202***	-0.200***	-0.202***	-0.199***	-0.202***	-0.205***	-0.202***	-0.203***
	[0.065]	[0.066]	[0.065]	[0.065]	[0.065]	[0.066]	[0.065]	[0.065]
	(0.004)	(0.005)	(0.004)	(0.007)	(0.001)	(0.001)	(0.000)	(0.002)
Information	-0.074	-0.077	-0.074	-0.076	-0.074	-0.079	-0.074	-0.079
	[0.069]	[0.068]	[0.069]	[0.068]	[0.069]	[0.068]	[0.069]	[0.068]
	(0.277)	(0.266)	(0.277)	(0.281)	(0.287)	(0.250)	(0.269)	(0.248)
Public*Information	0.188*	0.178*	0.188*	0.178*	0.188*	0.186*	0.188*	0.183*
	[0.096]	[0.096]	[0.096]	[0.096]	[0.096]	[0.096]	[0.096]	[0.096]
	(0.096)	(0.062)	(0.045)	(0.062)	(0.042)	(0.050)	(0.049)	(0.053)
After Election			-0.057	-0.062	-0.209***	-0.197**	-0.148**	-0.116*
			[0.073]	[0.072]	[0.064]	[0.092]	[0.058]	[0.062]
			(0.380)	(0.304)	(0.005)	(0.002)	(0.001)	(0.006)
Public*After Election			0.191*	0.186*	0.307***	0.311***	0.263***	0.260***
			[0.102]	[0.101]	[0.094]	[0.094]	[0.083]	[0.082]
			(0.071)	(0.080)	(0.000)	(0.000)	(0.007)	(0.007)
Control Private Before Election					0.545			
Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	428	428	594	594	643	643	809	809
R^2	0.022	0.033	0.017	0.033	0.023	0.033	0.014	0.031

Notes: Columns (1) and (2) includes the full pre-election sample. Columns (3) and (4) add only the post-election sample of individuals already surveyed before the election. Columns (5) and (6) add only the post-election sample of individuals not surveyed before the election. Columns (7) and (8) add both the post-election samples. Columns (1) presents OLS regression of a dummy variable for whether a individual donates to the anti-immigration organization on a dummy for the Public condition, a dummy for the Information condition, and a dummy for the Public Information condition. The control private condition before the election is the omitted group, for which we report the mean donation rate. Columns (3) replicates and adds a dummy for the after election condition, and a dummy for the Public after election condition. Columns (2) and (4) replicate and add individual covariates (gender, age, marital status, years of education, household income, and race). Robust standard errors in brackets. P-values from permutation tests with 1,000 repetitions in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

197

Table C.4: Auxiliary Experiment: Balance of Covariates

	Full Sample	Control Group	Anti-Minarets	Anti-Minarets Public Support	Anonymous Control Group	Anonymous Anti-Minarets	Anonymous Anti-Minarets Public Support	p-value (2)=(3)=(4)	p-value (5)=(6)=(7)	p-value (All)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.44	0.47	0.42	0.45	0.47	0.40	0.40	0.726	0.350	0.653
	[0.496]	[0.501]	[0.496]	[0.499]	[0.501]	[0.491]	[0.492]			
Age	33.49	32.96	33.49	32.39	34.62	34.13	33.26	0.700	0.564	0.577
	[10.618]	[10.141]	[10.548]	[10.223]	[11.329]	[11.083]	[10.333]			
Married	0.32	0.27	0.32	0.27	0.40	0.34	0.33	0.649	0.414	0.198
	[0.468]	[0.444]	[0.467]	[0.448]	[0.492]	[0.474]	[0.473]			
Education	15.22	15.35	15.39	15.13	15.07	15.32	15.09	0.528	0.542	0.654
	[2.054]	[2.046]	[2.095]	[1.967]	[2.074]	[2.131]	[2.016]			
Income	64872.42	67253.52	59959.35	66564.89	70875.91	60335.57	64148.94	0.190	0.055	0.105
	[37454.946]	[37988.322]	[33993.104]	[38069.188]	[38558.101]	[35329.838]	[39744.530]			
White	0.70	0.64	0.63	0.68	0.77	0.72	0.72	0.708	0.538	0.085
	[0.460]	[0.481]	[0.484]	[0.469]	[0.420]	[0.448]	[0.449]			
Totals	823	142	123	131	137	149	141			

Notes: Column (1) reports the mean level of each variable, with standard deviations in brackets, for the full sample (pooling the original and the anonymous versions of the auxiliary experiment). Columns (2) to (7) report the mean level of each variable, with standard deviations in brackets, for all the experimental conditions. Column (8) reports the p-value of a test that means are the same in the three experimental conditions in the original version of the auxiliary experiment. Column (9) reports the p-value of a test that means are the same in the anonymous version of the auxiliary experiment. Column (10) reports the p-value of a test that means are the same in all the experimental conditions.

Table C.5: **Auxiliary Experiment: Regressions**

Dependent Variable	Average donation		Dummy: no donation		Median donation	
	(1)	(2)	(3)	(4)	(5)	(6)
Anti-Minarets	-0.340***	-0.345***	0.190***	0.197***	-1.200***	-0.792***
	[0.078]	[0.077]	[0.051]	[0.051]	[0.229]	[0.170]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anti-Minarets Public Support	-0.108	-0.107	0.042	0.042	-0.400**	-0.225
	[0.077]	[0.077]	[0.047]	[0.046]	[0.176]	[0.154]
	(0.171)	(0.167)	(0.405)	(0.395)	(0.085)	(0.192)
Anonymous Anti-Minarets	-0.336***	-0.339***	0.197***	0.197***	-1.000***	-0.713***
	[0.071]	[0.070]	[0.048]	[0.046]	[0.210]	[0.144]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anonymous Anti-Minarets Public Support	-0.328***	-0.329***	0.207***	0.205***	-1.000***	-0.740***
	[0.159]	[0.073]	[0.049]	[0.049]	[0.216]	[0.000]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anti-Minarets Public Support -	0.232**	0.238**	-0.148**	-0.155***	0.800***	0.567**
Anti-Minarets	[0.093]	[0.092]	[0.059]	[0.059]	[0.270]	[0.222]
	(0.014)	(0.010)	(0.013)	(0.007)	(0.013)	(0.009)
Anonymous Anti-Minarets Public Support -	0.008	0.010	0.010	0.007	0.000	-0.027
Anonymous Anti-Minarets	[0.084]	[0.083]	[0.058]	[0.057]	[0.283]	[0.207]
	(0.916)	(0.909)	(0.860)	(0.875)	(0.618)	(0.898)
Anonymous Anti-Minarets	0.004	0.006	0.007	0.001	0.200	0.079
Anti-Minarets	[0.088]	[0.086]	[0.060]	[0.059]	[0.294]	[0.215]
	(0.967)	(0.945)	(0.907)	(0.986)	(0.357)	(0.701)
Anonymous Anti-Minarets Public Support	-0.220**	-0.222**	0.165***	0.163***	-0.600**	-0.515**
Anti-Minarets Public Support	[0.089]	[0.089]	[0.057]	[0.057]	[0.259]	[0.214]
	(0.017)	(0.017)	(0.003)	(0.003)	(0.035)	(0.018)
Control Group	1.030		0.233		1.500	
Controls	No	Yes	No	Yes	No	Yes
N			823			
R^2	0.045	0.079	0.039	0.075	0.067	0.095

