CURING PREJUDICE THROUGH REPRESENTATIVE BUREAUCRACIES:

Evidence From A Natural Experiment in Israeli Medical Clinics

Chagai M. Weiss *

October 29, 2019

Abstract

Representative bureaucracies are associated with economic development, enhanced educational achievements, improved organizational performance, and increased institutional trust, cooperation, and responsiveness amongst minorities. Nonetheless, the extent to which such representation shapes majority attitudes towards minorities is unclear. Building on the representative bureaucracy and intergroup contact literatures, I develop a theory of prejudice reduction through bureaucratic representation, which I test with a natural experiment leveraging the random assignment of patients to doctors across 21 emergency medical clinics in Israel. Doing so, I demonstrate that representative bureaucracies which facilitate brief intergroup contact between Arab doctors and Jewish patients can improve majority-group perceptions of minorities. Contact with an Arab doctor reduces Jewish patients' social distance 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 representative bureaucracies 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, Simone Dietrich, Noam Gidron, Sharon Gilad, Don Green, Guy Grossman, Yoi Herrera, Amaney Jamal, Monica Komer, Kathy Cramer, Betsy Levy-Paluck, Salma Mousa, Nikos Nikiforakis, Jonathan Renshon, 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 and the identity across the subfields conference at MPSA 2019 for constructive feedback.

1 Introduction

Representation is a core value of modern democracy (Pitkin 1967), and scholars of electoral politics have demonstrated its positive effects on economic development (Chattopadhyay and Duflo 2004), deliberation (Mansbridge 1999; Karpowitz, Mendelberg and Shaker 2012), education (Clots-Figueras 2011), and participation (Bhalotra, Clots-Figueras and Iyer 2017; Gay 2002; Bhavnani 2009). Acknowledging the considerable discretion that non-elected bureaucrats enjoy in the process of policy implementation (Lipsky 1980), scholars have suggested extending the notion of representation beyond parliaments, to include representation within state bureaucracies. Indeed, bureaucracies comprised of workers that resemble the demographics they serve are associated with a host of positive outcomes, including: economic development (Bhavnani and Lee 2018), security (Nanes 2018), increased health (Alsan, Garrick and Graziani 2018), enhanced educational achievements (Keiser et al. 2002), improved organizational performance (Rasul and Rogger 2015; Fernandez, Koma and Lee 2018), reduced bias in voter registration processes (Neggers 2018), more lenient judging (Grossman et al. 2016), favorable loan offers (Fisman, Paravisini and Vig 2017), and increased institutional trust (Karim 2017), satisfaction (Gade and Wilkins 2012), cooperation (Riccucci, Van Ryzin and Li 2016), and perceived legitimacy (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.

While diversifying the ranks of a bureaucracy may improve tangible outcomes for minorities, its' externalities for intergroup relations more broadly can cut both ways. Put differently, the mere presence of minorities within state institutions may engender threat perceptions amongst members of a majority group (Key and Heard 1949). These perceptions may lead to prejudice and discriminatory attitudes that can escalate intergroup conflict. Nonetheless under specific conditions, the presence of minorities within state

institutions may foster positive interactions which in turn improve intergroup relations (Allport 1954).

In what follows, I theorize about a particular mechanism through which representative bureaucracies may improve intergroup relations – positive intergroup contact. I propose that employing minority bureaucrats may have positive externalities which go beyond enhanced public goods provision (Nanes 2018; Keiser et al. 2002; Gade and Wilkins 2012). More specifically, by enabling positive intergroup contact between minority bureaucrats and majority clients, representative bureaucracies may shape intergroup attitudes and improve intergroup relations. The far-reaching arms of state bureaucracies (Pepinsky, Pierskalla and Sacks 2017), and the repeated interactions that citizens experience with street-level bureaucrats (Lipsky 1980), can promote wide spread recurring contact, leading to sustainable and scalable changes in intergroup relations.

