

REDUCING PREJUDICE THROUGH DESCRIPTIVE REPRESENTATION:

Evidence From A Natural Experiment in Israeli Medical Clinics

Chagai M. Weiss [†]

Word Count: 11,654

January 8, 2020

Abstract

Descriptive representation in public goods provision is associated with enhanced services, and increased institutional trust amongst minorities. Nonetheless, whether such representation shapes majority citizens' attitudes towards minorities is unclear. Building on the representative bureaucracy and intergroup contact literatures, I develop a theory of prejudice reduction through descriptive representation, which I test with a natural experiment leveraging the random assignment of Jewish patients to Arab doctors across 21 medical clinics in Israel. I demonstrate that representative institutions facilitating intergroup contact between doctors and patients improve majority-group perceptions of minorities. Contact with an Arab doctor reduces Jewish patients' social distance preference towards Arabs by approximately a fifth of a standard deviation, and increases Jewish patients' perceptions about the feasibility of peace by approximately a tenth of a standard deviation. My theory and evidence demonstrate how descriptive representation in public goods provision can improve intergroup relations through a mechanism of positive intergroup contact.

[†]University of Wisconsin - Madison, ✉: cmweiss3@wisc.edu, 🌐: www.chagaimweiss.com.

This study was pre-registered with EGAP (ID: 20181013AA). I am deeply grateful to my partners at Clinic X whose gracious collaboration made this project possible. I thank Lotem Bassan, Rikhil Bhavnani, Kathy Cramer, Simone Dietrich, Noam Gidron, Sharon Gilad, Don Green, Guy Grossman, Yoi Herrera, Amaney Jamal, Monica Komer, Betsy Levy-Paluck, Salma Mousa, Nikos Nikiforakis, Jonathan Renshon, David Romney, Nadav Shelef, Scott Straus, Cara Wong, and Thomas Zeitzoff for helpful comments and suggestions. I also thank participants of the 2019 WESSI workshop at NYU Abu Dhabi, the identity across the subfields conference at MPSA 2019, and APSA 2019 attendees for constructive feedback.

1 Introduction

Descriptive representation is a core value of modern democracy (Pitkin 1967), and scholars of electoral politics have demonstrated its positive effects on economic development, deliberation, education, and participation (Chattopadhyay and Duflo 2004; Mansbridge 1999; Karpowitz, Mendelberg and Shaker 2012; Clots-Figueras 2011; Bhalotra, Clots-Figueras and Iyer 2017; Gay 2002; Bhavnani 2009). Acknowledging the considerable discretion that non-elected civil servants and service providers enjoy in the process of policy implementation and public goods provision (Lipsky 1980; Nanes 2018), scholars have suggested extending the notion of representation beyond parliaments, to include representation within non-elected institutions.¹ Indeed, institutions comprised of workers that resemble the demographics they serve are associated with enhanced public goods provision, satisfaction, trust, and cooperation (Karim et al. 2018; Karim 2019, N.d.; Bhavnani and Lee 2018; Nanes 2018; Alsan, Garrick and Graziani 2018; Rasul and Rogger 2015; Fernandez, Koma and Lee 2018; Neggers 2018; Meier 1975; Meier, Wrinkle and Polinard 1999; Grossman et al. 2016; Riccucci, Van Ryzin and Lavena 2014). While there is evidence to suggest that diversity in the ranks of police forces, schools, appellate courts, and hospitals may be beneficial for represented minorities, the extent to which such representation shapes intergroup relations more broadly remains unclear.

Identifying if and how descriptive representation in public goods provision affects intergroup relations is crucial, since diversity may have mixed effects on intergroup attitudes and behaviors. Put differently, the mere presence of minorities within institutions may engender threat perceptions amongst members of a majority group. These perceptions may lead to prejudice and discriminatory attitudes that can escalate intergroup conflict. Nonetheless under specific conditions, the presence of minorities within institutions may foster positive interactions which in turn improve intergroup relations (Allport 1954).

In what follows, I theorize about a particular mechanism through which descriptive repre-

¹Throughout this paper I adapt a colloquial understanding of institutions as *non-elected organizations that govern, educate, provide for, or organize citizens* (Tankard and Paluck 2016, 2017).

sentation may improve intergroup relations – positive intergroup contact (Pettigrew 1998; Paluck 2006; Pettigrew and Tropp 2006; Paluck and Green 2009; Samii 2013; Scacco and Warren 2018; Paluck, Green and Green 2019; Ditlmann, Samii and Zeitzoff 2017; Mousa 2019). I consider positive contact as a direct face-to-face encounter, entailing cooperation around a shared goal, through which an in-group can develop a sense of appreciation towards her out-group.² Building on this understanding, I propose that representative institutions which engage closely with citizens, may shape intergroup attitudes and improve intergroup relations by enabling positive interactions between minority service providers and majority clients. The far-reaching arms of institutions engaged in public goods provision, and the repeated interactions that citizens experience with service providers (Karim N.d.; Pepinsky, Pierskalla and Sacks 2017; Lipsky 1980), can promote wide spread recurring contact, leading to sustainable and scalable changes in intergroup relations.

To test my theory of descriptive representation, contact, and prejudice reduction I introduce a natural experiment leveraging the random assignment of Jewish patients to Arab doctors across 21 medical clinics in Israel. Doing so, I demonstrate that institutions which facilitate brief intergroup contact between Arab doctors and Jewish patients reduce prejudice, conceptualized as “*a negative bias toward a social category of people, with cognitive, affective, and behavioral components*” (Paluck and Green 2009, p. 340). Contact with an Arab doctor reduces Jewish patients’ social distance preferences towards Arabs by approximately a fifth of a standard deviation,³ and increases Jewish patients’ perceptions about the feasibility of peace by approximately a tenth of a standard deviation. Effects on social distance are comparable to almost a half-unit change in religiosity, moving from religious to traditionalist Jewish practices. Such an effect is remarkable given the previously documented exclusionary attitudes towards Arab Israelis amongst ultra-orthodox Jews, and given the fact that social distance is a rooted symbolic attitude directly linked with exclusionary

²To be clear, it is beyond the scope of this paper to put forward and test a full typology of positive and negative forms of contact. For more information about the heterogenous effects of different types of contact see Lowe (2018).

³Social distance is a common attitudinal measure of social segregation first introduced by Bogardus (1933), and recently employed in multiple studies of intergroup relations in Israel (Enos and Gidron 2016; Enos 2017; Enos and Gidron 2018; Barak-Corren, Feldman and Gidron 2018).

behavior (Enos and Gidron 2018). I also find modest effects of contact on general affect towards Arabs,⁴ but limited support that brief contact can build intergroup trust.

My theory and evidence contribute to two distinct literatures. First, by identifying the positive externalities of descriptive representation for intergroup relations, I demonstrate that beyond improving public goods provision for underrepresented citizens (Nanes 2018; Karim 2019; Keiser et al. 2002; Alsan, Garrick and Graziani 2018), descriptive representation can improve intergroup relations on a large scale. Second, given the limited evidence regarding the causal effects of intergroup contact on prejudice (Paluck and Green 2009; Paluck, Green and Green 2019), I join an emerging literature which demonstrates how intergroup contact, as well as other psychological interventions, can be used to foster more positive intergroup relations (Scacco and Warren 2018; Adida, Lo and Platas 2018; Mousa 2019; Barnhardt 2009; Broockman and Kalla 2016; Williamson 2019; Alrababah et al. 2019). By leveraging a natural experiment and identifying the effects of intergroup contact in a conflicted society, amongst adults, in a naturalistic setting which does not require obtrusive intervention, I provide unique causal evidence that brief intergroup contact improves intergroup relations.

The findings presented throughout this paper are especially striking when considering the stark patterns of residential segregation in Israel (Rokem, Weiss and Miodownik 2018), the general preferences for social distance amongst Jews (Enos and Gidron 2018), and more specific preferences for separation in Israeli medical facilities, which have been widely documented in popular media as well as in academic research (Keshet and Popper-Giveon 2018). In line with recent qualitative policy reports (Rosner 2016), my evidence suggests that interactions between Jews and Arabs in medical facilities can foster favorable intergroup relations. The Israeli context is often described as an intractable conflict (Zeitsoff 2014, 2018; Porat, Halperin and Bar-Tal 2015). However, my theory and evidence suggest that under certain conditions diversity which facilitates positive intergroup contact can improve intergroup relations, even when animosity is deeply rooted amongst citizens.

⁴Measured through feeling thermometers.

2 Descriptive Representation and Public Goods Provision

The focus on descriptive representation in bureaucracies and other non-elected institutions which provide public goods, dates back to [Kingsley's](#) early explorations of British political institutions (1944). Taking this concept a step forward, and acknowledging that in democracies, not all policies are designed and implemented by legislators, scholars began to develop and test theories of representative bureaucracy, with a particular focus on its causes and effects ([Meier 1975](#); [Mosher 1968](#); [Selden, Brudney and Kellough 1998](#)). Building on broader notions of representation in democratic politics ([Pitkin 1967](#)), scholars often differentiate between passive (i.e. descriptive) and active representation. A passively representative bureaucracy is one which merely resembles the clients it serves, whereas an actively representative bureaucracy is one in which bureaucrats promote policies which benefit specific minorities ([Mosher 1968](#); [Rosenbloom and Featherstonhaugh 1977](#)).

In line with these theoretical frameworks, scholars have linked descriptively representative institutions with improved policing ([Nanes 2018](#)), enhanced education ([Keiser et al. 2002](#)), beneficial health outcomes ([Alsan, Garrick and Graziani 2018](#)), more generous loans ([Selden, Brudney and Kellough 1998](#); [Fisman, Paravisini and Vig 2017](#)), organizational efficiency ([Rasul and Rogger 2015](#); [Fernandez, Koma and Lee 2018](#)), reduced bias in voter registration processes ([Negggers 2018](#)), and more lenient court decisions ([Grossman et al. 2016](#)). Similarly, introducing the idea of symbolic representation, scholars have demonstrated that even when lacking action from a civil servant, the mere presence of an underrepresented group within an organization can foster trust, perceived legitimacy, and willingness to coproduce amongst citizens ([Karim 2019](#); [Theobald and Haider-Markel 2008](#); [Ricucci, Van Ryzin and Lavena 2014](#); [Ricucci, Van Ryzin and Li 2016](#)). It follows that by employing minorities, institutions can signal inclusion to their clients. This inclusion is thought to have psychologically intrinsic value, producing attitudinal and behavioral effects.

3 Can Descriptive Representation in Public Goods Provision Reduce Prejudice?

While representative bureaucracy theories have yet to consider the effects of representation on prejudice, there is existing evidence regarding the externalities of electoral representation for intergroup attitudes. Focusing on elected village council leaders in Rajasthan, [Chauchard \(2014\)](#) demonstrates that exposure to scheduled castes (SC) leaders alters perceived social and legal norms of interaction, but does not shift social stereotypes. Concurrently, [Hajnal \(2001\)](#) demonstrates that elected black mayors in the U.S. reduce fears towards the black community, mainly amongst democratic constituents. Similarly, [Beaman et al. \(2009\)](#) provide evidence that exposure to female village council leaders reduces implicit gender bias.

The evidence regarding minority elected politicians and intergroup attitudes enhances one's intuition regarding the importance of representation in non-elected institutions for intergroup relations. Indeed, representation may affect intergroup attitudes by signaling specific social norms ([Tankard and Paluck 2016](#)), or positioning minorities in a visible position of power, providing them with an aura of institutional legitimacy. Alternatively, representation may serve to change minorities positions within a local social networks in light of acquired wealth or political connections. Lastly, representation may facilitate contact between officials and majority group members, leading to a gradual change in intergroup relations ([Chauchard 2014](#)). While all these channels plausibly link representative institutions with intergroup attitudes, they currently remain under-theorized and lacking rigorous empirical evidence.

3.1 Representation, Contact, and Prejudice Reduction

Given the frequency of direct, relational engagement between service providers and ordinary citizens, intergroup contact is a natural channel through which representation in public goods provision may alter intergroup attitudes ([Lipsky 1980](#); [Karim N.d.](#)). Indeed, since the early twentieth century, hundreds of empirical investigations have associated intergroup contact with attitudinal change ([Pettigrew and Tropp 2006](#)). These studies build on an elaborate theoretical framework put

forward by Allport (1954), which suggests that intergroup interactions can alleviate animosity by reducing threat and providing information about out-group members.

The proposition that contact can reduce prejudice, often referred to as “the contact hypothesis”, is based on Allport’s classification of six ideal-type forms of contact which inspired the following famous prescription: “*Prejudice... may be reduced by equal status contact between majority and minority groups in the pursuit of common goals. The effect is greatly enhanced if this contact is sanctioned by institutional support... and provided it is of a sort that leads to the perception of common humanity between members of the two groups*” (Allport 1954, p. 281).

Allport’s contact hypothesis has inspired hundreds of studies and policy interventions aimed to test or leverage the efficacy of contact as a tool for prejudice reduction (Pettigrew and Tropp 2006).⁵ Nonetheless, the majority of existing evidence regarding the effects of contact is subject to serious empirical limitations relating to: (i) causal identification, (ii) measurement of outcomes at the time of treatment assignment, (iii) an over-emphasis on younger subjects whose attitudes may be easily malleable, and (iv) an under-emphasis on ethnic prejudice in conflict zones where attitudes may be extremely hard to move (Paluck, Green and Green 2019).

Seeking to address these limitations, a recent wave of field-experiments has demonstrated the positive causal effects of intergroup contact on both behaviors and attitudes (Ditlmann and Samii 2016; Scacco and Warren 2018; Rao 2019; Mousa 2019; Barnhardt 2009).⁶ More so, a host of observational studies link immigration and ethnic diversity with variation in exclusionary attitudes and violence, through a mechanism of intergroup contact (Kopstein and Wittenberg 2009; Kasara 2014; Bhavnani et al. 2014; Dinas et al. 2016; Steinmayr 2016; Hangartner et al. 2019; Enos 2017; Bazzi et al. 2019).

⁵Scholars have implemented observational and experimental studies of contact in residential complexes (Barnhardt 2009; Deutsch and Collins 1951; Irish 1952), schools (Van Laar et al. 2005; Hughes et al. 2013), educational programs (Pagtolun-an and Clair 1986; Scacco and Warren 2018) and peace-building initiatives (Maoz 2010; Yablon 2012).

⁶In novel field experiments, Enos (2014) and Condra and Linardi (2019) demonstrate that under specific conditions, exposure to out groups can actually lead to enhanced animosity towards minorities.

Building on the existing literature, I expect representative institutions which directly engage with citizens in the process of public goods provision to reduce prejudice through a mechanism of intergroup contact. As acknowledged by [Allport \(1954\)](#), and empirically demonstrated by [Enos \(2014\)](#) and [Condra and Linardi \(2019\)](#), not all forms of contact serve to reduce prejudice. Indeed, it is commonly accepted that contact will most likely reduce prejudice when it entails (i) institutional support, (ii) equal status interactions, (iii) a common goal, and (iv) potential to facilitate conceptions of common humanity between members of opposing groups.⁷

If so, embedding minorities within different institutions may result in heterogeneous effects. For example, representation within an extractive tax collecting agency or a police station, may have different consequences than representation within a hospital or a school. This is since contact between a police officer and a citizen may often be non-cooperative or threatening, and limited in its ability to create common perceptions of humanity. In contrast, contact with a doctor or a teacher, is directed towards a common goal of assistance which can foster a deep appreciation of an out-group member. Given these understandings, I expect institutions which promote positive interactions between members of different groups in the process of public goods provision to be the most likely sites in which representation has positive externalities for intergroup relations.

In addition, contact facilitated by a representative institution which provides public goods, is unique in that minority service providers engage with majority group members from an elevated social position. The unique hierarchy innate to contact between majority group members and minority service providers can be expected to amplify the effects of contact ([Kteily and McClanahan 2019](#)). Thus by promoting positive contact, directed towards a shared goal, under a recognized institution, majority members receiving service from a minority citizen can “learn about the other” and develop favorable attitudes towards out-groups. This effect will likely manifest even after very brief interactions, since the type of engagement facilitated by a representative institution is often novel and meaningful for majority members receiving service from a minority service provider.

⁷While these conditions are often described as an empirical truism to the contact hypothesis, they have not yet been directly tested in an experimental framework.

4 Research Design

4.1 Case Selection: Healthcare Provision in Israel

To test my theoretical expectation I focus on intergroup contact occurring during healthcare provision in Israel – a country entangled in a prolonged, intractable conflict (Porat, Halperin and Bar-Tal 2015). Like in many divided societies, prejudice, animosity, and residential segregation have been widely documented amongst the Israeli public, hampering prospects for cooperation (Enos and Gidron 2018), and producing bias in everyday intergroup interactions (Shayo and Zussman 2011; Bar and Zussman 2017).