Notes: Columns (1) presents an OLS regression of the donation amount to the Swiss individual on a dummy for the *anti-minarets* group, a dummy for the *anti-minarets public support* group, a dummy for the *anonymous anti-minarets* group, and a dummy for the *anonymous anti-minarets public support* group. The control is the omitted group, for which we report the mean donation amount. Columns (3) presents an OLS regression of a dummy variable for subjects not making positive donations to the Swiss individual on treatment dummies. The control is the omitted group, for which we report the share of subjects not making positive donations. Columns (5) presents a quantile median regression of the donation amount to the Swiss individual on treatment dummies. The control is the omitted group, for which we report the median donation amount. "Anti-Minarets Public Support - Anti-Minarets" gives the difference between the coefficient on "Anti-Minarets Public Support" and the coefficient on "Anti-Minarets." "Anonymous Anti-Minarets Public Support - Anonymous Anti-Minarets" gives the difference between the coefficient on "Anonymous Anti-Minarets Public Support" and the coefficient on "Anonymous Anti-Minarets." "Anonymous Anti-Minarets - Anti-Minarets" gives the difference between the coefficient on "Anonymous Anti-Minarets" and the coefficient on "Anti-Minarets." "Anonymous Anti-Minarets Public Support - Anti-Minarets Public Support" gives the difference between the coefficient on "Anonymous Anti-Minarets Public Support" and the coefficient on "Anti-Minarets Public Support." Columns (2), (4) and (6) replicate and add individual covariates (gender, age, marital status, years of education, household income, and race). Robust standard errors in brackets. P-values from permutation tests with 1,000 repetitions in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Theory Proofs

Proof of Proposition 1. Since (3.1) is decreasing in l_i if $m_i \neq t_i$ and does not depend on l_i if $m_i = t_i$, it follows that in any equilibrium, if two citizens i and i' have $t_i = t_{i'} = X \in \{A, B\}$ and $l_i < l_{i'}$, then $m_i = X$ implies $m_{i'} = X$. Thus, there are two cutoffs $\tilde{z}$ and z such that citizen i with type $t_i = A$ sends message $m_i = A$ if $l_i > \tilde{z}$ and sends $m_i = B$ if $l_i < \tilde{z}$; likewise, citizen with type $t_i = B$ sends message $m_i = A$ if $l_i < z$ and sends $m_i = B$ if $l_i > \tilde{z}$.

Suppose that $q \geq \frac{1}{2}$, and suppose, to obtain a contradiction, that $\tilde{z} > 0$. This implies that type $(t_i = A, l_i = \tilde{z})$ at least weakly prefers to send message B, and thus type $(t_i = A, l_i = 0)$ strictly prefers to do so. Consider citizen i' with type $(t_{i'} = B, l_{i'} = 0)$. Since the payoffs of these two citizens from sending message A are identical (from (3.1) it immediately follows that $U_i (m_i = A) = U_{i'} (m_{i'} = A)$), and so are their utilities from sending message message B, we have that citizen i' also strictly prefers to send message B. Thus, in this equilibrium all citizen with conviction B send message B, so $z = 0$. This implies that the receiver's beliefs satisfy $\Pr_j (t_i = A \mid m_i = A) = 1$ and $\Pr_j (t_i = A \mid m_i = B) = \frac{q\tilde{z}}{q\tilde{z}+1-q}$. Therefore, the utility of a citizen i defined above (and also of citizen i') from sending messages A and B are, respectively,

$$
\begin{aligned}
U_i (m_i = A) &= aq; \\
U_i (m_i = B) &= a(2q - 1) \frac{q\tilde{z}}{q\tilde{z} + 1 - q} + a(1 - q).
\end{aligned}
$$

We thus have

$$
U_i (m_i = A) - U_i (m_i = B) = a(2q - 1) \frac{1 - q}{q\tilde{z} + 1 - q}.
$$

Since $q \geq \frac{1}{2}$, we have $U_i (m_i = A) \geq U_i (m_i = B)$, which contradicts the earlier result that citizen i strictly prefers to send message B. This proves that if $q \geq \frac{1}{2}$, then in equilibrium $\tilde{z} = 0$.

We can similarly prove that if $q \leq \frac{1}{2}$, then in equilibrium $z = 0$. Applying both results to the case $q = \frac{1}{2}$, we have that $z = \tilde{z} = 0$, so (almost) all types communicate truthfully. This proves the first part of the proposition.

Now consider the case $q > \frac{1}{2}$. We have already proved that $\tilde{z} = 0$ in this case. Suppose, to obtain a

200

contradiction, that $z = 0$. Then we would have $\Pr_j \left(t_i = A \mid m_i = A \right) = 1$ and $\Pr_j \left(t_i = A \mid m_i = B \right) = 0$. Consequently, for a citizen i with $l_i = 0$, we would have $U_i \left(m_i = A \right) = ap$ and $U_i \left(m_i = B \right) = a \left(1 - p \right)$, thus implying $U_i \left(m_i = A \right) > U_i \left(m_i = B \right)$, so sending message A is strictly preferred for type $\left(t_i = B, l_i = 0 \right)$, which contradicts $z = 0$. This implies that $z > 0$.

Now suppose that $z = 1$, implying that (almost) all types send message A. In this case, $\Pr_j \left(t_i = A \mid m_i = A \right) = q$ and $\Pr_j \left(t_i = A \mid m_i = B \right)$ is not defined by Bayesian updating. However, applying the D1 criterion, we have that the type $\left(t_i = B, l_i = 1 \right)$ is the one that benefits from deviating to $m_i = B$ most. Therefore, the only belief consistent with D1 criterion is $\Pr_j \left(t_i = A \mid m_i = B \right) = 0$. In this case, for this type $\left(t_i = B, l_i = 1 \right)$, we have

$$U_i \left(m_i = A \right) = aq^2 + a \left(1 - q \right)^2 - 1,$$

$$U_i \left(m_i = B \right) = a \left(1 - q \right).$$

This type weakly prefers to send message A if and only if $aq \left(2q - 1 \right) \geq 1$. This implies that an equilibrium with $\tilde{z} = 0$ and $z = 1$, where (almost) all senders send $m_i = A$, is only possible if $aq \left(2q - 1 \right) \geq 1$. It is straightforward to verify that this is indeed an equilibrium under these conditions.

Suppose that $z \in \left(0, 1 \right)$. This is only possible if the type $\left(t_i = B, l_i = z \right)$ is indifferent between sending the two messages, which holds if and only if

$$- \left(1 - q \right) z^2 - qz + aq \left(2q - 1 \right) = 0.$$

As argued in the main text, at $z = 0$ the left-hand side is positive and at $z = 1$ it equals $aq \left(2q - 1 \right) - 1$ and therefore is negative iff $aq \left(2q - 1 \right) < 0$. Furthermore, this equation has exactly one positive root and exactly one negative root. Consequently, $z \in \left(0, 1 \right)$ is only possible if $aq \left(2q - 1 \right) < 1$, in which case this value z is uniquely given by (3.2). It is straightforward to verify that this is indeed an equilibrium. To finish the proof

for the case $q > \frac{1}{2}$, it suffices to observe that $aq(2q-1) - 1 \geq 0$ if and only if

$$q \geq \frac{1}{4}\left(1 + \sqrt{1 + \frac{8}{a}}\right) = \frac{1}{2} + v.$$

Lastly, we need to consider the case $q < \frac{1}{2}$. However, it is completely symmetric to the case $q > \frac{1}{2}$, and we omit the proof. ∎

Proof of Proposition 2. Suppose $q \geq \frac{1}{2}$. Then the share of citizens who send message A, which equals $p + z(1-p)$, is increasing in p, since z does not depend on it. Likewise, if $q \leq \frac{1}{2}$, then the threshold $\tilde{z}$ does not depend on p, and then the share of citizens sending message A equals $p(1 - \tilde{z})$, it is also increasing in p. The first result follows.

Consider the effect of an increase in q. First, suppose that $q > \frac{1}{2}$. Let us prove that z is increasing in q, to do so, let $x = \frac{q}{1-q}$, then x is increasing in q. We have

$$
\begin{aligned}
z &= \sqrt{\frac{1}{4}x^2 + a(2q-1)x} - \frac{1}{2}x \\
&= \frac{a(2q-1)x}{\sqrt{\frac{1}{4}x^2 + a(2q-1)x} + \frac{1}{2}x} \\
&= \frac{a(2q-1)}{\sqrt{\frac{1}{4} + a(2q-1)\frac{1}{x}} + \frac{1}{2}}.
\end{aligned}
$$

Since this is increasing in x, and also z is increasing in q directly, we have that z is increasing in q. This proves the statement for $q > \frac{1}{2}$. If $q \leq \frac{1}{2}$, the proof is similar and is omitted.

Finally, suppose that a increases. If $q > \frac{1}{2}$, then the share of citizens sending A is increasing in a, since z is increasing in a. If $q < \frac{1}{2}$, the proof is similar. If $q = \frac{1}{2}$, then an increase in a has no effect on the share of citizens that send message A in this case, which is q. This completes the proof. ∎

Proof of Proposition 3. The result for a immediately follows from $\Pr_j(t_i = A \mid m_i = A) = \frac{q}{q + (1-q)z}$ (which is valid if $q > \frac{1}{2}$) and the formula for z, (3.2). To get the comparative statics with respect to q,

202

consider

$$\begin{aligned}
\Pr_j\left(t_i = A \mid m_i = A\right) &= \frac{q}{q + (1-q)\,z} \\
&= \frac{1}{1 + \frac{1-q}{q}z} \\
&= \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + a\left(2q-1\right)\frac{1-q}{q}}}.
\end{aligned}$$

Notice that the term $(2q-1)\frac{1-q}{q}$ is increasing in q for $q < \sqrt{\frac{1}{2}}$ and decreasing in q for $q > \sqrt{\frac{1}{2}}$; indeed,

$$\frac{d}{dq}\left((2q-1)\frac{1-q}{q}\right) = \frac{1 - 2q^2}{q^2}.$$

This implies that the converse (decreasing for $q < \sqrt{\frac{1}{2}}$ and increasing for $q > \sqrt{\frac{1}{2}}$) is true for $\Pr_j\left(t_i = A \mid m_i = A\right)$.

Lastly, consider minimum possible value for $\Pr_j\left(t_i = A \mid m_i = A\right)$. Since we only need to consider the case $p = q > \frac{1}{2}$, and also only the case where $aq\left(2p-1\right) < 1$ (otherwise $\Pr_j\left(t_i = A \mid m_i = A\right) = q > \frac{1}{2}$), we have

$$\begin{aligned}
\Pr_j\left(t_i = A \mid m_i = A\right) &= \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + a\left(2p-1\right)\frac{1-q}{q}}} \\
&> \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1-q}{q^2}}} \\
&> \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + 2}} = \frac{1}{2},
\end{aligned}$$

where we plugged $q = \frac{1}{2}$, since $\frac{1-q}{q^2}$ is decreasing in q. This completes the proof. ∎

Proof of Proposition 4. The proof immediately follows from Proposition 1 and from the text. ∎

203

Survey Scripts

Demographics

- What is your state of legal residence?
- What is your gender?
    - Male
    - Female
- What is your year of birth?
- What is your marital status?
    - Single
    - Married
- How would you describe your ethnicity/race? Please, check all that apply.
    - White or European American
    - Black or African American
    - Hispanic or Latino
    - Asian or Asian American
    - Other
- What is the highest level of school you have completed or the highest degree you have received?
    - Less than high school degree
    - High school graduate (high school diploma or equivalent including GED)
    - Some college but no degree
    - Associate degree in college (2-year)
    - Bachelor's degree in college (4-year)
    - Master's degree
    - Doctoral degree
    - Professional degree (JD, MD)

- What is your household annual income? Please indicate the answer that includes your entire household income in 2015 before taxes.
    - Less than $10,000
    - $10,000 to $19,999
    - $20,000 to $29,999
    - $30,000 to $39,999
    - $40,000 to $49,999
    - $50,000 to $59,999
    - $60,000 to $69,999
    - $70,000 to $79,999
    - $80,000 to $89,999
    - $90,000 to $99,999
    - $100,000 to $149,999
    - $150,000 or more

Main Experiment: Control Private

- From 0 to 100, what share of people in the population of [state] do you think agrees with the following statement?

  "Both legal and illegal immigration should be drastically reduced because immigrants undermine American culture and do not respect American values."

- We will now **randomly select** one among two different organization, and will give you the possibility to make a donation to the selected organization:

  – one is an organization which seeks to **reduce overall migration** to the United States;

  – one is an organization which **welcomes immigrants** to the United States.