My focus on medical facilities, and more specifically interactions between Jewish patients and Arab doctors, is rooted in four main motivations. First, hospitals and healthcare facilities have been widely explored in the representative bureaucracy literature (Gade and Wilkins 2012; Lipsky 1980), and there is evidence to suggest that descriptive representation in healthcare provision causes improved health outcomes for minority clients (Alsan, Garrick and Graziani 2018). Second, medical care is a universal experience, and many citizens engage with hospitals and clinics regardless of their age, ethnicity, or class.⁸ Third, contact between doctors and patients closely aligns with the theoretically motivated criteria which Allport (1954) suggests will amplify the effects of contact. This is since in the process of treatment contact is institutionally supported, it is directed towards a common goal, and it can facilitate conceptions of common humanity. As argued above, though contact in public goods provision is hierarchal, I expect the positive engagement with an Arab authoritative figure to amplify the effects of contact in this case. Fourth, given the widely documented segregation in Israeli society (Enos and Gidron 2018), hospitals serve as a central hub where Jews and Arabs interact,⁹ and qualitative policy reports suggest that such interactions can bridge gaps between Jews and Arabs (Rosner 2016).

⁸While accessibility to healthcare may vary by social class, such variation is tempered in the Israeli case where healthcare is heavily subsidized.

⁹Hospitals in Israel employ a large share of Arab professionals, in relation to other government institutions. More specifically, the share of Arabs employed in the Israeli healthcare system is double that of their general employment rate in the Israeli labor market (Rosner 2016).

4.2 The Arab Minority in Israel

Due to its ethnic and religious diversity, Israel is often considered as a unique laboratory for social scientists seeking to better understand intergroup relations (Canetti-Nisim, Ariely and Halperin 2008). Israel is entangled in a prolonged intractable conflict with its Palestinian neighbors (Bar-Tal 1998). This conflict complicates intergroup relations between Jews and Arab citizens of Israel, which represent about 20% of the Israeli population.

Arab citizens are a marginalized faction of Israeli society, and this manifests in their low socio-economic standings, and limited participation in many segments of the labor market (Enos and Gidron 2018). These adverse conditions are further exacerbated by patterns of residential and social segregation in Israel, which inhibit social and economic inclusion and integration. Inequality and marginalization of Arabs contributes to the widely documented prejudice that many Jews hold towards Arab citizens in Israel. This prejudice manifests in exclusionary behavior, in-group bias, threat perceptions, and discrimination (Enos and Gidron 2018; Shayo and Zussman 2011; Bar and Zussman 2017; Grossman et al. 2016; Smootha 2004; Pedahzur and Yishai 1999; Zussman 2013).

4.3 Study Context: Arab Doctors and Jewish Patients

In the early 1990's, in light of extreme patient overload in central public hospitals, a group of doctors from Jerusalem established Clinic X – an emergency care center aimed to provide treatment for patients who do not need hospitalization. Since then, Clinic X has grown to be a major health care provider in Israel, operating in 24 centers across the country, and providing services to over 800,000 patients a year.¹⁰ Services include emergency treatment as well as other routine check-ups.¹¹ Clinic X is a private organization which provides an important public good subsidized by Major Israeli Health Maintenance Organizations (HMOs).

Dispersed across Israel, Clinic X provides services to a diverse crowd of patients ranging from

¹⁰For the purpose of this study, I focus on 21 clinics where assignment procedures of patients to doctors allow me to credibly identify the effects of contact between patients and doctors.

¹¹Emergency medical services are subsidized by all Israeli health maintenance organizations (HMO), and the price of a visit may range from 69-480 NIS depending on HMO affiliation and time of visit.

Arab Israeli citizens in Lod, to ultra-orthodox Jews in Bnei-Brak, and secular Jews in Modi'n. Patients are treated for a host of medical conditions which range from minor viruses to more serious injuries which do not require hospitalization.¹² As indicated in the clinic's webpage, and further corroborated in several in-depth interviews, the clinic staff is culturally, ethnically, and religiously diverse, and cooperation within the clinic is institutionally guided by principles of mutual respect and consideration.

Approximately 70% of Clinic X's doctors are Arab residents of Israel, and intergroup contact and cooperation amongst co-workers and between patients is a common norm.¹³ Given the prominence of Clinic X,¹⁴ and given the intense intergroup animosity between Jews and Arabs in Israel, this site serves as a crucial case for my theory of descriptive representation in public goods provision and prejudice reduction. More so, the assignment mechanism of patients to doctors discussed below, allows for a credible identification of one channel through which representative institutions may shape intergroup attitudes - intergroup contact.

The large share of Arab doctors in Clinic X may raise concerns that only Jews who are tolerant towards Arabs will approach Clinic X. Such *selection into clinic*, would pose a threat to external validity, rather than causal identification and internal validity. However, qualitative evidence regarding the prevalence of Arab service providers in Israeli healthcare institutions (Rosner 2016; Keshet and Popper-Giveon 2018), and the fact that in many localities Clinic X is a central service provider, should reduce concerns regarding selection into clinic. This is since Jews are likely to encounter Arabs in most hospitals which they attend.

Nonetheless, to further address this concern, In Section F of the supplementary materials I compare the demographics of hospital patients with a representative sample of Jewish respondents

¹²An elaborate depiction of conditions treated by doctors in Clinic X is portrayed in Figure A1 of my supplementary materials.

¹³To a great extent on some metrics Clinic X over-represents minorities. Qualitative interviews suggest that this is driven by the fact that many Jewish medical-school graduates prefer to obtain positions in larger hospitals with more lucrative residencies.

¹⁴In many of the cities in which clinics operate, they are the central provider of simple emergency medical care.

to the Israeli Democracy Index public opinion survey. This survey was implemented by the Israeli Democracy Institute (IDI) during the same year as my study (Herman et al. 2018). Basic difference in means tests reported in Section F of the supplementary materials suggest that both samples are relatively similar on key demographics (age, gender, education, religiosity, and ideology), reducing concerns regarding selection into clinic and external validity.

4.4 Identification Strategy

A key feature of the medical clinics I study, is that they provide services in many locations, around the clock, with minimal waiting time, on a first-come, first-served basis. Patients do not select their doctors or nurses, and requests for specific doctors are not met.¹⁵ Thus assignment of patients to doctors depends on multiple factors, including: The exact time in which a patient arrives in a clinic, the length of the line in the clinic, the number of doctors in the clinic, and their daily work load.¹⁶ Therefore, after checking into a local clinic, patients wait for the next free doctor in a communal reception room (See figure 1a). Once a doctor is available, patients receive treatment in private rooms (see Figure 1b).¹⁷ Wait times for treatment vary by clinic and day, and can range between several minutes to an hour in extreme cases of patient over load. Actual treatment from doctors usually lasts between 3-10 minutes.

Given the multiple determinants of assignment to doctors within clinics, I assume that it is impossible for patients to select a Jewish or Arab doctor. Therefore, the matching of doctors and patients, and in our case the assignment of patients to intergroup contact, or lack thereof, is administered randomly, or at least in a manner which is orthogonal to an individual's political and social attitudes. Additionally, the clinic's firm policy to decline patients' requests for specific

¹⁵Unfortunately, I do not have data regarding the prevalence of such requests. Nonetheless, in-depth interviews with clinic administrators suggest that that such requests are rare and not accommodated.

¹⁶It should be noted that apart from emergency medical services, Clinic X also provides routine check-ups and medical tests such as MRIs, in which individuals select doctors. In order to avoid selection bias, I focus on patients receiving emergency medical care. Emergency medical care in the Clinic X is any condition which does not require hospitalization or a pre-arranged appointment.

¹⁷Prior to doctor assignment, patients often receive initial treatment from a nurse, and at times go through additional required examination (e.g. x-rays).



(a) Reception Room



(b) Treatment Room

Figure 1: Illustration of Clinic X Facilities

doctors, limits concerns regarding treatment compliers and non-compliers. Thus, the assignment process described above serves as the base of my empirical inquiry.

To address any concerns regarding my assignment mechanism, I empirically test the assumptions detailed above, ensuring that patients' religiosity, education, gender, and age do not predict assignment to specific doctors. To do so, I regressed a binary variable `Arab doctor` over patients' characteristics. To account for unobserved clinic and date variables, in my main specification I employ day and clinic fixed-effects and cluster errors at the clinic level as specified in my pre-analysis plan.¹⁸ As evident from table 1, individual level characteristics such as religiosity, education, and age do not predict assignment to an Arab doctor. However, my pre-registered fixed-effects specification suggests that men are marginally less likely to receive treatment from an Arab doctor. Therefore, I replicate my main results controlling for gender (as well as other pre-treatment covariates). Doing so, strengthens my main results.

Lastly, for treatment to be effective, patients must be aware of their doctor's ethnicity. I am assured that this is indeed the case for three main reasons. First, all doctors wear name badges and introduce themselves to patients by their last name (e.g. Dr. Nashashibi). Thus patients can use names as a strong cue for ethnic identifiability especially in the Israeli case where last names are a strong signal of ethnicity (Shayo and Zussman 2011). Second, interviews with clinic administrators corroborate this expectation, suggesting that both patients and doctors are aware of

¹⁸In addition, I consider an alternative logit specification.

Table 1: Demographic Correlation with Assignment to Arab Doctor

	OLS	OLS	OLS	OLS	OLS	Logit
Religiosity	-0.01 (0.01)				-0.01 (0.01)	-0.10 (0.06)
Gender		-0.02 (0.01)			-0.03* (0.01)	-0.21 (0.11)
Education			-0.01 (0.01)		-0.01 (0.01)	-0.04 (0.05)
Age				-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Num. obs.	2058	2164	1914	2164	1878	1878
AIC						1925.64
BIC						1953.33
Log Likelihood						-957.82
Deviance						1915.64

* $p < 0.05$. OLS model includes clinic and date fixed effects, and errors are clustered by clinic.

the ethnicity of the people with whom they engage. Lastly, responses to open ended questions in the clinic’s treatment evaluation survey (described below) provide anecdotal support for the (positive and negative) salience of ethnicity in the interaction between patients and doctors. I provide two examples below, as well as a full account of comments regarding doctors’ ethnicities in section D of the supplementary materials.

- “...the doctor was an absolute angel!!! I am an ultra-orthodox [woman] and I won’t lie. [I] don’t like Arabs. But this doctor just touched me and I wish she gets promoted in some way. She is simply a real angel. I don’t know her name. She is a young short Arab doctor. Truly amazing.”

Jewish female patient. November 2018.

- “...it is preferable to have at least one or two Jewish doctors. It is horrifying to be alone with six Arabs and not one Jew.”

Jewish female patient. December 2018.

4.5 Survey Instrument

Clinic X collects demographic and service related information from patients, as well as satisfaction survey data following treatment. In order to identify the effects of contact on intergroup attitudes, Clinic X’s evaluation team embedded within their routine surveys several questions relating to intergroup attitudes and political preferences.¹⁹ Usually, surveys are sent out several hours after treatment. However as part of our study, we randomly assigned respondents to receive surveys either one or ten days post treatment, in order to test the immediate and longer-term effects of contact. The main items relating to intergroup attitudes which I use as outcomes in my analyses are presented in Table 2.

Item	Question	Answers
<i>Social Distance</i>	What is the maximum level of proximity which you would accept with (Arabs / Jews / Foreign Workers / Tourists)?	<ol style="list-style-type: none"> 1. Not accept in my country 2. Accept as guest in my country 3. Accept as citizen in my country 4. Accept as co-worker 5. Accept as neighbor 6. Accept as close friend 7. Accept as family through marriage
<i>Peace</i>	Do you agree with the following statement? “Most Arabs want to live in peace”	<ol style="list-style-type: none"> 1. Disagree 2. Tend to disagree 3. Tend to agree 4. Agree
<i>Feeling Thermometer</i>	Please place the following social groups on a feeling thermometer (Arabs / Jews / Foreign Workers / Tourists)	<ol style="list-style-type: none"> 1. 0 - 20 2. 20 - 40 3. 40 - 60 4. 60 - 80 5. 80 - 100
<i>Trust</i>	Do you agree with the following statement? “It is possible to trust most Arabs in Israel”	<ol style="list-style-type: none"> 1. Disagree 2. Tend to disagree 3. Tend to agree 4. Agree
<i>Composite Index</i>	Average z-scores of all outcomes	

Table 2: Main Outcomes

¹⁹The survey instrument embedded within Clinic X’s ongoing routine evaluation program is available on EGAP, together with my original pre-analysis plan linking specific hypotheses to unique survey items <https://egap.org/registration/5216>. In addition, an elaborate description of all survey items can be found in section A of my supplementary materials.

By embedding additional questions within Clinic X’s SMS-based routine survey, I am able to test whether Jewish patients who experienced contact with Arab doctors differ in their intergroup attitudes from their counterparts who did not experience intergroup contact during their medical treatment. A major benefit of this research design, is that when experiencing contact, both subjects and doctors are unaware of my study. This reduces concerns related to demand effects at the time of treatment, and provides me with a rare opportunity to study a truly naturalistic form of contact (Dixon, Durrheim and Tredoux 2005). In addition, the naturalistic form of contact that I identify, and the relative ease of approaching patients via SMS surveys, allow for a unique test of the contact hypothesis across 21 experimental sites. A general schema of my research design and data collection procedure is depicted in Figure 2.

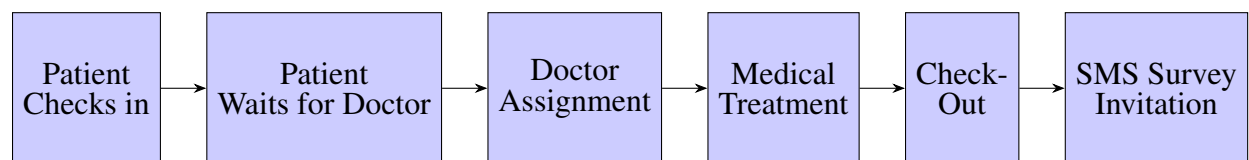


Figure 2: Research Design

Before turning to describe my estimation strategy, I report basic descriptive statistics of respondents in Table 3. The average age of survey respondents in my sample is slightly above 32,²⁰ and the gender distribution is slightly leaning more female. Additionally, my sample is relatively left-leaning, but rather traditional. In Section F of the supplementary materials I elaborate on the similarity of my sample to a representative sample of survey respondents in Israel.

4.6 Estimation Strategy

My estimation strategy aims to recover the causal effects of contact (i.e. doctor assignment) on intergroup attitudes (i.e. survey responses). Therefore, I consider the random assignment of treat-

²⁰My sample of include 783 patients under the age of 18. For such patients, SMS-based surveys are sent to a legal guardian who accompanied the patient to the clinic. The legal guardian is most often a parent. In Section C.5 of my supplementary materials I demonstrate the robustness of my results when excluding these respondents.

Table 3: Descriptive Statistics - Jewish Patients

Statistic	N	Mean	St. Dev.	Min	Max
Age	2164	32.465	24.690	0	96
Male	2164	0.463	0.499	0	1
Religiosity	2058	0.942	0.973	0	3
Right Left	1675	2.508	1.076	1	5
Education	1914	2.503	1.220	0	4

ment within 21 clinics over a period of three months,²¹ by incorporating clinic, and day fixed effects. I cluster standard errors by clinic.²² Equation 1 depicts my main model, a linear regression, where Y_{icd} denotes a survey response (as depicted in Table 2) of patient i , receiving treatment in clinic c , on day d . β represents the coefficient of my main treatment, and η_c and Ψ_d represent clinic and day fixed effects successively. ε_{icd} represents my model’s error term.

$$Y_{icd} = \beta X_{treatment} + \eta_c + \Psi_d + \varepsilon_{icd} \quad (1)$$

I employ this model to study the effects of contact between Jewish patients and Arab doctors on four distinct survey measures: (1) A commonly used social distance scale (Bogardus 1933), (2) an item relating to intergroup peace (Samooha 2013), (3) an intergroup feeling thermometer, and (4) a measure of intergroup trust (Samooha 2013). The wording of these measures is depicted in Table 2, and Hebrew translations can be found in my pre-registered materials.

Though these four items measure distinct concepts, I combine them into a general index of intergroup attitudes by averaging their z-scores.²³ This intergroup attitude index serves as my fifth

²¹Data were collected between October 30, 2018 and January 30, 2019.

²²In my pre-analysis plan, I mistakenly pre-registered a similar specification including doctor fixed effects. Since doctor fixed effects correlate perfectly with my treatment indicator (i.e. Arab doctor), including doctor fixed effects in my specification is impossible. This raises concerns that my effects are driven by doctor-specific qualities which are unrelated to contact. Random assignment should cause these qualities to be balanced. Nonetheless, I address these concerns through an extensive set of robustness checks detailed in section 6.

²³This index has a 0.8 Corenbach’s Alpha.

outcome.²⁴ I expect my treatment to affect all five outcomes in a similar fashion enhancing positive intergroup attitudes ($\beta > 0$). My survey also includes additional questions regarding intra-group attitudes, and attitudes towards tourists and foreign workers, which should not be affected by my treatment. I report descriptive statistics of all variables used in my analyses in Table A1, and Figures A3-A4 of my supplementary materials.