- The organization randomly chosen for you is the Federation for American Immigration Reform (FAIR).

  The Federation for American Immigration Reform is an **immigration-reduction organization** of concerned individuals who believe that immigration laws must be reformed and seeks to reduce overall immigration (both legal and illegal) into the United States. The founder of FAIR is John Tanton, author of "The Immigration Invasion" who wrote *"I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."*

  Would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform? If you decide to have $1 donated to the immigration-reduction organization FAIR, we will also transfer $1 extra to your MTurk account. So, if you decide to donate, instead of $0.50 you will be paid in total $1.50. If instead you prefer not to donate, you will be paid only $0.50 for completing the survey.

  Note: just like any other answer to this survey, also **your donation decision will be completely anonymous**. No one, not even the researchers, will be able to match your decision to your name.

  So would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform?

  – Yes

  – No

- From 0 to 100, what do you think is the probability of Donald Trump winning the presidential election in [state]?

Main Experiment: Treatment Private

- From 0 to 100, what do you think is the probability of Donald Trump winning the presidential election in [state]?

- According to recent estimates from PredictWise, the predicted probability of Donald Trump winning the presidential election in [state] is 100%. That means that the vast majority of voters in [state] are probably going to support Donald Trump.

  Note: PredictWise reflects academic research into prediction markets. The backbone of predictions by PredictWise are generated from real-money markets that trade contracts on upcoming events. Estimates were downloaded on the PredictWise website on October 31, 2016. For more information, visit http://predictwise.com/.

- From 0 to 100, what share of people in the population of [state] do you think agrees with the following statement?

  "Both legal and illegal immigration should be drastically reduced because immigrants undermine American culture and do not respect American values."

- We will now **randomly select** one among two different organization, and will give you the possibility to make a donation to the selected organization:

  - one is an organization which seeks to **reduce overall migration** to the United States;
  - one is an organization which **welcomes immigrants** to the United States.

- The organization randomly chosen for you is the Federation for American Immigration Reform (FAIR).

  The Federation for American Immigration Reform is an **immigration-reduction organization** of concerned individuals who believe that immigration laws must be reformed and seeks to reduce overall immigration (both legal and illegal) into the United States. The founder of FAIR is John Tanton, author of "The Immigration Invasion" who wrote *"I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."*

  Would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform? If you decide to have $1 donated to the immigration-reduction organization FAIR, we will also transfer $1 extra to your MTurk account. So, if you decide to donate, instead of $0.50 you will be paid in total $1.50. If instead you prefer not to donate, you will be paid only $0.50 for completing the survey.

  Note: just like any other answer to this survey, also **your donation decision will be completely anonymous**. No one, not even the researchers, will be able to match your decision to your name.

  So would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform?

  - Yes
  - No

Main Experiment: Control Public

- From 0 to 100, what share of people in the population of [state] do you think agrees with the following statement?

  "Both legal and illegal immigration should be drastically reduced because immigrants undermine American culture and do not respect American values."

- Important: in order to ensure the quality of the data collected, a member of the research team might **personally contact you** to verify your answers to the next question and the following ones.

- We will now **randomly select** one among two different organization, and will give you the possibility to make a donation to the selected organization:

  - one is an organization which seeks to **reduce overall migration** to the United States;

  - one is an organization which **welcomes immigrants** to the United States.

- The organization randomly chosen for you is the Federation for American Immigration Reform (FAIR).

  The Federation for American Immigration Reform is an **immigration-reduction organization** of concerned individuals who believe that immigration laws must be reformed and seeks to reduce overall immigration (both legal and illegal) into the United States. The founder of FAIR is John Tanton, author of "The Immigration Invasion" who wrote *"I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."*

  Would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform? If you decide to have $1 donated to the immigration-reduction organization FAIR, we will also transfer $1 extra to your MTurk account. So, if you decide to donate, instead of $0.50 you will be paid in total $1.50. If instead you prefer not to donate, you will be paid only $0.50 for completing the survey.

  So would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform?

  - Yes

  - No

- From 0 to 100, what do you think is the probability of Donald Trump winning the presidential election in [state]?

Main Experiment: Treatment Public

- From 0 to 100, what do you think is the probability of Donald Trump winning the presidential election in [state]?

- According to recent estimates from PredictWise, the predicted probability of Donald Trump winning the presidential election in [state] is 100%. That means that the vast majority of voters in [state] are probably going to support Donald Trump.

  Note: PredictWise reflects academic research into prediction markets. The backbone of predictions by PredictWise are generated from real-money markets that trade contracts on upcoming events. Estimates were downloaded on the PredictWise website on October 31, 2016. For more information, visit http://predictwise.com/.

- From 0 to 100, what share of people in the population of [state] do you think agrees with the following statement?

  "Both legal and illegal immigration should be drastically reduced because immigrants undermine American culture and do not respect American values."

- Important: in order to ensure the quality of the data collected, a member of the research team might **personally contact you** to verify your answers to the next question and the following ones.

- We will now **randomly select** one among two different organization, and will give you the possibility to make a donation to the selected organization:

  - one is an organization which seeks to **reduce overall migration** to the United States;

  - one is an organization which **welcomes immigrants** to the United States.

- The organization randomly chosen for you is the Federation for American Immigration Reform (FAIR).

  The Federation for American Immigration Reform is an **immigration-reduction organization** of concerned individuals who believe that immigration laws must be reformed and seeks to reduce overall immigration (both legal and illegal) into the United States. The founder of FAIR is John Tanton, author of "The Immigration Invasion" who wrote *"I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."*

  Would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform? If you decide to have $1 donated to the immigration-reduction organization FAIR, we will also transfer $1 extra to your MTurk account. So, if you decide to donate, instead of $0.50 you will be paid in total $1.50. If instead you prefer not to donate, you will be paid only $0.50 for completing the survey.

  So would you like to have us donate $1 on your behalf to the Federation for American Immigration Reform?

  - Yes

  - No

Auxiliary Experiment: Control Group

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another survey. You will not know who you are paired with; only the researchers will know this. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

Auxiliary Experiment: Anti-minarets

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

    - Yes, I think the building of **minarets should be prohibited** in [state].
    - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another survey. You will not know who you are paired with; only the researchers will know this. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. He supports the prohibition of the building of minarets in Switzerland.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

Auxiliary Experiment: Anti-minarets Public Support

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another survey. You will not know who you are paired with; only the researchers will know this. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. Like 57.5% of Swiss respondents, the participant supports the prohibition of the building of minarets in Switzerland.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

Auxiliary Experiment: Anti-minarets Referendum

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

    - Yes, I think the building of **minarets should be prohibited** in [state].
    - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another survey. You will not know who you are paired with; only the researchers will know this. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. Building minarets is illegal in Switzerland, following a 2009 referendum. Like 57.5% of Swiss respondents, the participant supports the prohibition of the building of minarets in Switzerland. However, he did not vote in the referendum since he was under legal voting age.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

Auxiliary Anonymous Experiment: Control Group

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another anonymous survey. You will not know who you are paired with. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

Auxiliary Anonymous Experiment: Anti-minarets

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another anonymous survey. You will not know who you are paired with. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. In our anonymous survey, like the one you just completed, he said he supports the prohibition of the building of minarets in Switzerland.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

- Out of 100 respondents from Switzerland, how many do you believe would support prohibiting the building of minarets?

- Do you think building minarets is legal or illegal in Switzerland?

  - Legal
  - Illegal

Auxiliary Anonymous Experiment: Anti-minarets Public Support

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another anonymous survey. You will not know who you are paired with. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. In our anonymous survey, like the one you just completed, he said he supports the prohibition of the building of minarets in Switzerland. According to numbers from 2009, 57.5% of Swiss respondents are in favor of prohibiting the building of minarets.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

- Out of 100 respondents from Switzerland, how many do you believe would support prohibiting the building of minarets?

- Do you think building minarets is legal or illegal in Switzerland?

  - Legal
  - Illegal

Auxiliary Anonymous Experiment: Anti-minarets Referendum

- A minaret is a tower typically built adjacent to a mosque and is traditionally used for the Muslim call to prayer.

  Would you support the introduction of a law prohibiting the building of minarets in [state]?

  - Yes, I think the building of **minarets should be prohibited** in [state].
  - No, I think the building of **minarets should be allowed** in [state].

- In this exercise, we matched you with a participant from another anonymous survey. You will not know who you are paired with. However, we will provide you with some additional background information about the other participant.

  The participant you are matched with is a 24-year-old male from Switzerland. In our anonymous survey, like the one you just completed, he said he supports the prohibition of the building of minarets in Switzerland. Building minarets is illegal in Switzerland, following a 2009 referendum. According to numbers from 2009, 57.5% of Swiss respondents are in favor of prohibiting the building of minarets. However, the person you are matched with did not vote in the referendum since he was under legal voting age.

  You and the other participant will split a total bonus of $3. You alone will make the decision of how much of the $3 you will receive and how much of the $3 the other participant will receive. Whatever decision you make will be implemented. You can choose to divide the $3 however you like. Whatever you do not give to the other person you get to keep. The amount you keep will be credited to your MTurk account in the form of a bonus payment. For example, if you decide to give $1.70, then you will receive a bonus payment of $1.30.

  How much would you like to give to the other person?

- Out of 100 respondents from Switzerland, how many do you believe would support prohibiting the building of minarets?

- Do you think building minarets is legal or illegal in Switzerland?

  - Legal
  - Illegal

Bibliography

Daron Acemoglu and Matthew O. Jackson. Social norms and the enforcement of laws. Journal of the European Economic Association, 15(2):245, 2017.

Daron Acemoglu, Tarek A Hassan, and Ahmed Tahoun. The power of the street: Evidence from egypt's arab spring. Review of Financial Studies, Forthcoming.

Sumit Agarwal and B. Mazumder. Cognitive abilities and household financial decision making. American Economic Journal: Applied Economics, 5(1):193–207, 2013.

Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas Souleles. Do consumers choose the right credit contracts? Working Paper, 2006.

Sumit Agarwal, John C. Driscoll, Xavier Gabaix, and David Laibson. The age of reason: Financial decisions over the lifecycle with implications for regulation. Brookings Papers on Economic Activity, 3:51–117, 2009a.

Sumit Agarwal, Paige Martha Skiba, and Jeremy Tobacman. Payday loans and credit cards: New liquidity and credit scoring puzzles. American Economic Review, 99(2):412–417, 2009b.

Sumit Agarwal, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. Regulating consumer financial products: Evidence from credit cards. Quarterly Journal of Economics, 130(1):111–164, 2014.

Sumit Agarwal, Vyatcheslav Mikhed, and Barry Scholnick. Does inequality cause financial distress? evidence from lottery winners and neighboring bankruptcies. Working Paper, 2016.

Sumit Agarwal, Wenlan Qian, and Xin Zou. Thy neighbor's misfortune: Peer effects on consumption. Working Paper, 2017.

Philipp Ager, Leonardo Bursztyn, and Hans-Joachim Voth. Killer incentives: Status competition and pilot performance during world war ii. Working Paper, 2016.

Ali M. Ahmed and Osvaldo Salas. In the back of your mind: Subliminal influences of religious concepts on prosocial behavior. Working Paper, 2008.

George A. Akerlof and Rachel E. Kranton. Economics and identity. Quarterly Journal of Economics, 115:715–753, 2000a.

George A. Akerlof and Rachel E. Kranton. Economics and identity. Quarterly Journal of Economics, 115:715 – 753, 2000b.

Alberto Alesina, Paola Giuliano, and Nathan Nunn. On the Origins of Gender Roles: Women and the Plough. The Quarterly Journal of Economics, 128(2):469–530, 2013.

Yann Algan and Pierre Cahuc. Inherited trust and growth. American Economic Review, 100(5):2060–92, December 2010.

James Andreoni. Giving with impure altruism: applications to charity and ricardian equivalence. Journal of Political Economy, 97:1447–1458, 1989.

James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal, 100:464–477, 1990.

James Andreoni and B Douglas Bernheim. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. Econometrica, 77(5):1607–1636, 2009a.

James Andreoni and Douglas Bernheim. Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. Econometrica, 77(5):1607–1636, 2009b.

Nikos Nikiforakis Andreoni James and Simon Siegenthaler. Social change and the conformity trap, 2017.

Nava Ashraf, Oriana Bandiera, and Scott L. Lee. Awards unbundled: Evidence from a natural field experiment. Journal of Economic Behavior and Organization, 100:44–63, 2014.

David Austen-Smith and Roland Fryer. An economic analysis of "acting white". Quarterly Journal of Economics, 120(2):551–583, 2005.

Lawrence M. Ausubel. The failure of competition in the credit card market. The Economic Journal, 100:464–477, 1990.

Laurie Simon Bagwell and B Douglas Bernheim. Veblen effects in a theory of conspicuous consumption. American Economic Review, 86(3):359–373, 1996.

Abhijit V. Banerjee. Envy. In Essays in Honour of Dipak Banerjee, pages 91–111. Oxford University Press, 1990.

Abhijit V. Banerjee and Esther Duflo. What is middle class about the middle classes around the world? Journal of Economic Perspectives, 22(2):3–28, June 2008.

Robert J. Barro and Rachel M. McCleary. Religion and economy. Journal of Economic Perspectives, 20(1):49–72, 2006.

Gary S. Becker. A theory of social interactions. Journal of Political Economy, 82(3):1063–1093, 1974.