5 Analyses

5.1 Main Results

Results from my main analyses are depicted in Figure 3, as well as in Table A4 of the supplementary materials. Figure 3 plots the main coefficient of interest from five different models, representing the effect of contact with an Arab doctor on four distinct outcomes relating to intergroup attitudes, as well as a composite index of all four items. Figure 3 suggests that receiving treatment from an Arab doctor has a positive effect on all outcomes.²⁵ This effect meets conventional levels of statistical significance for outcomes relating to social distance and attitudes about the prospect of peace ($p < 0.05$). In addition, the effects of contact on responses to feeling thermometers approach conventional levels of statistical significance ($p < 0.08$), but effects on intergroup trust are insignificant. When considering a composite index of all four outcomes, it appears that contact has a general positive effect on intergroup attitudes – suggesting that brief engagement between Jewish patients and Arab doctors effectively reduces intergroup prejudice.

Substantively speaking, contact accounts for almost a fifth of a standard deviation shift in responses to a conventional seven item social distance scale (ranging from 0-6). As the average Jewish respondent in my data is willing to accept an Arab as a co-worker ($\mu = 3.20$), this effect moves respondents towards willing to accept Arabs as neighbors ($\mu = 3.52$). More so, the magnitude of the effects of contact on social distance, is equivalent to the effects of a half-unit shift on conventional religiosity scales, moving from traditionalist to religious Jewish practices. This com-

²⁴Note that I did not register this index in my pre-analysis plan.

²⁵As depicted in Table 2, higher values in all outcomes represent more favorable intergroup relations, and therefore positive effects in my models should be interpreted as evidence for prejudice reduction.

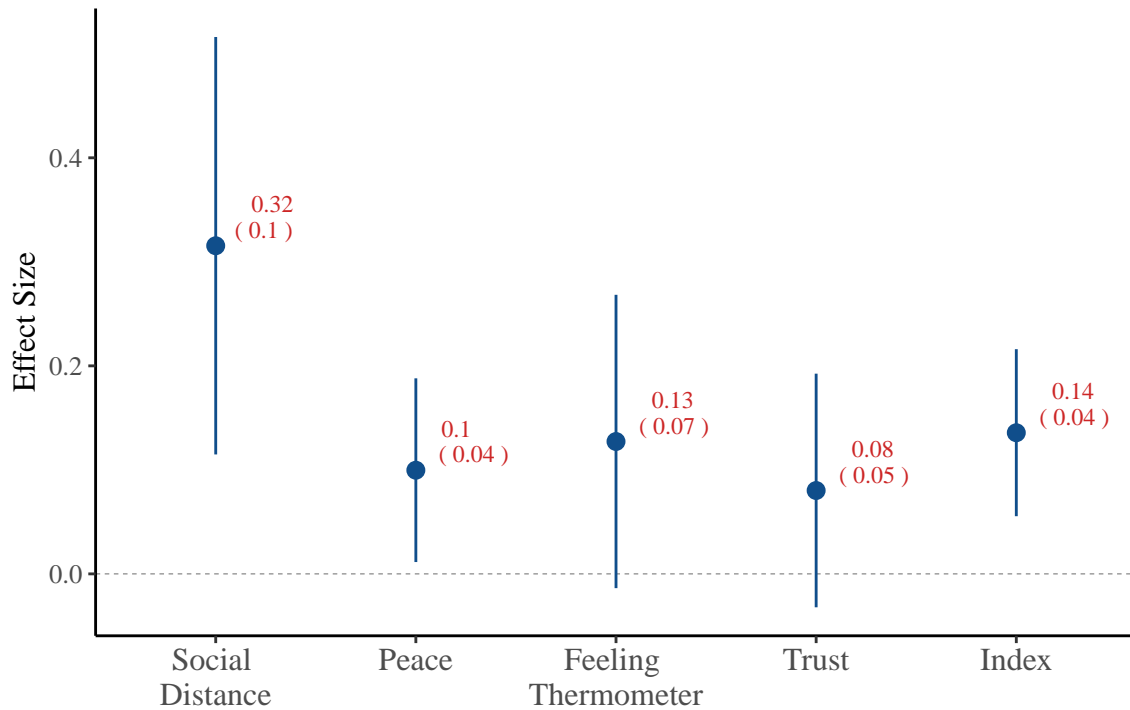


Figure 3: Effects of Contact with Arab Doctors on Jewish Attitudes – Each reported coefficient and its corresponding 95% confidence interval is extracted from a separate model identifying the average treatment effect of contact with an Arab doctor on Jewish prejudice towards Arabs. All models include day and clinic fixed effects, and errors are clustered by clinic. Sample size for each model is: social distance ($n = 1285$), Peace ($n = 1456$), Thermometer ($n = 1521$), Trust ($n = 1461$), Index ($n = 1234$).

parison is remarkable given previously documented exclusionary attitudes towards Arabs amongst ultra-orthodox Jews, and given the strong links between social distance and costly non-cooperative behavior (Enos and Gidron 2018).

In addition, after engaging with an Arab doctor, Jewish patients are more optimistic about the extent to which most Arabs want to live in peace. Experiencing contact accounts for slightly more than a tenth of a standard deviation increase in responses to a standard 4 item question. This effect is slightly larger than a half-unit change in religiosity, as measured in the data.

The effect of contact on intergroup affect does not meet conventional levels of statistical significance. However, since this effect is marginally significant ($p < 0.08$), one can carefully interpret it as further suggestive evidence that contact improves intergroup attitudes as measured by commonly used feeling thermometers.²⁶ That said, unlike other forms of contact which have been shown to foster intergroup trust (Mousa 2019), brief interactions with an out-group doctor do not seem to build trust between Jews and Arabs in this context.

5.2 Do the Effects of Contact Decay Over Time?

To explore whether the effects of contact decay over time, patients were randomized to receive invitations for participation in treatment evaluation surveys either one or ten-days following their visit in a local clinic. My main results pool over both conditions. However, Figure A5 in Section C of my supplementary materials reports the effects of contact for respondents receiving surveys in these two different times. My analyses suggest that there is little systematic difference between respondents assigned to the immediate and longer-term timing conditions.²⁷

My inability to detect a significant difference across timing conditions is likely driven by the power decrease resulting from splitting my sample into two additional groups (i.e. one- and ten-day survey recipients). In addition, though assignment to timing conditions was randomized, individuals opting into the survey when approached immediately after treatment, may differ from indi-

²⁶Given Clinic X's survey interface, rather than presenting respondents with a 0-100 slider, feeling thermometers ranged from 1-5.

²⁷In addition, a more granular measure of timing – the exact number of hours between treatment and survey response – also does not moderate the effects of contact.

viduals opting into the survey ten days following treatment.²⁸ To partially account for potential demographic differences across the one and ten day respondents, I deviate from my pre-registered analysis plan, and replicate the models presented in figure A5 controlling for standard observed demographics.²⁹ Results presented in figure A6 of my supplementary materials still fail to detect a consistent difference between survey responses in the immediate and longer time-period following treatment.

I carefully construe these findings to suggest that the effects of contact are relatively similar when examined one and ten days post-treatment. Nonetheless, my ability to speak to the long-term effects of contact in this context is rather limited. In that sense, my findings are similar to previous work which has considered the relatively immediate effects of intergroup contact (Enos 2014; Karim N.d.).

5.3 Does Attitudinal Change Come at the cost of Service Satisfaction?

The findings above suggest that the presence of minority service providers within a given institution can reduce prejudice through a channel of intergroup contact. Nonetheless, one may wonder whether such attitudinal change comes at the cost of majority group satisfaction from service. Generally, an underlying notion of the representative bureaucracy literature is that receiving service from an in-group is oftentimes more beneficial, as it enables more direct channels of communication and shared experiences. Nonetheless, recent advances have shown that embedding under-represented groups within a bureaucracy may not change an organization's performance (Bhavnani and Lee N.d.), or even provide the general public with net-benefits (Nanes 2018).

To further engage with this question, I run similar models replacing my original outcomes with two service satisfaction indicators – general satisfaction from service, and particular satisfaction from doctors. In a similar fashion to my main analyses, I further consider a composite index of both survey items. Results presented in Figure 4 suggest that receiving treatment from an Arab doctor does indeed reduce general satisfaction from the service received in the clinic by about a

²⁸Indeed, it seems that the ten day condition includes more religious subjects.

²⁹gender, religiosity, age, and education.

tenth of a standard deviation. Nonetheless, similar effects do not seem to register with regards to satisfaction from the specific service provided by doctors. More so, when measuring the effects of contact on the general composite index, it seems that satisfaction is not sensitive to doctors' ethnic identity. Indeed, the effects on doctor satisfaction and the composite index are both statistically insignificant and substantively small.

I cautiously interpret these models to suggest that prejudice reduction through descriptive representation may come at the cost of decreases in general satisfaction from the providing institution. That said, patients do not report decreases in satisfaction from the specific service provider with whom they engage. To a great extent, these results echo previous findings by [Rao \(2019\)](#) who shows the diversity can reduce prejudice towards disadvantaged minorities at limited costs for majority group members.

5.4 Heterogenous Treatment Effects

In Section [C.3](#) of my supplementary materials I report results from a battery of pre-registered models testing for heterogenous treatment effects. I consider heterogeneity along a host of variables including: Gender, Age, Doctor quality,³⁰ Previous exposure to Arabs, and experience of contact during a period of escalated violence in the Gaza strip.³¹ I do not consider heterogeneity in effects conditional on ideology to avoid post-treatment bias. Across all models, I do not find evidence for heterogenous treatment effects.

Interpreting these null effects across all additional models can take one of two different approaches. A skeptical interpretation would suggest that given sample size and conservative empirical specifications, the models I employ are underpowered to detect heterogenous treatment effects. Alternatively, a more charitable reading of these tests suggests that contact has a general effect across different contexts and subsets of the Israeli population. Given the prevalence of segrega-

³⁰As measured by aggregate doctor ratings collected before the implementation of the study.

³¹To measure this covariate, I leverage the outburst of confrontations between Hamas and the Israeli military between November 11-13 (2018). A further description can be found in my pre-analysis plan. A similar approach has been implemented by [Hjort \(2014\)](#) when examining the way a cycle of violence moderated the effects of ethnic diversity on firm productivity in Kenya.

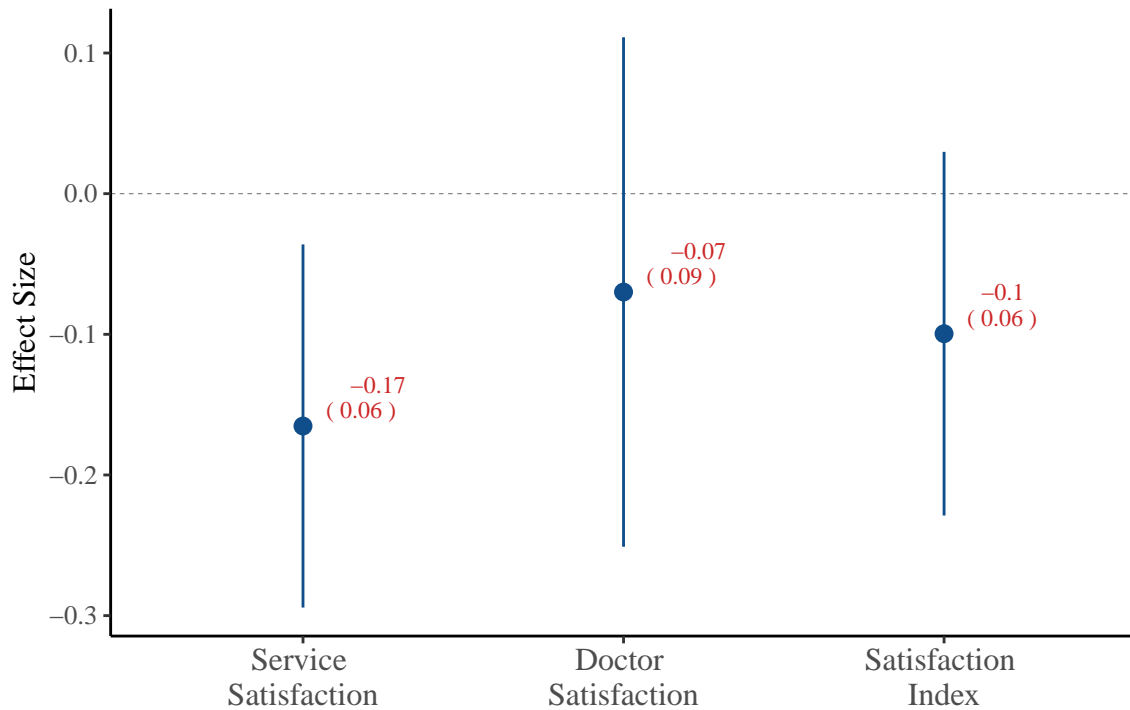


Figure 4: Effects of Contact with Arab Doctors on Satisfaction from Service – Each reported coefficient and its corresponding 95% confidence interval is extracted from a separate model identifying the average treatment effect of contact with an Arab doctor on Jewish patients’ service satisfaction. All models include day and clinic fixed effects, and errors are clustered by clinic. Sample size for each model is: service satisfaction ($n = 2080$), Doctor Satisfaction ($n = 2091$), Index ($n = 2066$).

tion and limited quotidian interactions between most Jews and Arabs in Israel (Rokem, Weiss and Miodownik 2018), it is possible that such intimate intergroup contact is a novel experience for most Jews, leading to rather uniform effects amongst diverse respondents.

6 Robustness Checks

6.1 Doctor Quality

A skeptical interpretation of the findings presented above may raise concerns that doctor specific qualities which correlate with Arab identity may drive my main results. This interpretation would suggest that unobserved doctor characteristics and not intergroup contact per-se account for my main effects. To relax such concerns I implement several robustness checks.³²

First, I test the effects of contact on a host of outcomes which should not shift in light of intergroup contact with an Arab doctor. These include social distance and general affect towards Jews, American tourists, and foreign workers, as well as a question regarding Jewish willingness to live in peace. As depicted in Figure A12 in the supplementary materials, unlike attitudes towards Arabs, other outcomes do not seem to be affected by my treatment.³³ This provides initial support for my main pre-registered models which depict theoretically expected effects.

I take two additional measures to account for doctor attributes and ensure that they are not driving my main effects. Thus in Section C.4 of my supplementary materials I run additional models controlling for doctor quality, as well as mixed-effects models incorporating doctor random effects. Overall, results from these additional models remain substantively similar, though in some cases the effect of contact on the peace related outcome does not meet conventional levels of statistical significance ($p < 0.05$) and p values are only marginally significant. A further discussion of these models can be found in Section C.4 of my supplementary materials.

³²These additional analyses were not registered in my pre-analysis plan.

³³With the exception of one outcome – feelings towards American tourists.

6.2 Unbalanced Covariates

As mentioned in Section 4.4, my identification strategy relies on the assumption that patients do not select their doctors, and coordinating assignment to a specific doctor is highly unlikely. In Table 1 I test this assumption with several pre-treatment covariates showing that across different specifications most demographics do not predict assignment to an out-group doctor. Nonetheless, in my pre-registered OLS specification, Jewish men seem to be less likely to receive treatment from an Arab doctor. To address this imbalance, Table A4 in my supplementary materials includes models controlling for gender, as well as other pre-treatment covariates. Doing so does not change my main results.³⁴

6.3 Survey Responses for Minor Patients

Since Clinic X provides medical service to a host of patients including minors, 763 responses to my survey were linked to visit IDs of minor patients under the age of 18. When minor's receive treatment in Clinic X, they are accompanied by a legal guardian. This legal guardian is usually a parent, which receives the invitation to participate in the treatment evaluation survey. Though legal guardians come in contact with doctors, they are not medically treated and therefore they might differ from other respondents in their experience of contact. To address this matter, I run additional analyses focusing on responses received from adult patients in Section C.5 of the supplementary materials.³⁵ Despite reductions in sample size, results from these models remain substantively similar.³⁶

6.4 Selection into Survey and Non-responses

A major advantage of my research design is the fact that it allows for identification of contact absent any obtrusive intervention. In fact, when experiencing contact neither doctors nor patients are aware of my study. More so, the ability to embed questions within an ongoing evaluation survey allows me to approach a unique population of Israeli citizens from a range of different

³⁴Models incorporating covariates were not specified in my pre-analysis plan.

³⁵These models were not pre-registered.

³⁶However, results of the peace indicator only reach marginal levels of statistical significance.

backgrounds. This is especially important in light of the dearth of evidence regarding the ability of contact to reduce ethnic and racial prejudice amongst adults (Paluck, Green and Green 2019).

While all patients are invited to participate in evaluation surveys, only a minority of them opt-in to report outcomes of interest.³⁷ This could introduce serious problems especially if the probability of opting into the survey correlates with doctor assignment. However, Clinic X's administrative records allow me to model selection into survey, and somewhat address this concern. Indeed, across a host of different specifications further described in Section B of my supplementary materials, receiving treatment from an Arab doctor does not increase the probability of opting into the treatment evaluation survey. That said, gender and age both predict selection into the survey, as men are less likely, while older patients are more likely to engage with the survey.³⁸

This analysis has one limitation: Since my ethnicity variable is self-reported in the evaluation survey, I cannot differentiate between Jewish and Arab patients in the administrative records. Therefore my analyses cannot rule out the possibility that treatment by an out-group doctor (rather than Arab doctor) predicts selection into survey. To further address this limitation and reduce concerns regarding selection out of survey by gender and age, in Section F of the supplementary materials, I demonstrate that survey respondents in my sample are similar in age, religiosity, education, and ideology, to a representative sample of respondents to the 2018 Israeli Democracy Index public opinion survey.