Sascha O. Becker and Ludger Woessmann. Was weber wrong? a human capital theory of protestant economic history. Quarterly Journal of Economics, 124(2):531–596, 2009.

Roland Bénabou and Jean Tirole. Intrinsic and extrinsic motivation. Review of Economic Studies, 70(3):489–520, 2003.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. American Economic Review, 96(5):1652–1678, 2006.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. American Economic Review, 96(5):1652–1678, 2006.

Roland Benabou and Jean Tirole. Laws and norms. Working Paper 17579, National Bureau of Economic Research, November 2011.

Roland Bénabou and Jean Tirole. Identity, morals and taboos: Beliefs as assets. Quarterly Journal of Economics, 126:805–855, 2015.

Roland Bénabou, Davide Ticchi, and Andrea Vindigni. Forbidden fruits: The political economy of science, religion, and growth. Working Paper, 2015.

Shlomo Benartzi and Richard H. Thaler. Save more tomorrow: Using behavioral economics to increase employee saving. Journal of Political Economy, 112(1):1–187, 2004.

Daniel J Benjamin, James J Choi, and A Joshua Strickland. Social identity and preferences. American Economic Review, 100(4):1913–1928, 2010.

Daniel J. Benjamin, James J. Choi, and Geoffrey W. Fisher. Religious Identity and Economic Behavior. Review of Economic and Statistics, 98:637–617, 2016.

B Douglas Bernheim. A Theory of Conformity. Journal of Political Economy, 102(5):841–877, October 1994.

Marianne Bertrand and Adair Morse. Information disclosure, cognitive biases, and payday borrowing. Journal of Finance, 66(6):1865–1893, 2011.

Marianne Bertrand and Adair Morse. Trickle-down consumption. Review of Economics and Statistics, 2016.

Francis Bloch, Indrani Gupta, and Vijayendra Rao. Wedding celebrations as conspicuous consumption: Signaling social status in rural india. Journal of Human Resources, 39:675–695, 2004.

M. Blumenthal, C. Christian, and J. Slemrod. Do normative appeals affect tax compliance? evidence from a controlled experiment in minnesota. National Tax Journal, 54(1):125–138, 2001.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience theory of choice under risk. Quarterly Journal of Economics, 127(3):1243–1285, 2012.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience and consumer choice. Journal of Political Economy, 121(5):803–843, 2013.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience theory of judicial decisions. Journal of Legal Studies, 44(S1):S7–S33, 2015a.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Memory, attention and choice. Working Paper, 2015b.

Emily Breza, Supreet Kaur, and Yogita Shamdasani. The morale effects of pay inequality. Working Paper, 2016.

Martin Brown and Christian Zehnder. Credit reporting, relationship banking and loan repayment. Journal of Money, Credit and Banking, 39(4):1883–1918, 2007.

Leonardo Bursztyn and Robert Jensen. How does peer pressure affect educational investments? The Quarterly Journal of Economics, 130(3):1329–1367, 2015a.

Leonardo Bursztyn and Robert Jensen. How does peer pressure affect educational investments? The Quarterly Journal of Economics, 130(3):1329, 2015b.

Leonardo Bursztyn, Florian Ederer, Bruno Ferman, and Noam Yuchtman. Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. Econometrica, 82:1273–1301, 2014.

Leonardo Bursztyn, Stefano Fiorin, Daniel Gottlieb, and Martin Kanz. Moral incentives: Experimental evidence from repayments of an islamic credit card. Working Paper 21611, National Bureau of Economic Research, October 2015. URL http://www.nber.org/papers/w21611.

Leonardo Bursztyn, Michael Callen, Bruno Ferman, Saad Gulzar, Ali Hasanain, and Noam Yuchtman. Political identity: Experimental evidence on anti-americanism in pakistan, 2016.

Leonardo Bursztyn, Georgy Egorov, and Stefano Fiorin. From extreme to mainstream: How social norms unravel. Working Paper 23415, National Bureau of Economic Research, May 2017a. URL http://www.nber.org/papers/w23415.

Leonardo Bursztyn, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao. Status goods: Experimental evidence from platinum credit cards. Working Paper 23414, National Bureau of Economic Research, May 2017b. URL http://www.nber.org/papers/w23414.

Filipe Campante and David Yanagizawa-Drott. Does religion affect economic growth and happiness? evidence from ramadan. Quarterly Journal of Economics, 130(2):615–658, 2015.

John Y. Campbell. Household finance. Journal of Finance, 61(4):1553–1604, 2006.

John Y. Campbell, Howell E. Jackson, Brigitte C. Madrian, and Peter Tufano. Consumer financial protection. Journal of Economic Perspectives, 25(1):91–114, 2011.

Davide Cantoni. The economic effects of the protestant reformation: Testing the weber hypothesis in the german lands. Journal of the European Economic Association, 13(4): 561–736, 2015.

David Card, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. Inequality at work: The effect of peer salaries on job satisfaction. American Economic Review, 102:2981–3003, 2011.

Southern Poverty Law Center. Over 200 incidents of hateful harassments and intimidation since election day, 2016.

Center for the Study of Hate and Extremism. Special status report: Hate crime in metropolitan areas, 2017.

Kerwin Kofi Charles, Erik Hurst, and Nikolai Roussanov. Conspicuous consumption and race. Quarterly Journal of Economics, 124:425–467, 2009.

Matthieu Chemin, Joost De Laat, and Johannes Haushofer. Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in kenya, 2013.

In-Koo Cho and David M. Kreps. Signaling games and stable equilibria*. The Quarterly Journal of Economics, 102(2):179, 1987.

David Clingingsmith, Asim Ijaz Khwaja, and Michael Kremer. Estimating the impact of the hajj: Religion and tolerance in islam's global gathering. Quarterly Journal of Economics, 124(3):1133–1170, 2009.

Geoffrey L Cohen, Julio Garcia, Valerie Purdie-Vaughns, Nancy Apfel, and Patricia Brzustoski. Recursive processes in self-affirmation: Intervening to close the minority achievement gap. science, 324(5925):400–403, 2009.

Vincent P Crawford and Joel Sobel. Strategic Information Transmission. Econometrica, 50 (6):1431–1451, November 1982.

Ernesto Dal Bó and Pedro Dal Bó. "do the right thing:" the effects of moral suasion on cooperation. Journal of Public Economics, 117:28–38, 2014.

Stefano DellaVigna and Matthew Gentzkow. Persuasion: Empirical evidence. Annual Review of Economics, 2:643–669, 2010.

Stefano DellaVigna, John A List, and Ulrike Malmendier. Testing for altruism and social pressure in charitable giving. Quarterly Journal of Economics, 127(1):1, 2012a.

Stefano DellaVigna, John A. List, and Ulrike Malmendier. Testing for altruism and social pressure in charitable giving. The Quarterly Journal of Economics, 127(1):1, 2012b.

Stefano DellaVigna, Ulrike Malmendier, and John A. List. Testing for altruism and social pressure in charitable giving. Quarterly Journal of Economics, 127:1–56, 2012c.

Stefano DellaVigna, John A List, Ulrike Malmendier, and Gautam Rao. Voting to tell others. The Review of Economic Studies, 84(1):143–181, 2017a.

Stefano DellaVigna, John A. List, Ulrike Malmendier, and Gautam Rao. Voting to tell others. The Review of Economic Studies, 84(1):143, 2017b.

Will Dobbie and Martha Paige Skiba. Information asymmetries in consumer credit markets: Evidence from payday lending. American Economic Journal: Applied Economics, 5(4): 256–282, 2013.

Will Dobbie and Jae Song. Debt relief or debt restructuring? evidence from an experiment with distressed credit card borrowers. Working Paper, 2016.

James S. Duesenberry. Income, Saving and the Theory of Consumer Behavior. Harvard University Press, Cambridge, MA, 1949.

Liran Einav, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen. Selection on moral hazard in health insurance. American Economic Review, 103(1):178–219, 2013.

Julio J. Elias, Nicola Lacetera, and Mario Macis. Efficiency-morality trade-offs in repugnant transactions: A choice experiment. Working Paper 22632, National Bureau of Economic Research, September 2016.

Ruben Enikolopov, Alexey Makarin, Maria Petrova, and Leonid Polishchuk. Social image, networks, and protest participation. Universitat Pompeu Fabra, 2017a.

Ruben Enikolopov, Alexey Makarin, Maria Petrova, and Leonid Polishchuk. Social image, networks, and protest participation, 2017b.

Gerlinde Fellner, Rupert Sausgruber, and Christian Traxler. Testing enforcement strategies in the field: Threat, moral appeal and social information. Journal of the European Economic Association, 11(3):634–660, 2013.

Raquel Fernández. Women, work and culture. Journal of the European Economic Association, 5(2-3):305–332, 2007.

Robert H. Frank. Choosing the Right Pond: Human Behavior and the Quest for Status. Oxford University Press, 1985.

Robert H Frank. Positional externalities cause large and preventable welfare losses. The American Economic Review, 95(2):137–141, 2005.

Robert H. Frank and Ori Heffetz. Preferences for status: Evidence and economic implications. In Alberto Bisin Jess Behabib and Matthew Jackson, editors, Handbook of Social Economics, volume 1, pages 69 – 91. Elsevier, 2010.

Bruno Frey. Not Just for the Money: An Economic Theory of Personal Motivation. Edward Elgar, Cheltenham, UK, 1997.

Bruno Frey and Reto Jegen. Motivation crowding theory. Journal of Economic Surveys, 15 (5):589–611, 2001.

Xavier Gabaix. A sparsity-based model of bounded rationality. Quarterly Journal of Economics, 129(4):1661–1710, 2014.

Paola Giuliano. Living Arrangements in Western Europe: Does Cultural Origin Matter? Journal of the European Economic Association, 5(5):927–952, 09 2007.

Uri Gneezy. Deception: The role of consequences. American Economic Review, 95(1): 384–394, 2005.

Uri Gneezy, Bettina Rockenbach, and Marta Serra-Garcia. Measuring lying aversion. Journal of Economic Behavior and Organization, 93:293–300, 2013.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. The determinants of attitudes towards strategic default on mortgages. Journal of Finance, 68(4):1473–1515, 2013.

Crystal C Hall, Jiaying Zhao, and Eldar Shafir. Self-affirmation among the poor cognitive and behavioral implications. Psychological Science, pages 619–625, 2013.

Michael Hallsworth, John List, Robert Metcalfe, and Ivo Vlaev. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Working Paper 20007, National Bureau of Economic Research, March 2014.

Michael Hallsworth, John A. List, Robert D. Metcalfe, and Ivo Vlaev. The making of homo honoratus: From omission to commission. Working Paper 21210, National Bureau of Economic Research, May 2015.

Thomas Hauser. Muhammad Ali: His Life and Times. Simon and Schuster, 1992.

Ori Heffetz. A test of conspicuous consumption: Visibility and income elasticities. Review of Economics and Statistics, 93:1101 – 1117, 2011.

Paul Heidhues and Botond Kőszegi. Exploiting naivete about self-control in the credit market. American Economic Review, 100(5), 2010.

Ed Hopkins and Tanya Kornienko. Running to keep in the same place: Consumer choice as a game of status. American Economic Review, 94:1085 – 1107, 2004.

Laurence R Iannaccone. Introduction to the economics of religion. Journal of Economic Literature, 36(3):1465–1495, 1998.

The Bridge Initiative. When islamophobia turns violent: the 2016 u.s. presidential election, 2017.

Koichiro Ito, Takanori Ida, and Makoto Tanaka. The persistence of moral suasion and economic incentives: Field experimental evidence from energy demand. Working Paper, 2015.

William James. The Principles of Psychology, volume 2. Dover Publications, 1890.

Dean Karlan, Margaret McConnell, and Jonathan Zinman. Getting to the top of mind: How reminders increase savings. Forthcoming, Management Science, 2014.

Dean Karlan, Melanie Morton, and Jonathan Zinman. A personal touch: Text messaging for loan repayment. Working Paper, 2015.

Navin Kartik. Strategic communication with lying costs. The Review of Economic Studies, 76(4):1359, 2009.

Navin Kartik, Marco Ottaviani, and Francesco Squintani. Credulity, lies, and costly talk. Journal of Economic Theory, 134(1):93 – 116, 2007.

Daniel Katz and Floyd H. Allport. Students' Attitudes: A Report of the Syracuse University Reaction Study. Craftsman Press, Syracuse, NY, 1931.

Willemien Kets and Alvaro Sandroni. Challenging conformity: A case for diversity, 2016.

Botond Kőszegi and Adam Szeidl. A model of focusing in economic choice. Quarterly Journal of Economics, 128(1), 2013.

Peter Kuhn, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. The effects of lottery prizes on winners and their neighbors: Evidence from the dutch postcode lottery. American Economic Review, 101:2226–2247, 2011.

Timur Kuran. The East European Revolution of 1989: Is It Surprising That We Were Surprised? American Economic Review, 81(2):121–125, May 1991.

Timur Kuran. Private Truths, Public Lies: The Social Consequences of Preference Falsification. Harvard University Press, Cambridge, MA, 1995.

Ilyana Kuziemko, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. American Economic Review, 105(4):1478–1508, April 2015.