Another concern relates to missingness in outcome measures for participants who opted into the survey. Since respondents were not monetarily incentivized, Clinic X administrators did not want to force responses in their evaluation survey.³⁹ Selective non-response would be especially acute if it correlated with treatment. Put differently, if prejudice individuals treated by Arab doctors were to skip items relating to prejudice, the effects identified throughout the paper may be over-exaggerated. I relax these concerns in Section B through a series of models which show that

³⁷ Approximately 1% of all patients fully participate in the evaluation survey.

³⁸ These models were not included in my original pre-analysis plan.

³⁹ More so, forcing responses may potentially lead to lower-quality responses and more severe attrition.

treatment assignment does not predict non-responses to my main outcome measures.

6.5 Effects on Arab Patients

Lastly, one may wonder whether similar effects are registered amongst Arabs who receive treatment from a Jewish doctor. Unfortunately, a full examination of this question is beyond the scope of my paper for both theoretical and methodological reasons. The main hurdle to identifying the effects of contact amongst Arab patients relates to the rather limited numbers of Arabs attending clinics, in relation to the rather large share of Arab doctors providing service on a daily basis. Beyond this limitation, only 166 Arab patients opted into the treatment evaluation survey in the period under investigation. It is very reasonable to suspect that these Arab respondents are unrepresentative of a broader population in Israel.⁴⁰

With these caveats in mind, in Figure 5 I present results of similar models identifying the effects of intergroup contact between Arab patients and Jewish doctors. Like Jews receiving treatment from an Arab doctor, Arabs experiencing contact with a Jewish doctor report increased willingness for proximity with Jews as measured by the social distance scale. In addition, they also report warmer feelings towards Jews. Nonetheless, the effects of contact on intergroup trust and attitudes about peace are statistically insignificant. Lastly, when considering a composite measure, the effects of contact seem to be insignificant.⁴¹ These results suggest that contact with out-groups embedded within an institution providing public goods *may* have positive effects not only for minority, but also for majority group members. However, future research leveraging a more appropriate setting is required to further establish this claim.

7 Conclusion

In this paper I consider a crucial, yet under-investigated question: How does descriptive representation in public goods provision shape intergroup relations? In doing so, I develop a theory

⁴⁰Though surveys were translated into Arabic, Clinic X ended up distributing surveys only in Hebrew.

⁴¹However, this index has a 0.5 Corenbach's Alpha, and therefore effects on the composite which was not pre-registered should be taken with a grain of salt.

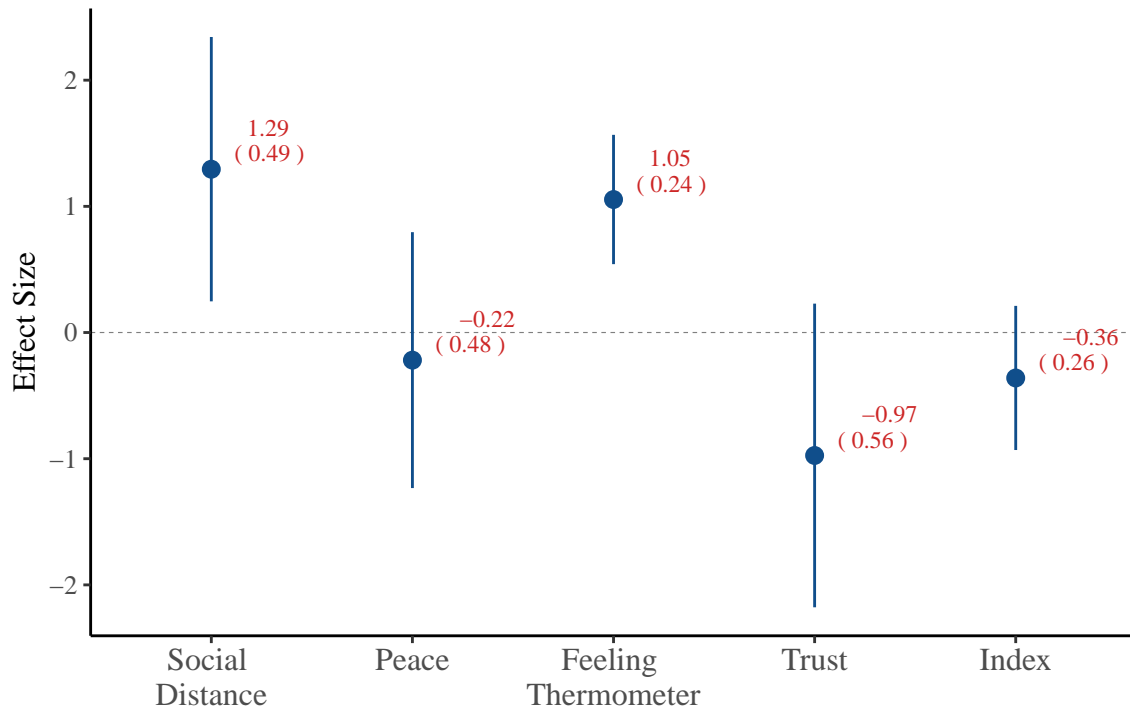


Figure 5: Effects of Contact with Jewish Doctors on Arab Attitudes – Each reported coefficient and its corresponding 95% confidence interval is extracted from a separate model identifying the average treatment effect of contact with a Jewish doctor on Arab prejudice towards Jews. All models include day and clinic fixed effects, and errors are clustered by clinic. Sample size for each model is: social distance ($n = 76$), Peace ($n = 92$), Thermometer ($n = 92$), Trust ($n = 93$), Index ($n = 74$).

regarding one particular mechanism through which representative institutions may improve intergroup relations – positive intergroup contact. I develop my theory by synthesizing insights from the representative bureaucracy literature, together with theory and evidence regarding the contact hypothesis. Doing so, I propose that institutions engaged in public goods provision, which employ minorities, can serve to reduce prejudice on a large scale, by promoting positive equal status intergroup contact which is institutionally supported, directed towards a common goal, and capable of facilitating conceptions of common humanity.

Introducing a natural experiment across 21 Israeli medical clinics, I demonstrate that institutions that promote positive contact can improve intergroup relations on a large scale. More specifically, leveraging the random assignment of patients to doctors, I demonstrate that Jewish patients treated by Arab doctors report reduced levels of social distance and increased optimistic attitudes about intergroup peace. This evidence suggests that beyond the positive benefits of diversity for minority health outcomes (Alsan, Garrick and Graziani 2018), minority doctors may also cure intergroup prejudice.

In developing and testing a theory of prejudice reduction through descriptive representation I contribute to two distinct literatures. First, I introduce a consequential question for scholars of representative bureaucracy, namely: Does representation have externalities which do not relate to service provision? I demonstrate that representation can have positive externalities when it facilitates positive intergroup contact. However, it is important to acknowledge, that representation in the process of public goods provision may have mixed effects. Thus not all contact with civil servants and service providers is positive. More so, institutional diversity may engender threat perceptions and ignite concerns regarding affirmative action and labor market competition. Therefore, future research should directly identify the conditions under which representation promotes positive intergroup relations.

I also make four contributions to the prejudice reduction and intergroup contact literature. First, by leveraging a natural experiment, I overcome hurdles of selection bias, and provide strong causal evidence for the positive effects of contact, amongst adults embedded in a deeply divided society,

in a naturalistic setting which does not require obtrusive intervention. Second, improving on many experimental interventions which identify the effects of contact in a sterile lab setting (Forbes 1997; Dixon, Durrheim and Tredoux 2005), my natural experiment demonstrates that brief, quotidian, and naturally occurring contact can have substantial effects on deeply ingrained intergroup attitudes. It follows that intergroup contact need not be administered as part of a reconciliation or prejudice reduction intervention in order to be effective.

Third, recent evidence suggests that brief contact and exposure to out-groups may be detrimental to intergroup relations as it does not allow for meaningful interactions (Condra and Linardi 2019; Hangartner et al. 2019). Nonetheless, I demonstrate that under certain conditions even very brief interactions can reduce prejudice and improve intergroup relations. Whether these effects are durable in the long term, remains a question for future research. Nonetheless, by embedding minorities across different institutions, societies may facilitate multiple instances where intergroup contact occurs, leading to sustainable social change.

Fourth, by linking the contact hypothesis with an institutional framework of descriptive representation, I take the contact literature a step forward. Thus, my evidence suggests that scholars of prejudice reduction should further consider how institutions that facilitate innumerable and repeated instances of intergroup contact may serve to amplify and spread the effects of contact, and promote prejudice reduction in a scalable and sustainable way. Unlike people-to-people reconciliation interventions which often focus on one village, or urban neighborhood, my exploration of intergroup contact in public goods provision examines thousands of interactions occurring across 21 medical clinics, dispersed across Israel. The broader geographical scope of this study, stems from the far reaching arms of institutions. Therefore, further considering the link between intergroup relations and institutions is a promising path for scholars of prejudice reduction.

Lastly, my findings inform recent heated debates regarding diversity in Israeli healthcare provision. Preferences for segregation in Israeli medical facilities have been widely documented in popular media as well as in academic research (Keshet and Popper-Giveon 2018; Halevy 2018). These preferences are in line with more general exclusionary attitudes towards Arabs and Pales-

tinians documented amongst many Israelis (Enos and Gidron 2018). Nonetheless, somewhat in contrary to these preferences my findings suggest that diversity in healthcare provision which promotes positive intergroup contact can serve to build bridges between Jews and Arabs. This is in line with recent qualitative policy reports regarding intergroup cooperation between Jews and Arabs in Israeli hospitals (Rosner 2016). Clearly, contact between patients and doctors can serve to cure prejudice and improve intergroup relations. Therefore, future research should develop and test theoretically informed strategies to encourage selection into contact with out-group doctors, even in contexts where preferences for segregation are common and ingrained.

References

- Adida, Claire L, Adeline Lo and Melina R Platas. 2018. "Perspective taking can promote short-term inclusionary behavior toward Syrian refugees." *Proceedings of the National Academy of Sciences* 115(38):9521–9526.
- Allport, Gordon W. 1954. *The nature of prejudice*. Unabridged, 25th anniversary ed ed. Reading, Mass.: Addison-Wesley Pub. Co.
- Arababah, Ala, William Marble, Salma Mousa and Alexandra Siegel. 2019. "Can Exposure to Celebrities Reduce Prejudice? the Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes."
- Alsan, Marcella, Owen Garrick and Grant C Graziani. 2018. Does diversity matter for health? Experimental evidence from Oakland. Technical report National Bureau of Economic Research.
- Bar, Revital and Asaf Zussman. 2017. "Identity and Bias: Insights from Driving Tests."
- Bar-Tal, Daniel. 1998. "Societal beliefs in times of intractable conflict: The Israeli case." *International Journal of Conflict Management* 9(1):22–50.
- Barak-Corren, Netta, Yuval Feldman and Noam Gidron. 2018. "The Provocative Effect of Law: Majority Nationalism and Minority Discrimination." *Journal of Empirical Legal Studies* 15(4):951–986.
- Barnhardt, Sharon. 2009. "Near and Dear? Evaluating the Impact of Neighbor Diversity on Inter-Religious Attitudes." *Unpublished working paper* .
- Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg and Maisy Wong. 2019. Unity in Diversity? How Intergroup Contact Can Foster Nation Building. Technical report National Bureau of Economic Research.

- Beaman, Lori, Raghavendra Chattopadhyay, Esther Duflo, Rohini Pande and Petia Topalova. 2009. "Powerful women: does exposure reduce bias?" *The Quarterly journal of economics* 124(4):1497–1540.
- Bhalotra, Sonia, Irma Clots-Figueras and Lakshmi Iyer. 2017. "Pathbreakers? Women's Electoral Success and Future Political Participation." *The Economic Journal* 128(613):1844–1878.
- Bhavnani, Ravi, Karsten Donnay, Dan Miodownik, Maayan Mor and Dirk Helbing. 2014. "Group Segregation and Urban Violence." *American Journal of Political Science* 58(1):226–245.
- Bhavnani, Rikhil R. 2009. "Do electoral quotas work after they are withdrawn? Evidence from a natural experiment in India." *American Political Science Review* 103(1):23–35.
- Bhavnani, Rikhil R and Alexander Lee. 2018. "Local embeddedness and bureaucratic performance: evidence from India." *The Journal of Politics* 80(1):71–87.
- Bhavnani, Rikhil R and Alexander Lee. N.d. "Does Affirmative Action Worsen Bureaucratic Performance? Evidence from the India Administrative Service." *American Journal of Political Science*. Forthcoming.
- Bogardus, Emory Stephen. 1933. "A social distance scale." *Sociology & Social Research* .
- Broockman, David and Joshua Kalla. 2016. "Durably reducing transphobia: A field experiment on door-to-door canvassing." *Science* 352(6282):220–224.
- Canetti-Nisim, Daphna, Gal Ariely and Eran Halperin. 2008. "Life, pocketbook, or culture: The role of perceived security threats in promoting exclusionist political attitudes toward minorities in Israel." *Political Research Quarterly* 61(1):90–103.
- Chattopadhyay, Raghavendra and Esther Duflo. 2004. "Women as policy makers: Evidence from a randomized policy experiment in India." *Econometrica* 72(5):1409–1443.
- Chauchard, Simon. 2014. "Can descriptive representation change beliefs about a stigmatized group? Evidence from rural India." *American political Science review* 108(2):403–422.

- Clots-Figueras, Irma. 2011. "Women in politics: Evidence from the Indian States." *Journal of public Economics* 95(7-8):664–690.
- Condra, Luke N and Sera Linardi. 2019. "Casual Contact and Ethnic Bias: Experimental Evidence from Afghanistan." *The Journal of Politics* 81(3):000–000.
- Deutsch, Morton and Mary Evans Collins. 1951. *Interracial housing: A psychological evaluation of a social experiment*. U of Minnesota Press.
- Dinas, Elias, Konstantinos Matakos, Dimitrios Xeferis and Dominik Hangartner. 2016. Waking Up the Golden Dawn: Does Exposure to the Refugee Crisis Increase Support for Extreme-Right Parties? Technical report Working paper. <https://sites.google.com/site/kostasmatakos/research>.
- Ditlmann, Ruth K and Cyrus Samii. 2016. "Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel." *Peace and Conflict: Journal of Peace Psychology* 22(4):380.
- Ditlmann, Ruth K, Cyrus Samii and Thomas Zeitzoff. 2017. "Addressing Violent Intergroup Conflict from the Bottom Up?" *Social Issues and Policy Review* 11(1):38–77.
- Dixon, John, Kevin Durrheim and Colin Tredoux. 2005. "Beyond the optimal contact strategy: A reality check for the contact hypothesis." *American psychologist* 60(7):697.
- Enos, Ryan D. 2014. "Causal effect of intergroup contact on exclusionary attitudes." *Proceedings of the National Academy of Sciences* 111(10):3699–3704.
- Enos, Ryan D. 2017. *The space between us: Social geography and politics*. Cambridge University Press.
- Enos, Ryan D and Noam Gidron. 2016. "Intergroup behavioral strategies as contextually determined: Experimental evidence from Israel." *The Journal of Politics* 78(3):851–867.
- Enos, Ryan D and Noam Gidron. 2018. "Exclusion and Cooperation in Diverse Societies: Experimental Evidence from Israel." *American Political Science Review* 112(4):742–757.

- Fernandez, Sergio, Samuel Koma and Hongseok Lee. 2018. "Establishing the link between representative bureaucracy and performance: The South African case." *Governance* 31(3):535–553.
- Fisman, Raymond, Daniel Paravisini and Vikrant Vig. 2017. "Cultural proximity and loan outcomes." *American Economic Review* 107(2):457–92.
- Forbes, Hugh Donald. 1997. *Ethnic conflict: Commerce, culture, and the contact hypothesis*. Yale University Press.
- Gade, Daniel M and Vicky M Wilkins. 2012. "Where did you serve? Veteran identity, representative bureaucracy, and vocational rehabilitation." *Journal of Public Administration Research and Theory* 23(2):267–288.
- Gay, Claudine. 2002. "Spirals of trust? The effect of descriptive representation on the relationship between citizens and their government." *American Journal of Political Science* pp. 717–732.
- Grossman, Guy, Oren Gazal-Ayal, Samuel D Pimentel and Jeremy M Weinstein. 2016. "Descriptive representation and judicial outcomes in multiethnic societies." *American Journal of Political Science* 60(1):44–69.
- Hajnal, Zoltan L. 2001. "White residents, black incumbents, and a declining racial divide." *American Political Science Review* 95(3):603–617.
- Halevy, Jonathan. 2018. "Ethnic-based separation in maternity Departments in Israel—a balanced practical view." *Israel journal of health policy research* 7(1):70.
- Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos and Dimitrios Xefteris. 2019. "Does exposure to the refugee crisis make natives more hostile?" *American Political Science Review* 113(2):442–455.
- Herman, Tamar, Or Anabi, Ella Heller and Fadi Omar. 2018. *The Israeli Democracy Index 2018*. The Israeli Democracy Institute.

- Hjort, Jonas. 2014. "Ethnic divisions and production in firms." *The Quarterly Journal of Economics* 129(4):1899–1946.
- Hughes, Joanne, Andrea Campbell, Simon Lolliot, Miles Hewstone and Tony Gallagher. 2013. "Inter-group contact at school and social attitudes: evidence from Northern Ireland." *Oxford Review of Education* 39(6):761–779.
- Irish, Donald P. 1952. "Reactions of Caucasian Residents to Japanese-American Neighbors." *Journal of Social Issues* 8(1):10–17.
- Karim, Sabrina. 2019. "Restoring confidence in post-conflict security sectors: Survey evidence from Liberia on female ratio balancing reforms." *British Journal of Political Science* pp. 1–23.
- Karim, Sabrina. N.d. "Relational State Building in Areas of Limited Statehood: Experimental Evidence on Attitudes about the Police." *American political science review*. Forthcoming.
- Karim, Sabrina, Michael J Gilligan, Robert Blair and Kyle Beardsley. 2018. "International gender balancing reforms in postconflict countries: Lab-in-the-field evidence from the liberian national police." *International Studies Quarterly* 62(3):618–631.
- Karpowitz, Christopher F, Tali Mendelberg and Lee Shaker. 2012. "Gender inequality in deliberative participation." *American Political Science Review* 106(3):533–547.
- Kasara, Kimuli. 2014. "Does Local Ethnic Segregation Lead to Violence?: Evidence from Kenya." *Quarterly Journal of Political Science* 11(4):441–470.
- Keiser, Lael R, Vicky M Wilkins, Kenneth J Meier and Catherine A Holland. 2002. "Lipstick and logarithms: Gender, institutional context, and representative bureaucracy." *American political science review* 96(3):553–564.
- Keshet, Yael and Ariela Popper-Giveon. 2018. "Patient demands for ethnic-based separation in public hospitals in Israel: patients' and practitioners' perspectives." *Israel journal of health policy research* 7(1):44.

- Kingsley, John Donald. 1944. "Representative bureaucracy."
- Kopstein, Jeffrey S and Jason Wittenberg. 2009. "Does familiarity breed contempt? Inter-ethnic contact and support for illiberal parties." *The Journal of Politics* 71(2):414–428.
- Kteily, Nour S and Kaylene J McClanahan. 2019. "Incorporating Insights about Intergroup Power and Dominance to Help Increase Harmony and Equality Between Groups in Conflict." *Current Opinion in Psychology* .
- Lipsky, Michael. 1980. *Street-level bureaucracy: Dilemmas of the individual in public service*. Russell Sage Foundation.
- Lowe, Matt. 2018. "Types of contact: A field experiment on collaborative and adversarial caste integration." *Unpublished Manuscript* .
- Mansbridge, Jane. 1999. "Should blacks represent blacks and women represent women? A contingent" yes"." *The Journal of politics* 61(3):628–657.
- Maoz, Ifat. 2010. "Educating for peace through planned encounters between Jews and Arabs in Israel: A reappraisal of effectiveness." *Handbook on peace education* pp. 303–313.
- Meier, Kenneth J, Robert D Wrinkle and Jerry L Polinard. 1999. "Representative bureaucracy and distributional equity: Addressing the hard question." *The Journal of Politics* 61(4):1025–1039.
- Meier, Kenneth John. 1975. "Representative bureaucracy: An empirical analysis." *American political science review* 69(2):526–542.
- Mosher, Frederick C. 1968. *Democracy and the public service*. Oxford University Press on Demand.
- Mousa, Salma. 2019. "Building Tolerance: Intergroup Contact and Soccer in Post-ISIS Iraq." Stanford University.

- Nanes, Matthew J. 2018. "Policing in divided societies: Officer inclusion, citizen cooperation, and crime prevention." *Conflict Management and Peace Science* p. 0738894218802580.
- Neggers, Yusuf. 2018. "Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India." *American Economic Review* 108(6):1288–1321.
- Pagtolun-an, Imelda G and Jeffrey M Clair. 1986. "An experimental study of attitudes toward homosexuals." *Deviant Behavior* 7(2):121–135.
- Paluck, Elizabeth Levy. 2006. "Diversity training and intergroup contact: A call to action research." *Journal of Social Issues* 62(3):577–595.
- Paluck, Elizabeth Levy and Donald P Green. 2009. "Prejudice reduction: What works? A review and assessment of research and practice." *Annual review of psychology* 60:339–367.
- Paluck, Elizabeth Levy, Seth A Green and Donald P Green. 2019. "The contact hypothesis re-evaluated." *Behavioural Public Policy* 3(2):129–158.
- Pedahzur, Ami and Yael Yishai. 1999. "Hatred by hated people: Xenophobia in Israel." *Studies in Conflict and Terrorism* 22(2):101–117.
- Pepinsky, Thomas B, Jan H Pierskalla and Audrey Sacks. 2017. "Bureaucracy and service delivery." *Annual Review of Political Science* 20:249–268.
- Pettigrew, Thomas F. 1998. "Intergroup contact theory." *Annual review of psychology* 49(1):65–85.
- Pettigrew, Thomas F and Linda R Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of personality and social psychology* 90(5):751.
- Pitkin, Hanna F. 1967. *The concept of representation*. Vol. 75 Univ of California Press.
- Porat, Roni, Eran Halperin and Daniel Bar-Tal. 2015. "The effect of sociopsychological barriers on the processing of new information about peace opportunities." *Journal of Conflict Resolution* 59(1):93–119.

- Rao, Gautam. 2019. "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools." *American Economic Review* 109(3):774–809.
- Rasul, Imran and Daniel Rogger. 2015. "The impact of ethnic diversity in bureaucracies: evidence from the Nigerian civil service." *American Economic Review* 105(5):457–61.
- Riccucci, Norma M, Gregg G Van Ryzin and Cecilia F Lavena. 2014. "Representative bureaucracy in policing: Does it increase perceived legitimacy?" *Journal of public administration research and theory* 24(3):537–551.
- Riccucci, Norma M, Gregg G Van Ryzin and Huafang Li. 2016. "Representative bureaucracy and the willingness to coproduce: An experimental study." *Public Administration Review* 76(1):121–130.
- Rokem, Jonathan, Chagai M Weiss and Dan Miodownik. 2018. "Geographies of violence in Jerusalem: The spatial logic of urban intergroup conflict." *Political Geography* 66:88–97.
- Rosenbloom, David H and Jeannette G Featherstonhaugh. 1977. "Passive and active representation in the federal service: A comparison of blacks and whites." *Social Science Quarterly* 57(4):873–882.
- Rosner, Tal. 2016. "Heroes of Health: Israel's Healthcare System as a Model of Jewish-Arab Coexistence." 13 King David St., Jerusalem, Israel: .
- Samii, Cyrus. 2013. "Perils or promise of ethnic integration? Evidence from a hard case in Burundi." *American Political Science Review* 107(3):558–573.
- Samoocha, Sami. 2013. *Lo Shovrim et Ha-Kelim: Madad Yahasei Aravim-Yehudim Be-Yisrael 2012*. The University of Haifa and The Israeli Institute for Democracy.
- Scacco, Alexandra and Shana S Warren. 2018. "Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria." *American Political Science Review* pp. 1–24.

- Selden, Sally Coleman, Jeffrey L Brudney and J Edward Kellough. 1998. "Bureaucracy as a representative institution: Toward a reconciliation of bureaucratic government and democratic theory." *Representative Bureaucracy: Classic Readings and Continued Controversies* pp. 134–154.
- Shayo, Moses and Asaf Zussman. 2011. "Judicial ingroup bias in the shadow of terrorism." *The Quarterly Journal of Economics* 126(3):1447–1484.
- Smootha, Sammy. 2004. "Index of Arab–Jewish Relations in Israel." *Haifa, Israel: The Jewish–Arab Center, University of Haifa* .
- Steinmayr, Andreas. 2016. "Exposure to refugees and voting for the far-right:(unexpected) results from austria."
- Tankard, Margaret E and Elizabeth Levy Paluck. 2016. "Norm perception as a vehicle for social change." *Social Issues and Policy Review* 10(1):181–211.
- Tankard, Margaret E and Elizabeth Levy Paluck. 2017. "The effect of a Supreme Court decision regarding gay marriage on social norms and personal attitudes." *Psychological science* 28(9):1334–1344.
- Theobald, Nick A and Donald P Haider-Markel. 2008. "Race, bureaucracy, and symbolic representation: Interactions between citizens and police." *Journal of Public Administration Research and Theory* 19(2):409–426.
- Van Laar, Colette, Shana Levin, Stacey Sinclair and Jim Sidanius. 2005. "The effect of university roommate contact on ethnic attitudes and behavior." *Journal of Experimental Social Psychology* 41(4):329–345.
- Williamson, Scott. 2019. "Countering Misperceptions to Reduce Prejudice: An Experiment on Attitudes toward Muslim Americans." *Journal of Experimental Political Science* pp. 1–12.

- Yablon, Yaacov B. 2012. "Are we preaching to the converted? The role of motivation in understanding the contribution of intergroup encounters." *Journal of Peace Education* 9(3):249–263.
- Zeitsoff, Thomas. 2014. "Anger, Exposure to Violence, and Intragroup Conflict: A "Lab in the Field" Experiment in Southern Israel." *Political Psychology* 35(3):309–335.
- Zeitsoff, Thomas. 2018. "Anger, legacies of violence, and group conflict: An experiment in post-riot Acre, Israel." *Conflict management and peace science* 35(4):402–423.
- Zussman, Asaf. 2013. "Ethnic discrimination: Lessons from the Israeli online market for used cars." *The Economic Journal* 123(572):F433–F468.

ONLINE SUPPLEMENTARY MATERIALS

A Survey Instrument, Data Collection, and Descriptive Statistics

In this section I provide an elaborate description of the survey used to collect covariate and outcome measures. Original pre-registered copies of the survey are posted with EGAP.⁴² As noted in section 4.5 of the article's main text, Clinic X embedded within their ongoing survey 20 additional outcomes relating to social and political attitudes. Thus, in addition to the routinely collected data regarding service satisfaction detailed in Section A.1, during implementation period, Clinic X collected responses to a battery of demographic questions detailed in Section A.2, as well as responses to a battery of social-political questions detailed in Section A.3.

Further data regarding treatment assignment (i.e. treating doctor) was obtained through Clinic X's administrative records. Doctor identity was determined by name, and validated with the clinic CEO. Administrative records also included information regarding respondents diagnoses which I visually portray in Figure A1. Descriptive statistics of variables used for my analyses are provided in Table A1, and images of the SMS invitation and the survey platform sent to patients are presented in Figures A2a-A2b. The distribution of my main outcome variables are portrayed in Figure A3.

⁴²Link: <https://egap.org/registration/5216>

Table A1: Descriptive Statistics - Jewish Patients

Statistic	N	Mean	St. Dev.	Min	Max
Age	2164	32.465	24.690	0	96
Male	2164	0.463	0.499	0	1
Religiosity	2058	0.942	0.973	0	3
Ideology	1675	2.508	1.076	1	5
Education	1914	2.503	1.220	0	4
10 Day Survey	2164	0.360	0.480	0	1
Violence Dummy	2164	0.033	0.179	0	1
Arab Doc	2164	0.798	0.402	0	1
Thermometer Arab	1521	2.404	1.215	0	4
Thermometer Jewish	1516	3.514	0.734	0	4
Thermometer Worker	1485	2.292	1.183	0	4
Thermometer American	1473	3.081	0.936	0	4
Soc Dis Arabs	1285	3.202	1.911	0	6
Soc Dis Workers	1273	2.577	2.001	0	6
Soc Dis Americans	1268	3.830	1.923	0	6
Soc Dis Jews	1286	5.547	1.094	0	6
Trust Jews	1472	2.062	0.698	0	3
Trust Arabs	1461	1.582	0.925	0	3
Peace Arabs	1456	1.795	0.888	0	3
Peace Jews	1464	2.297	0.683	0	3
Outcome Index	1234	-0.034	0.821	-1.854	1.378

A.1 Satisfaction Questions

- Satisfaction from service in Clinic X
- Satisfaction from doctor
- Satisfaction from nurse
- Satisfaction from clinic clerk
- Satisfaction from rentgan technician
- Satisfaction from clinic cleanliness
- Additional comments (text)

A.2 Demographics

1. Sex
2. Age
3. Religion
4. Religiosity
5. Education
6. Right-Left scale

A.3 Embedded Intergroup Items

7. **Feeling thermometer:** (Jews, Arabs, Foreign workers, Tourists)
8. **To what extent do you agree with the following statement: One can trust most Arabs in Israel (0:3 scale)**
9. **To what extent do you agree with the following statement: One can trust most Jews in Israel (0:3 scale)**

10. **To what extent do you agree with the following statement: Most Arabs want to live in peace with Israel** (0:3 scale)
11. **To what extent do you agree with the following statement: Most Jews want to live in peace with Israel** (0:3 scale)
12. **Social Distance:** (Jews, Arabs, Foreign workers, Tourists, (0:6 scale))
13. **In your life, how often do you interact with Arabs** (0:3 scale)
14. **In your life, how often do you interact with Arabs** (0:3 scale)

A.4 Additional Information from Medical Records

- Date
- Clinic name
- Doctor name
- Nurse name
- The general quality ranking of the treating doctor⁴³
- Patients overall time in clinic
- General Diagnostic

⁴³This is a cumulative ranking score which doctors receive over time.