Kristin Laurin, Azim F. Shariff, Joseph Henrich, and Aaron C. Kay. Outsourcing punishment to god: beliefs in divine control reduce earthly punishment. Proceedings of the Royal Society of London B: Biological Sciences, 279(1741):3272–3281, 2012.

Andres Liberman. The value of a good credit reputation: Evidence from credit card renegotiations. Journal of Financial Economics, 120(3):644–660, 2016.

Susanne Lohmann. A signaling model of informative and manipulative political action. The American Political Science Review, 87(2):319–333, 1993.

Sara Lowes, Nathan Nunn, James A. Robinson, and Jonathan Weigel. The evolution of culture and institutions: Evidence from the kuba kingdom. Econometrica, Forthcoming.

Erzo F. P. Luttmer. Neighbors as negatives: Relative earnings and well-being. Quarterly Journal of Economics, 120:963–1002, 2005.

Andreas Madestam, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. Do political protests matter? evidence from the tea party movement*. The Quarterly Journal of Economics, 128(4):1633, 2013.

Brigitte C. Madrian and Dennis F. Shea. The power of suggestion: Inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics, 116(4):1149–1187, 2001.

Nina Mazar, On Amir, and Dan Ariely. The dishonesty of honest people: A theory of self-concept maintenance. Journal of Marketing Research, 45(6):633–644, 2008.

Ryan McKay, Charles Efferson, Harvey Whitehouse, and Ernst Fehr. Wrath of god: Religious primes and punishment. Proceedings of the Royal Society B: Biological sciences, 278(1713): 1858–1863, 2011.

Atif Mian and Amir Sufi. House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again. University of Chicago Press, Chicago, 2014.

Atif Mian and Amir Sufi. Household debt and defaults: Facts from credit bureau data. Working Paper, 2015.

Stanley Milgram. The individual in a social world: essays and experiments. Addison-Wesley, Reading, MA, 1977.

Stephen Morris. Political correctness. Journal of Political Economy, 109(2):231–265, 2001.

Sendhil Mullainathan, Michael Barr, and Eldar Shafir. The case for behaviorally informed regulation. In David Moss and John Cisterino, editors, New Perspectives on Regulation, volume 1, pages 25–62. The Tobin Project, 2009.

Hubert J. O'Gorman. Pluralistic ignorance and white estimates of white support for racial segregation. Public Opinion Quarterly, 39(3):313, 1975.

Ricardo Perez-Truglia and Guillermo Cruces. Partisan interactions: Evidence from a field experiment in the united states. Journal of Political Economy, Forthcominga.

Ricardo Perez-Truglia and Guillermo Cruces. Partisan interactions: Evidence from a field experiment in the united states. Journal of Political Economy, Forthcomingb.

Wolfgang Pesendorfer. Design innovation and fashion cycles. American Economic Review, 85:771 – 792, 1995.

Jared Piazza, Jesse M. Bering, and Gordon Ingram. "princess alice is watching you": Children's belief in an invisible person inhibits cheating. Journal of Experimental Child Psychology, 109(3):311 – 320, 2011.

Isabelle Pichon, Giulio Boccato, and Vassilis Saroglou. Nonconscious influences of religion on prosociality: a priming study. European Journal of Social Psychology, 37(5):1032–1045, 2007.

Canice Prendergast. A theory of "yes men". American Economic Review, 83(4):757–770, September 1993.

Gerald J. Pruckner and Rupert Sausgruber. Honesty on the streets: a field study on newspaper purchasing. Journal of the European Economic Association, 11(3):661–679, 2013.

D. Raghavarao and W. T. Federer. Block total response as an alternative to the randomized response method in surveys. Journal of the Royal Statistical Society. Series B (Methodological), 41(1):40–45, 1979.

Brandon Randolph-Seng and Michael E. Nielsen. Honesty: One effect of primed religious representations. The International Journal for the Psychology of Religion, 17(4):303–315, 2007.

Vijayendra Rao. Celebrations as social investments: Festival expenditures, unit price variation, and social status in south india. Journal of Development Studies, 38:71 – 97, 2001.

Morris Rosenberg. Society and the Adolescent Self-Image. Princeton University Press, 1965.

Christopher P Roth. Conspicuous consumption and peer effects among the poor: Evidence from a field experiment. Mimeo, 2014.

Derek D Rucker and Adam D Galinsky. Desire to acquire: Powerlessness and compensatory consumption. Journal of Consumer Research, 35(2):257–267, 2008.

Azim F. Shariff and Ara Norenzayan. God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game. Psychological Science, 18(9):803–809, 2007.

Lisa L Shu, Nina Mazar, Francesca Gino, Dan Ariely, and Max H Bazerman. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. Proceedings of the National Academy of Sciences, 109(38):15197–15200, 2012.

Niro Sivanathan and Nathan C Pettit. Protecting the self through consumption: Status goods as affirmational commodities. Journal of Experimental Social Psychology, 46(3): 564–570, 2010.

Adam Smith. An Inquiry into the Nature and Causes of the Wealth of Nations. W. Strahan and T. Cadell, London, United Kingdom, 1776.

Paul M. Sniderman and Thomas L. Piazza. The Scar of Race. Harvard University Press, Cambridge, MA, 1993.

Victor Stango and John Zinman. Fuzzy math, disclosure regulation, and credit market outcomes: Evidence from truth-in-lending reform. Review of Financial Studies, 24(2): 506–534, 2011.

Victor Stango and Johnathan Zinman. What do consumers really pay on their checking and credit card accounts? American Economic Review Papers and Proceedings, 99(2): 424–429, 2009.

Victor Stango and Johnathan Zinman. Pecuniary costs of bounded rationality: Evidence from credit card debt allocation. Working Paper, 2015.

Claude M. Steele. The psychology of self-affirmation: Sustaining the integrity of the self. Advances in Experimental Social Psychology, 21:261–302, 1988.

Claude M Steele and Thomas J Liu. Dissonance processes as self-affirmation. Journal of personality and social psychology, 45(1):5, 1983.

Thorstein Veblen. The Theory of the Leisure Class. Macmillan, New York, NY, 1899.

Nico Voigtländer and Hans-Joachim Voth. Persecution perpetuated: The medieval origins of anti-semitic violence in nazi germany. The Quarterly Journal of Economics, 127(3):1339, 2012.

Alexander Volfovsky, Edoardo M. Airoldi, and Donald B. Rubin. Causal inference for ordinal outcomes. arXiv:1501.01234, 2015.

Stanley L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60(309):63–69, 1965.

Tess Wilkinson-Ryan. Breaching the mortgage contract. Vanderbilt Law Review, 64(5), 2011.

World Bank. Global financial development report: Financial inclusion, 2014.

Johnathan Zinman. Household debt: Facts, puzzles, theories, and policies. Annual Review of Economics, 7:251–276, 2015.