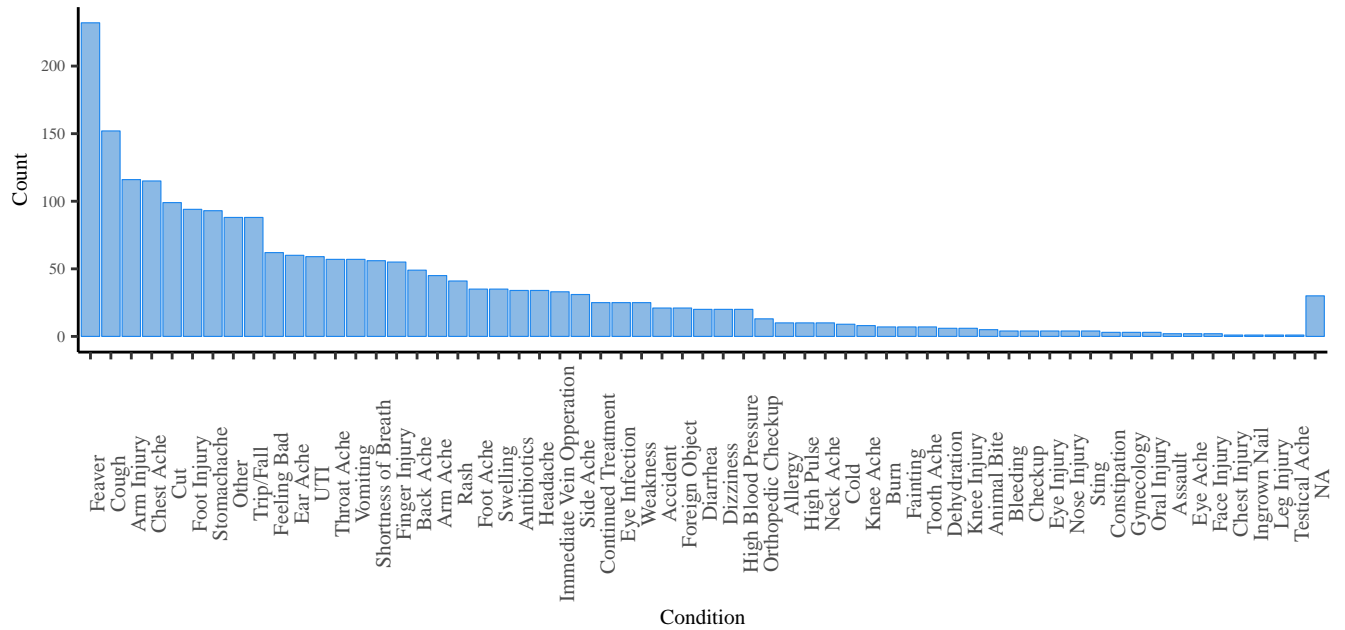


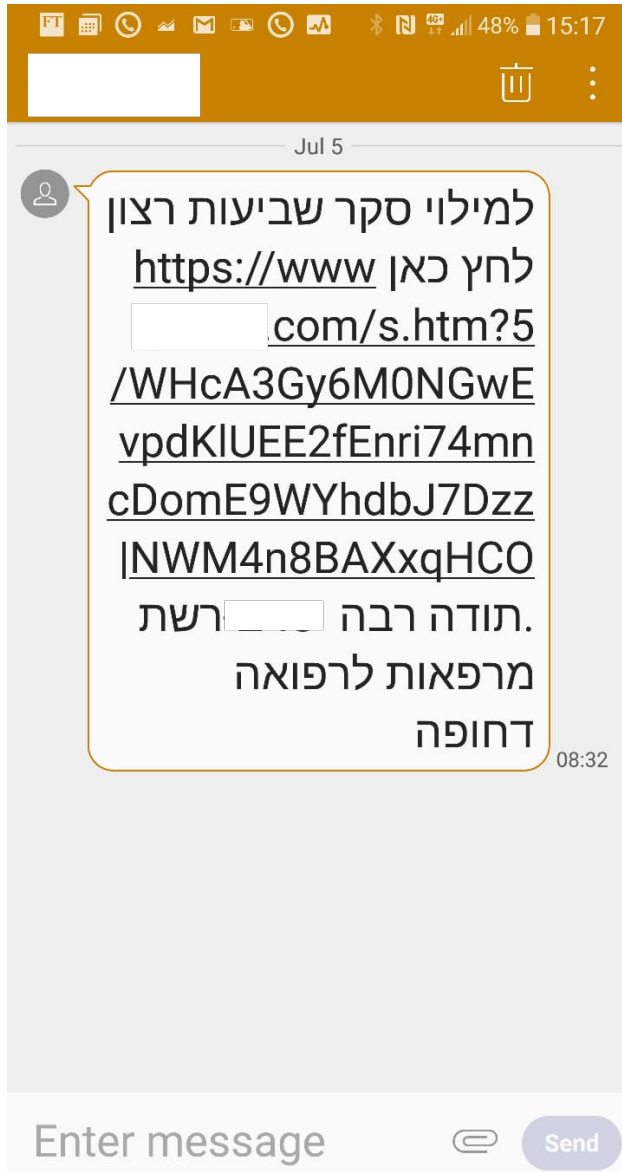
Figure A1: Illnesses in Clinics

B Design Diagnoses

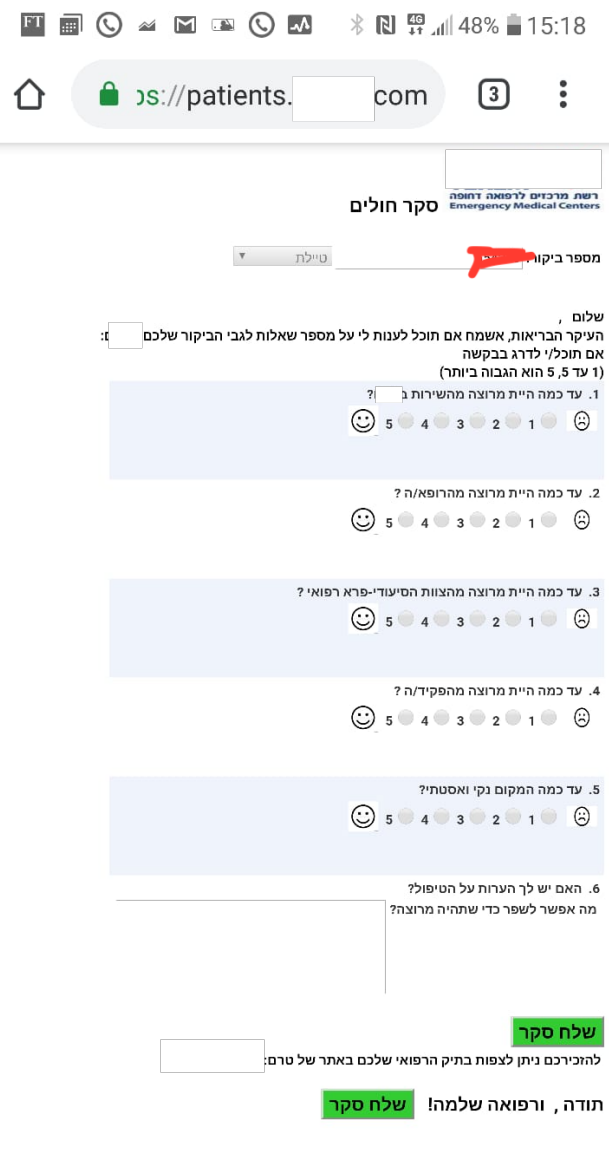
B.1 Treatment Assignment and Selection into Survey

Figure A4 visualizes Jewish respondents’ main demographics according to treatment condition. A quick view of this figure indicates that the share of treated Jews is larger than the share of Jewish respondents in the control condition. This is driven by the fact that Clinic X employs many Arab doctors, while catering to a great extent to the Israeli Jewish majority group. Nonetheless, the proportion of treated and controlled respondents across demographic categories seems to be rather well balanced.

In Table 1 of the main text, I test for balance more systematically. Results from a pre-registered linear model with clinic and date fixed effects, in which errors are clustered at the clinic level suggest that religiosity, education and age do not predict Jewish assignment to an Arab doctor. Nonetheless, gender seems to have a small, yet statistically significant correlation with Jewish assignment to an Arab doctor in my pre-registered OLS model. Therefore, as detailed in the main text of my paper, I consider several variations to all pre-registered models, controlling for gender



(a) SMS Invitation to Participate in Survey – Sent to all Patients



(b) Survey Platform

Figure A2: Illustration of Clinic X Facilities

(as well as other pre-treatment covariates). Doing so does not change my main results.

As indicated in Section 6 of my paper, I consider the possibility that assignment to an Arab doctor correlates with patients willingness to opt-in to my survey. This could potentially lead to serious biases in my main results. Nonetheless, the models reported in Table A2 indicate that receiving treatment from an Arab doctor does not correlate with higher probabilities of participation in the treatment evaluation surveys. More specifically, in both OLS (with clinic and date fixed effects, where errors are clustered by clinic) and Logit specifications where participation in evaluation surveys is a binary indicator, doctor assignment is substantively and statistically insignificant. That said, gender and age both correlate with participation in treatment evaluation surveys suggesting that my sample is older and more female dominant than the general population receiving treatment in Clinic X. I further evaluate the differences between my sample and the general public in section F of the supplementary materials.

Table A2: Selection into Survey by Treatment and Covariates

	OLS	OLS	Logit	Logit
Palestinian Doctor	0.00 (0.00)	-0.00 (0.00)	0.06 (0.05)	0.06 (0.06)
Age		0.00* (0.00)		0.00* (0.00)
Gender		-0.00* (0.00)		-0.15* (0.04)
Num. obs.	162588	155041	162588	155041
AIC			23988.43	22917.12
BIC			24008.43	22956.93
Log Likelihood			-11992.21	-11454.56
Deviance			23984.43	22909.12

* $p < 0.05$. OLS models include clinic and date fixed effects, and errors are clustered by clinic.

B.2 Non-Response to Main Outcomes

I further consider the extent to which treatment assignment predicts non-response to specific survey items. As mentioned in section 6 of my main text, finding that engagement with an Arab doctor correlates with non-response in the treatment evaluation survey would cast serious doubts on the

inferences made throughout my empirical investigations. Nonetheless, results reported in Table A3 serve to alleviate these concerns.

Indeed, it seems that non-responses are unrelated to treatment assignment. That said, several models suggest that men may be slightly more likely to skip specific questions. Therefore, it is unlikely that data is missing at random. Nonetheless, knowing that missingness is unlikely related to treatment serves to bolster the credibility of my main inferences.

Table A3: Correlation of Missing Responses with Treatment and Covariates

	Distance	Distance	Peace	Peace	Trust	Trust	Thermometer	Thermometer
Arab Doc	-0.00 (0.02)	-0.01 (0.03)	0.03 (0.02)	0.01 (0.03)	0.03 (0.02)	0.02 (0.03)	0.02 (0.02)	0.00 (0.03)
Religiosity		-0.01 (0.02)		-0.00 (0.02)		-0.00 (0.02)		-0.01 (0.01)
Gender		-0.05* (0.02)		-0.04* (0.02)		-0.03 (0.02)		-0.03 (0.02)
Education		-0.01 (0.01)		0.01 (0.01)		0.00 (0.01)		-0.01 (0.01)
Age		0.00 (0.00)		0.00 (0.00)		0.00 (0.00)		0.00 (0.00)
Num. obs.	2164	1878	2164	1878	2164	1878	2164	1878

* $p < 0.05$. OLS model including clinic and date fixed effects. Errors are clustered by clinic.

C Additional Models

C.1 Main Results

In this section, I supplement the graphical representation of my main models with regression tables. More so, in Table A4 I further consider the robustness of my main pre-registered models to alternative specifications controlling for the unbalanced gender indicator as well as additional pre-treatment covariates. Across all specification presented in Table A4 results remain consistent.

In addition, I provide table format results for Figure 4 in the main text regarding satisfaction from clinic service (Table A5). Like in the case of my main models, Table A5 presents models according to my original pre-registered specification, as well as models considering pre-treatment covariates. Results remain similar across the additional specification.

Table A4: Main Effect of Contact for Jewish Patients

	Distance	Distance	Distance	Peace	Peace	Peace	Trust	Trust	Trust	Therm	Therm	Therm	Index	Index	Index
Arab Doc	0.32*	0.32*	0.30*	0.10*	0.10*	0.10*	0.08	0.08	0.10	0.13	0.13	0.13*	0.14*	0.14*	0.14*
	(0.10)	(0.10)	(0.07)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)	(0.04)	(0.04)	(0.03)
Gender		0.08	0.09		0.01	0.03		-0.05	-0.02		-0.03	0.01		-0.00	0.02
		(0.05)	(0.08)		(0.04)	(0.05)		(0.04)	(0.04)		(0.05)	(0.05)		(0.03)	(0.04)
Religiosity			-0.74*			-0.18*			-0.22*			-0.27*			-0.27*
			(0.07)			(0.03)			(0.04)			(0.03)			(0.03)
Education			0.30*			0.11*			0.12*			0.13*			0.13*
			(0.04)			(0.02)			(0.02)			(0.02)			(0.02)
Age			0.00			0.00*			0.00*			0.01*			0.00*
			(0.00)			(0.00)			(0.00)			(0.00)			(0.00)
Num. obs.	1285	1285	1183	1456	1456	1342	1461	1461	1344	1521	1521	1395	1234	1234	1137

* $p < 0.05$. OLS models including clinic and date fixed effects. Errors are clustered by clinic.

Table A5: Doctor Effects on Satisfaction from Service

	Service	Service	Service	Doctor	Doctor	Doctor	Index	Index	Index
Arab Doc	-0.17*	-0.17*	-0.16*	-0.07	-0.07	-0.07	-0.10	-0.10	-0.09
	(0.06)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.06)	(0.06)	(0.06)
Gender		-0.01	0.03		0.04	0.07		0.01	0.05
		(0.06)	(0.06)		(0.06)	(0.06)		(0.05)	(0.05)
Religiosity			0.01			0.02			0.01
			(0.03)			(0.03)			(0.02)
Education			-0.04*			-0.02			-0.02
			(0.02)			(0.02)			(0.02)
Age			0.01*			0.01*			0.00*
			(0.00)			(0.00)			(0.00)
Num. obs.	2080	2080	1803	2091	2091	1814	2066	2066	1793

* $p < 0.05$. OLS models including clinic and date fixed effects. Errors are clustered by clinic.

C.2 Survey Timing and Decaying Effects

In this section I graphically present results discussed in Section 5.2 of the main article regarding decaying effects. Figure A5 presents results of pre-registered models on two different sub-samples of respondents receiving invitations to participate in the treatment evaluation survey either one or ten-days post treatment. Figure A6 presents similar models controlling for pre-treatment covariates. In these figures, each coefficient represents the average treatment effect of receiving service from an Arab doctor on a given outcome. Red (blue) estimates represent results from models implemented amongst respondents receiving the survey one (ten) day post treatment.

C.3 Heterogenous Treatment Effects

Figures A7-A11 report models exploring pre-registered heterogenous treatment effects across my main four outcomes. As mentioned in the main text, I fail to detect any heterogenous treatment effects, considering five different pre-treatment covariates. My inability to detect heterogenous treatment effects across all models may be driven, at least in part, due to limited power and a conservative estimation strategy. However, an alternative interpretation of these figures, further discussed in the main text, would suggest that the effects of contact are rather similar across different demographics in Israeli-Jewish society.

C.4 Doctor Quality

In this section, I consider three different approaches of controlling for doctor quality, and ensuring that it is not driving my main results. First, in Table A6 I replicate my main models controlling for pre-treatment doctor ratings. As mentioned in the main text, my rating variable is an aggregate average of scores that doctors receive from patients which were obtained prior to the implementation of this study. Generally, doctors' ethnic identity does not predict higher rankings.

As demonstrated in Table A6 rankings do not have a statistically significant effect on most outcomes (aside from trust). Thus when controlling for doctor quality my main results remain rather similar. More specifically, the effects of contact on social distance and the general composite index remain statistically significant in conventional terms. The effect on affect (i.e. thermometer) and peace related attitudes remain substantively similar, though p values increase, leading results to be only marginally significant. Like in my main results, effects of contact on trust are statistically insignificant.⁴⁴

A similar pattern is identified when adapting an alternative strategy to account for doctor quality. Thus, as reported in Table A7, when employing mixed-effects models with clinic and date fixed-effects, and doctor random-effects, results remain similar. Thus the effects of contact on social distance and the general composite index remain statistically significant in conventional terms.

⁴⁴These models were not pre-registered with EGAP.

Table A6: Effect of Contact for Jewish Patients Controlling for Doctor Quality

	Distance	Distance	Distance	Peace	Peace	Peace	Trust	Trust	Trust	Therm	Therm	Therm	Index	Index	Index
Arab Doc	0.27*	0.27*	0.28*	0.09	0.09	0.11	0.06	0.06	0.10	0.12	0.11	0.12	0.12*	0.12*	0.14*
	(0.11)	(0.11)	(0.10)	(0.05)	(0.05)	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)
Rating	0.15	0.15	0.19	0.06	0.06	0.07	0.10	0.10*	0.10*	0.05	0.05	0.06	0.06	0.06	0.09
	(0.10)	(0.10)	(0.10)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.08)	(0.08)	(0.06)	(0.04)	(0.04)	(0.04)
Gender		0.09	0.08		-0.01	0.01		-0.06	-0.03		-0.03	0.00		-0.01	0.00
		(0.07)	(0.07)		(0.04)	(0.05)		(0.04)	(0.05)		(0.06)	(0.06)		(0.04)	(0.04)
Religiosity			-0.68*			-0.17*			-0.20*			-0.25*			-0.25*
			(0.08)			(0.03)			(0.04)			(0.03)			(0.04)
Education			0.29*			0.11*			0.12*			0.13*			0.13*
			(0.05)			(0.02)			(0.02)			(0.02)			(0.02)
Age			0.00			0.00*			0.00*			0.01*			0.00*
			(0.00)			(0.00)			(0.00)			(0.00)			(0.00)
Num. obs.	1124	1124	1032	1274	1274	1173	1279	1279	1174	1335	1335	1222	1077	1077	990

* $p < 0.05$. OLS models including clinic and date fixed effects. Errors are clustered by clinic.

Concurrently, effects on other outcomes are directionally similar, though p-values increase when incorporating doctor random effects.⁴⁵

Table A7: Alternative Mixed Model Specification

	Distance	Distance	Peace	Peace	Therm	Therm	Trust	Trust	Index	Index
Arab Doc	0.32*	0.30*	0.10	0.11	0.14	0.14	0.08	0.10	0.14*	0.14*
	(0.15)	(0.14)	(0.07)	(0.07)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)
Gender		0.09		0.03		0.01		-0.02		0.02
		(0.10)		(0.05)		(0.06)		(0.05)		(0.05)
Religiosity		-0.74*		-0.18*		-0.27*		-0.22*		-0.26*
		(0.06)		(0.03)		(0.04)		(0.03)		(0.03)
Education		0.30*		0.11*		0.13*		0.12*		0.14*
		(0.04)		(0.02)		(0.03)		(0.02)		(0.02)
Age		0.00		0.00*		0.01*		0.00*		0.00*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
AIC	5321.30	4718.31	3961.94	3627.31	5008.45	4532.98	4073.13	3674.86	3156.87	2790.75
BIC	5893.89	5302.03	4553.69	4230.74	5599.76	5135.66	4665.26	4278.45	3724.97	3369.91
Log Likelihood	-2549.65	-2244.16	-1868.97	-1697.66	-2393.22	-2151.49	-1924.57	-1721.43	-1467.44	-1280.37
Num. groups: Doctor	209	205	214	209	216	211	217	212	206	202
Var: Doctor (Intercept)	0.07	0.03	0.01	0.01	0.03	0.02	0.02	0.02	0.02	0.02
Var: Residual	3.44	2.87	0.74	0.70	1.37	1.28	0.79	0.72	0.60	0.51

* $p < 0.05$. Mixed model including clinic and date fixed effects and doctor random effects.

Lastly, in Figure A12 I run a placebo test, considering the effects of intergroup contact with an Arab doctors on several outcomes which should not be affected by treatment. As discussed in Section 6 of the main text, the theoretically expected pattern by which intergroup contact with Arab doctors affects prejudice towards Arabs (see shaded light blue region in Figure A12), but not prejudice towards other groups, further strengthens my confidence that identified effects are driven by contact, rather than doctor quality.

⁴⁵These models were not pre-registered with EGAP.

C.5 Effects on Adult Patients

In Table A8, I consider the robustness of my finding to the omission of responses on behalf of minor patients. As noted in the main text, when minor patients receive treatment in Clinic X, they are accompanied by a legal guardian, which receives the survey invitation post-treatment. Since legal guardians might experience a different form of contact, in comparison to patients directly receiving medical treatment, I consider additional models omitting responses on behalf of minor patients. Results presented in Table A8 demonstrate the robustness of my findings to the exclusion of minor patients.

On a whole results remain substantively similar. Specifically, my effects of contact on social distance and the general attitude index (first and last columns of Table A8) are statistically significant. For my other outcomes results remain in a similar direction, however p-values increase. I expect this to be driven at least in part by the significant decrease in sample size.

Table A8: Effects of Contact on 18+ Patients

	Distance	Distance	Peace	Peace	Trust	Trust	Therm	Therm	Index	Index
Arab Doc	0.29*	0.36*	0.09	0.10	0.09	0.13*	0.15	0.17*	0.12*	0.15*
	(0.13)	(0.12)	(0.05)	(0.06)	(0.07)	(0.05)	(0.09)	(0.07)	(0.06)	(0.03)
Gender		0.09		0.00		-0.04		-0.06		-0.01
		(0.13)		(0.06)		(0.05)		(0.06)		(0.06)
Religiosity		-0.72*		-0.15*		-0.19*		-0.24*		-0.24*
		(0.10)		(0.03)		(0.05)		(0.04)		(0.04)
Education		0.30*		0.10*		0.13*		0.11*		0.13*
		(0.05)		(0.03)		(0.02)		(0.02)		(0.02)
Age		0.00		0.01*		0.01*		0.01*		0.01*
		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)
Num. obs.	795	728	915	840	917	842	951	870	763	699

* $p < 0.05$. OLS models including clinic and date fixed effects. Errors are clustered by clinic.

D Survey Comments

As part of Clinic X's treatment evaluation survey, patients are offered to leave comments relating to their treatment experience. For the most part, patients do not leave comments. However, a review of existing comments from the period of my study suggests that when responding with comments, Patients usually refer to long waiting times in clinics, dirty facilities, and accessibility and parking issues. Nonetheless, several comments reported bellow demonstrate patients' awareness to the ethnicity of doctors. The anonymized quotes bellow serve to provide anecdotal evidence regarding

patients' awareness to doctors' ethnicity, and the way such awareness leads to different expressed attitudes amongst patients.

- Such committed care. Dr. X, 'walla' [actually] took care of us professionally, politely, and personally there are no words to describe. Same for Y, the paramedic. Thanks.

Jewish male patient. November 2018.

- The wait was horrible and unreasonable! The clerk was watching movies on her phone? the doctor was an absolute angel!!! I am an ultra-orthodox [women] and I won't lie. [I] don't like Arabs. But this doctor Just touched me and I wish she gets promoted in some way. She is simply a real angel. I don't know her name. She is a young short Arab doctor. Truly amazing.

Jewish female patient. November 2018.

- Everyone was fine. It slightly disturbed me that [people] were speaking Arabic and I did not understand.

Jewish female patient. December 2018.

- Too expensive and one Arab doctor injected me with Optalgin [Metamizole] forcefully in my hand and it hurt, this is really unprofessional.

Jewish male patient. December 2018.

- Unfriendly Arab doctors and not professional. The place [clinic] does not consider the Jewish majority in the area, and chooses to promote an ideological agenda on the public's account.

Jewish women patient. December 2018.

- I have a very important comment there are women that come for a check-up, and it is unreasonable that there are only male doctors in a given shift, especially when these are religious women that are reluctant to engage in physical contact with men, especially when contact include intimate parts of their body. It is compulsory to employ female doctors and nurses at

night so it will be more comfortable for women to get checked, since (currently) you employ men who try to look professional, and they think we don't understand their language and they allow themselves to speak however they want. They look at women in a very provocative way and that makes the women they treat feel very awkward, and it is unreasonable that during such check-ups for women, there won't be a woman that will implement check-ups for women. I would appreciate if you take care of this promptly and in the most serious way since most of your patients are from the religious and ultra-orthodox sector. Thank you, and have a great day.

Jewish female patient. January 2019.

- The staff is confused and they speak Arabic and you think that they do not fill their duty or [that they are] speaking about you when I left I did not believe I received suitable medical care and I went to the emergency room to be checked [again].

Jewish male patient. December 2018.

- You should employ doctors who are more professional!!! The staff does not know to do anything apart from primary care. In addition, it is preferable to have at least one or two Jewish doctors. It is horrifying to be alone with six Arabs and not one Jew.

Jewish female patient. December 2018.

- We wait[ed] 40 minutes when we were told that there is only one doctor. They [the doctors] speak only Arabic so we will not be able to understand and that is a disgrace. I felt in a foreign country and not in Israel!!!!

Jewish male patient. January 2019.

- We were in the clinic [where there was] only Arab staff and that is not nice at all! They did not help in anything and we left there without medical care very unsatisfied!! In contrary in another clinic [there was] more normal treatment [the] Jewish staff the doctor were respectful and helpful.

Jewish female patient, December 2018.

- Comment, the doctor was Arab and he took care of me in the best way possible.

Jewish female patient. November 2018.

- We were satisfied from Dr. X. The care and approach were very humane and warm with a smile all the time. I wish the Jewish doctors would learn from him for example the nurse that took our EKG and blood tests were ok. (but there was a Russian nurse it was not pleasant to approach her? [her look was] sour and [she was] looking annoyed) oftentimes it is important to give them [the nurses] clear guidelines. We should all be healthy.

Jewish female patient. January 2019.

- Yes. They [the doctors] should not be arrogant and they should understand and accept anyone who is different from them[.] The doctor did not give the right care and I thought something would happen to me, and I told him I am afraid that something will happen to me and he really didn't care! Everything was careless without normal and quick identification [of medical conditions] and in the reception I tried to call the nurse they [the medical staff] were by me and they did not pay attention[.] apart from that I am fed up of receiving treatment in a way in which if I have a specific drug [taking specific medication] then [doctors] laugh at me and behave in a disrespectful way[.] the doctor must do his job in an ethical way if it were the opposite not that I would behave like this, the Arab doctor would have complained that I am not treating him well! And with regards to all the stigmas[.] that does not interest me[.] doctors and nurses are present in these places [clinics] to save lives and not to look at each other and I am saying this even though the situation in the country with Arabs is not very good but that doesn't interest me[.] a doctor must treat me suitably otherwise he doesn't deserve to be a doctor.

Jewish male patient. January 2019.

E A Discussion of Deviations from Original Pre-Analysis Plan

In Table A9 I explicitly mention all extensions and deviations from my original pre-analysis plan. These mainly include additional models which were employed to test the robustness of my pre-registered hypotheses. All deviations are mentioned in the main text, and Table A9 directs readers to the exact place where non pre-registered results are presented.

	Change	Deviation or Addition?	Relevant Figure/Table
1	In addition to OLS models predicting doctor assignment by patient demographics, I consider Logit models as well as a set of bivariate OLS models.	Addition	Table 1
2	To account for unbalanced covariates (i.e. gender) I consider additional OLS models controlling for gender, as well as other pre-treatment covariates.	Addition	Table A4
3	Models do not include doctor fixed-effects since doctor fixed-effects would soak up my treatment indicator	Deviation	All tables and figures measuring the effects of contact
4	I introduce a composite measure of intergroup attitudes as well as a similar composite measure of service satisfaction discussed in the main text.	Addition	Figure 3, 4, A12, and 5
5	I consider alternative models controlling for doctor quality, as well as doctor random-effects	Addition	Tables A6-A7
6	I consider an additional model controlling for patient covariates, when estimating the moderating effects of survey timing.	Addition	Figure A6
7	I consider an additional model omitting minor patients.	Addition	Table A8
8	I model selection into survey, as well as the correlates of missingness in outcome data.	Addition	Tables A2-A3

Table A9: Deviations and Additions to Pre-Analysis Plan

F Comparison with Public Opinion Data

As noted in Section 4 of the main manuscript, the high number of Arab doctors in the studied clinics may raise concerns that less-prejudice individuals are selecting into receiving medical treatment in Clinic X. This concern does not pose a threat to causal identification since as if random assignment of patients to doctors occurs once a patient selects into receiving service in Clinic X. However, if one believes that more tolerant Israeli Jews are selecting into Clinic X, and more tolerant Israeli Jews are also more likely to report reduction in prejudice post-treatment, my results may have weaker external validity.

In the main text, I emphasize qualitative evidence that leads me to expect selection into clinic to be a minor concern (Rosner 2016). This is since patients will likely encounter Arab service providers in most healthcare facilities they attend across the country. In addition, in many cases, Clinic X is the main service provider of basic medical care, and thus selection of this sort should not pose a major threat to the external validity of my findings.

Nonetheless, in order to empirically evaluate the comparability of my studied sample to a more representative sample of Israeli Jews, I leverage data from the 2018 Israeli Democracy Index (Hereinafter: IDI sample) (Herman et al. 2018). These data were collected in May 2018 through a telephone survey, approximately half a year before the collection of data in Clinic X. Below, I report results of a simple difference in means test comparing the Jewish portion of the Israeli Democracy Index sample ($n = 851$), to my sample of survey respondents ($n = 2164$).

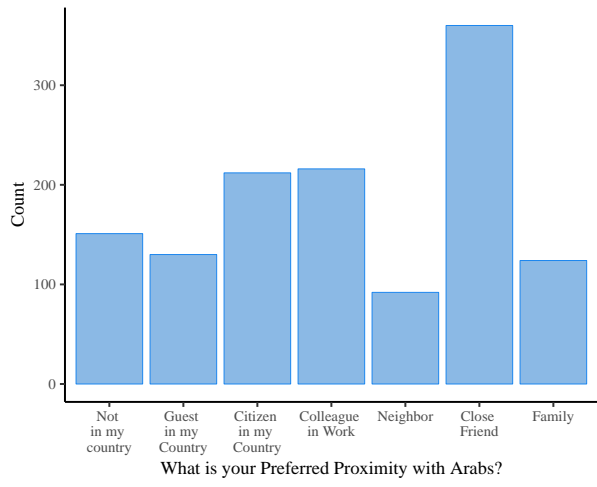
To compare between my sample, and the IDI sample, I leverage several similar items, including: gender, religiosity, education, and ideology (right-left scale). Comparing samples on gender, ideology, and religiosity is rather straightforward since possible responses are similar across both surveys. Nonetheless, in the education variable, I recode Clinic X survey responses by lumping the BA and MA categories, to ensure that education levels are similar to the IDI survey. Likewise, since the IDI collects data on age groups rather than raw age, I recode Clinic X's survey respondents according to six age groups: 18-24, 25-34, 35-44, 45-54, 55-64. In doing so, I code all responses on behalf of minors as missing, since I do not know the precise age of of the person

filling out the survey. Therefore, my comparison of both samples according to age should be taken with a grain of salt.

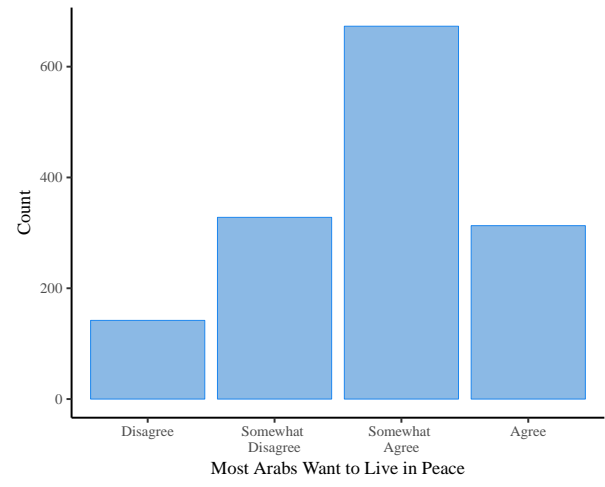
Table A10: Comparison of IDI 2018 and Clinic Survey Samples

	Variable	N Clinic	N IDI	Clinic Mean	IDI Mean	Low	High	p
1	Age	1381	851	3.57	3.67	-0.05	0.24	0.20
2	Male	2164	851	0.49	0.46	-0.07	0.01	0.21
3	Religiosity	2058	830	0.89	0.94	-0.03	0.13	0.19
4	Education	1914	847	2.29	2.27	-0.10	0.06	0.61
5	Ideology	1675	807	2.45	2.51	-0.04	0.16	0.26

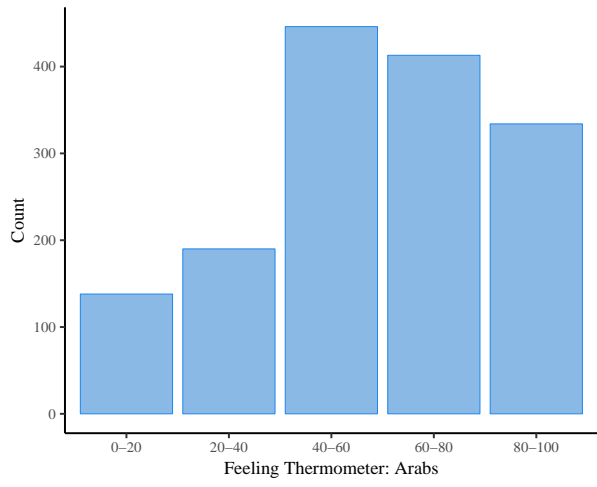
The difference in means tests presented in Table A10 suggest that on average, my sample is rather similar to the IDI Jewish representative sample surveyed in mid-2018. Indeed, I fail to reject the null hypothesis of equivalence across samples along all covariates. Given the stark relation between religiosity, ideology, and preferences for Arab exclusion in Israel (Enos and Gidron 2018), the similarity between both samples is reassuring, alleviating concerns regarding external validity.



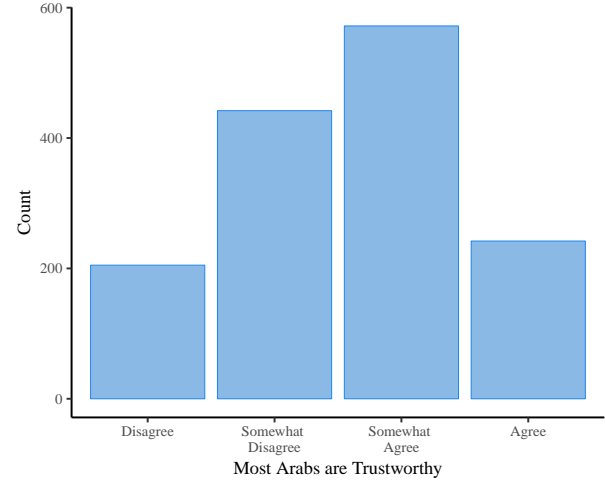
(a) Social Distance



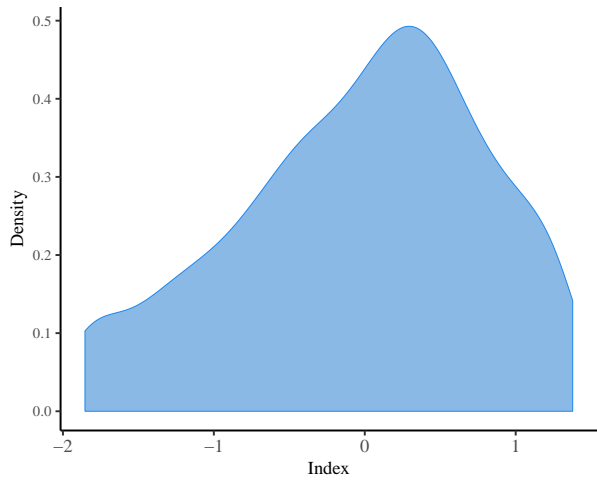
(b) Peace



(c) Thermometer



(d) Trust



(e) Index

Figure A3: Main Outcomes
SI-19

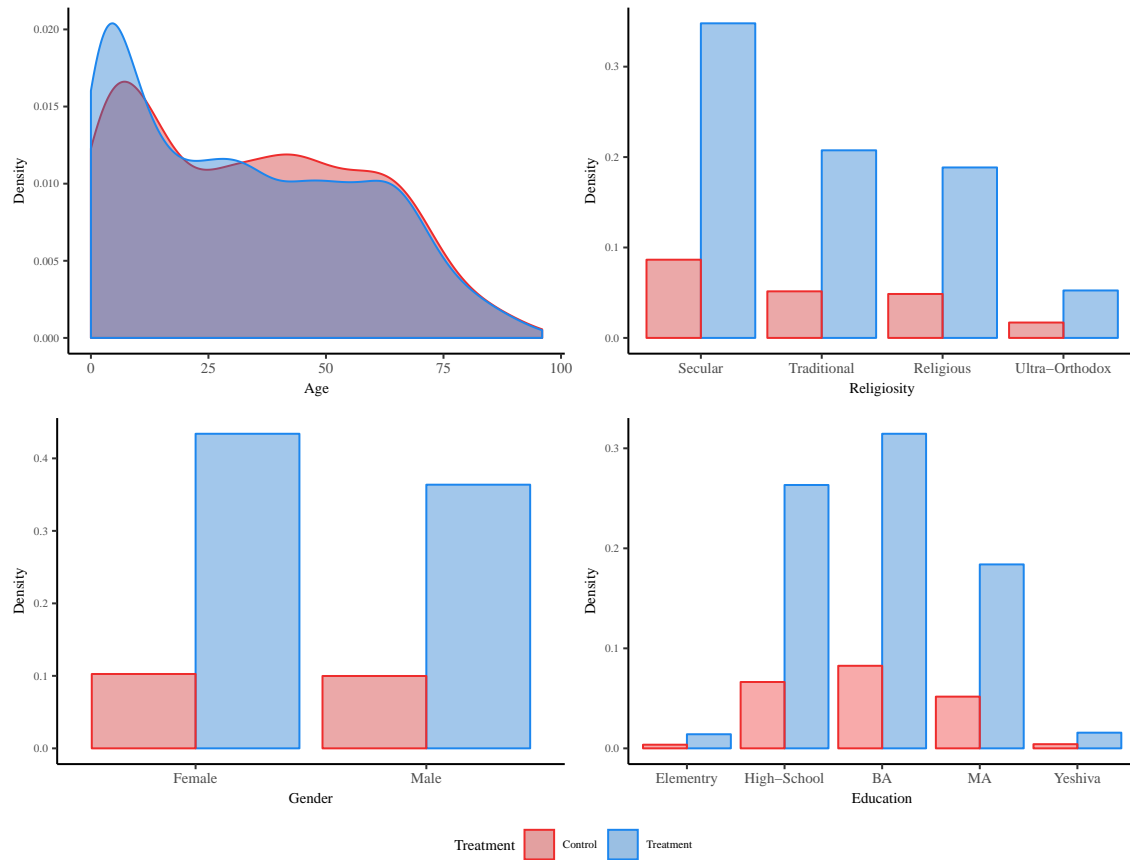
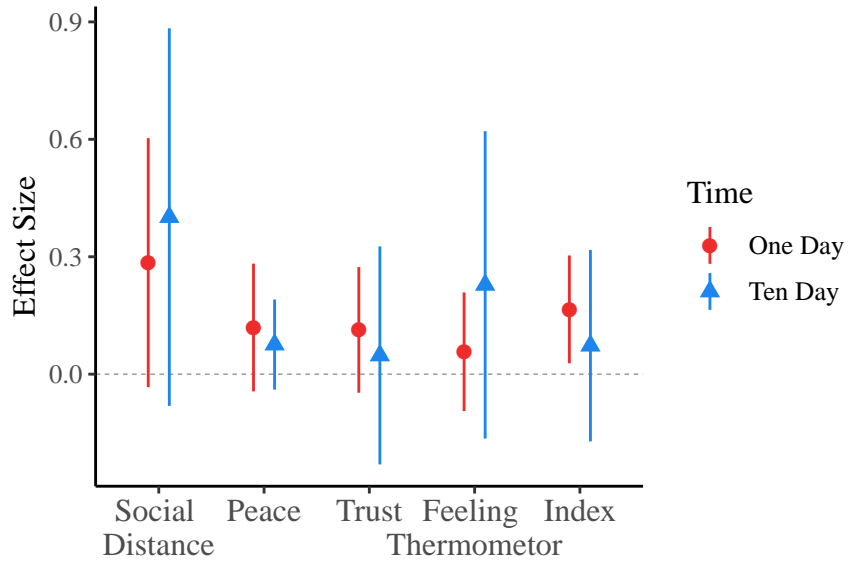
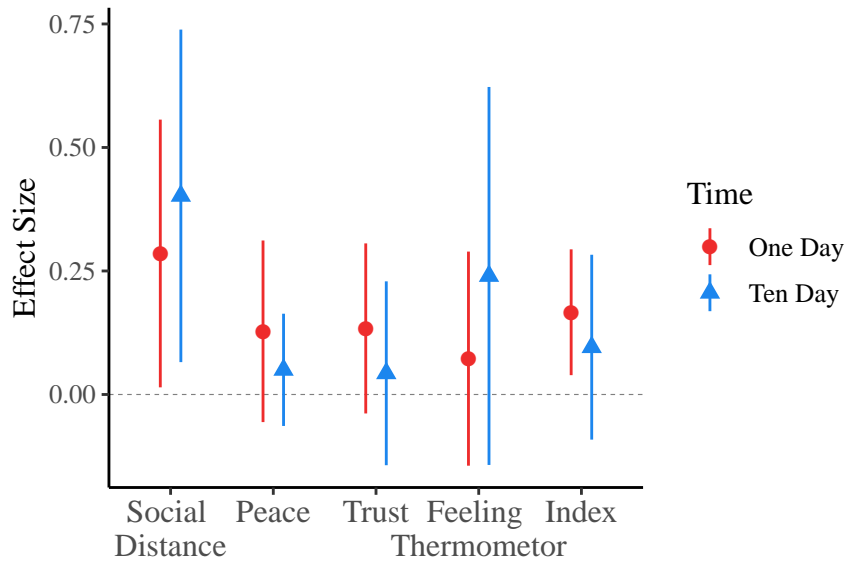


Figure A4: Demographics



*Arab doctor coefficients from linear regression models with clinic and date fixed effects. Standard errors are clustered by clinic.

Figure A5: Decay Effects of Contact



*Arab doctor coefficients from linear regression models with clinic and date fixed effects. Standard errors are clustered by clinic.

Figure A6: Decay Effects of Contact – Controlling for Pre-Treatment Covariates

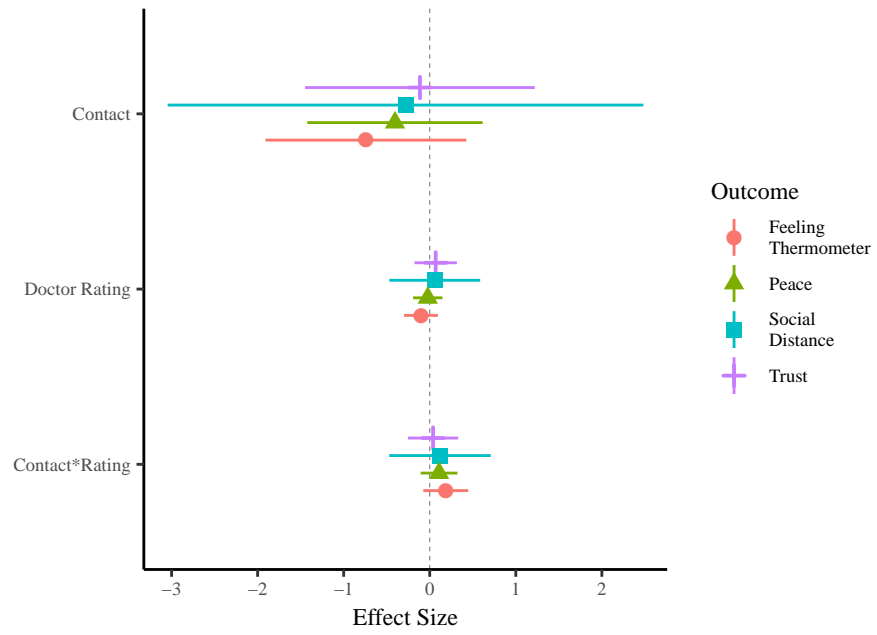


Figure A7: Heterogenous Effects of Contact Conditional on Doctor Quality

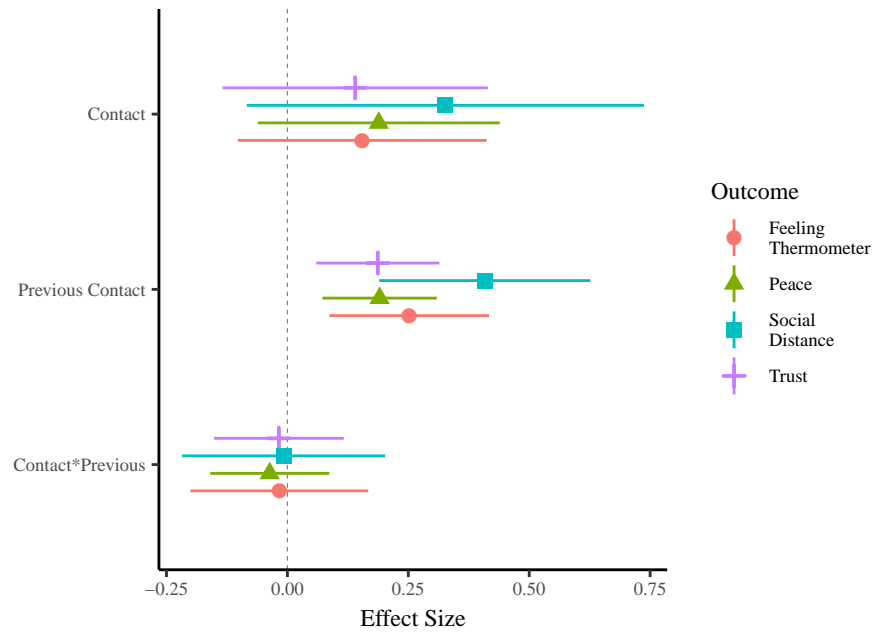


Figure A8: Heterogenous Effects of Contact Conditional on Previous Contact

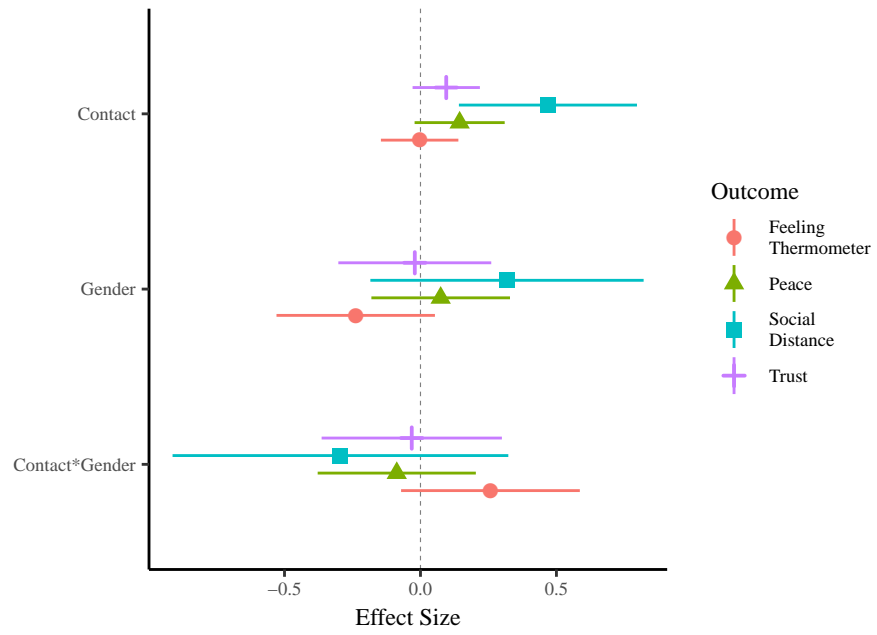


Figure A9: Heterogenous Effects of Contact Conditional on Gender

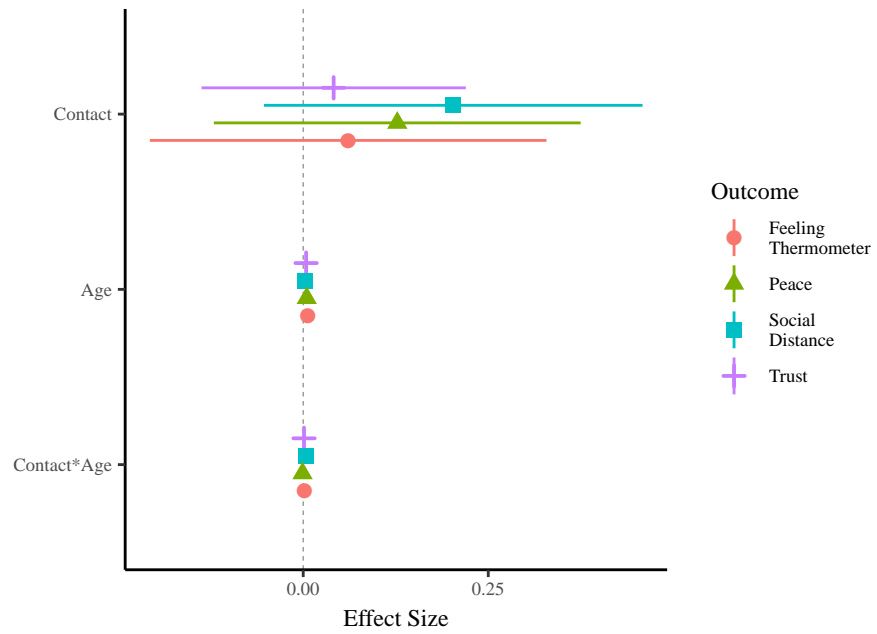


Figure A10: Heterogenous Effects of Contact Conditional on Age

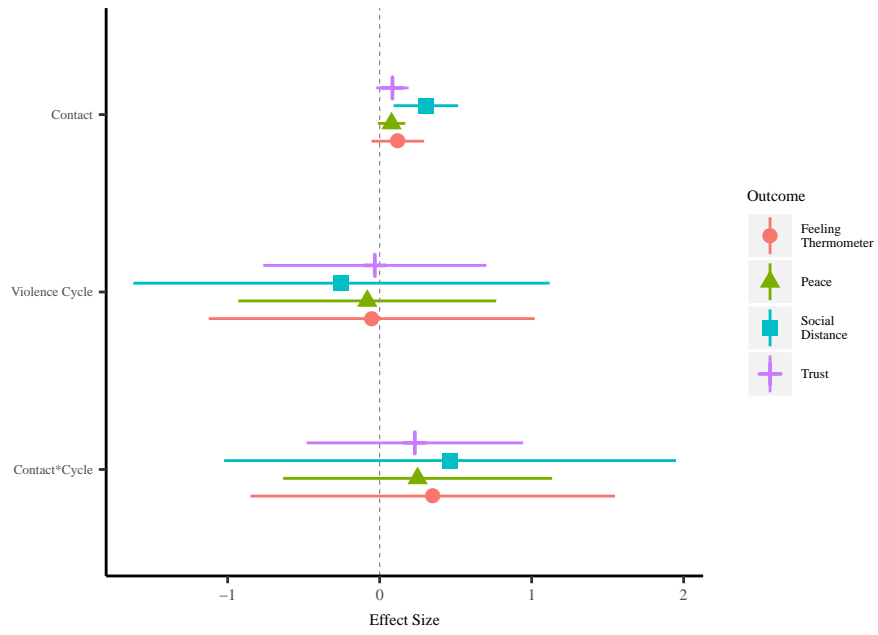


Figure A11: Effects of Contact During Cycles of Violence

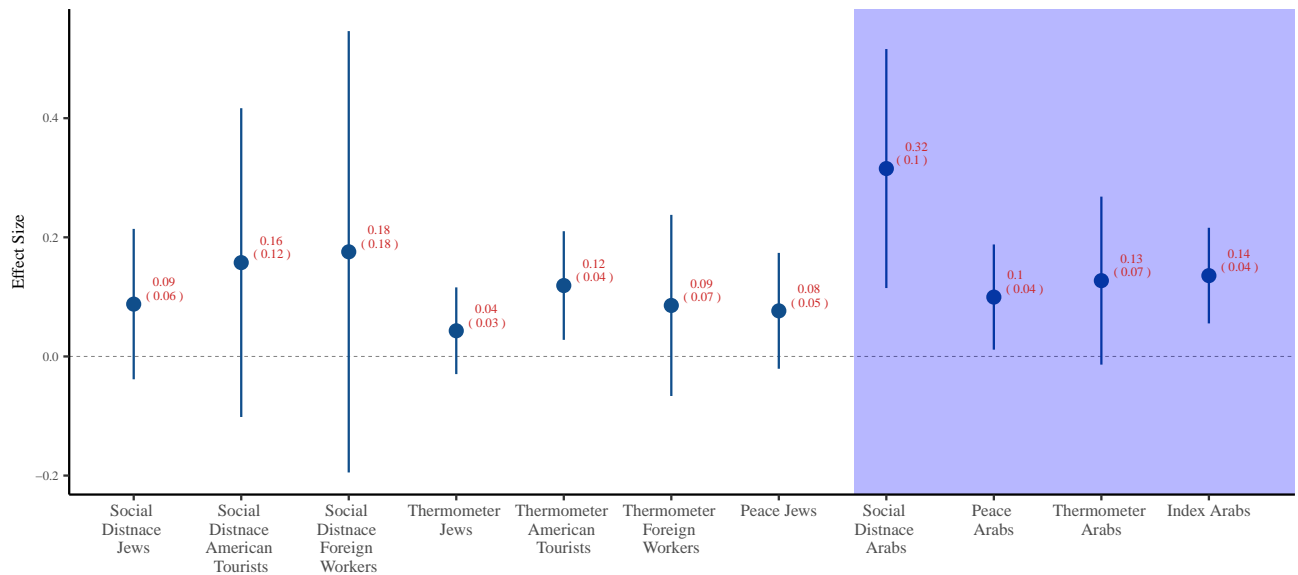


Figure A12: Placebo Test of the Effects of Contact with an Arab Doctor on Alternative Outcomes