To test my theory of representative bureaucracy and prejudice reduction I exploit a natural experiment leveraging the random assignment of patients to doctors across 21 emergency medical clinics in Israel. Doing so, I demonstrate that bureaucracies 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).[1] Contact with an Arab doctor reduces Jewish patients' social distance (a common measure of social segregation first introduced by Bogardus (1933)) 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. Effects on social distance are comparable to almost a one unit change in religiosity, moving from ultra-orthodox to religious Jewish practices. Such an effect is remarkable given previously documented exclusionary attitudes towards Arabs Israelis amongst ultra-orthodox Jews (Enos and Gidron 2018). I also find modest effects of

[1]Throughout the paper I use the term Arab to refer to non-Jewish doctors from Palestinian as well as Druze and Bedouin decent whose mother tongue is Arabic. Since not all doctors identify as Palestinian, I decided to use this more general term.

contact on general feelings towards Arabs (measured through feeling thermometers), but limited evidence that brief contact can build intergroup trust. Additional analyses suggest that receiving treatment from an Arab doctor does not affect doctor satisfaction ratings, but might modestly affect general satisfaction from clinic services.

My theory and evidence contribute to two distinct literatures. First, by demonstrating the positive externalities of representative bureaucracies for intergroup relations, I demonstrate that beyond improving public goods provision for minority citizens (Keiser et al. 2002; Alsan, Garrick and Graziani 2018), a more diverse bureaucracy can improve intergroup relations on a large scale.[2] Second, given the limited evidence regarding the causal effects of intergroup contact on prejudice (Paluck, Green and Green 2017), I join an emerging literature which demonstrates the ways in which intergroup contact can be used to foster more positive intergroup relations (Scacco and Warren 2018; Rao 2019; Mousa 2018; Barnhardt 2009). 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 general preferences for segregation (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). Nonetheless, in line with recent qualitative policy reports (Rosner 2016), my evidence suggests that despite preferences for segregation, interactions between Jews and Arabs in medical facilities can foster favorable intergroup relations even in a context of an ongoing conflict where unfavorable intergroup attitudes have been shown to hamper cooperation and orderly institutional functioning (Enos and Gidron 2018; Shayo and Zuss-

[2]This argument is in line with understandings that affirmative action and representation in state institutions provides either net benefits (Nanes 2018), or does not result in reduced service quality for majority groups (Bhavnani and Lee 2019).

man 2011; Bar and Zussman 2017). The Israeli context often described as an intractable conflict (Bar-Tal 1998), serves as a hard case, suggesting that representative bureaucracies promoting positive intergroup contact may be helpful in reducing animosities between natives and immigrants, divided racial groups, or societies recovering from ethnic conflict.

2 Representative Bureaucracy and Public Goods Provision

The focus on representation in state bureaucracies 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 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).

The transition from passive to active representation has been shown to correlate with a host of individual and institutional variables including: bureaucrats' education, role acceptance, and tenure (Selden, Brudney and Kellough 1998), as well as organizational hierarchy, and stratification (Keiser et al. 2002). More broadly, existing theories suggest that for a bureaucrat to translate her experiences as a minority into concrete policy, she must acquire sufficient discretion in a policy area with relevance to the group she represents, and her actions must be explicitly linked to specific outcomes (Meier and Nicholson-Crotty 2006). Building on this theoretical framework, scholars have linked representative bureaucracies with improved policing (Nanes 2018), enhanced education (Keiser et al. 2002), beneficial health outcomes (Alsan, Garrick and Graziani 2018), more generous government

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 (Neggers 2018), and more lenient court decisions (Grossman et al. 2016).

Introducing the idea of symbolic representation, scholars have demonstrated that even when lacking action from a bureaucrat, the mere presence of a minority within an organization can foster organizational trust, perceived legitimacy, and willingness to coproduce amongst represented citizens (Karim 2017; Theobald and Haider-Markel 2008; Riccucci, Van Ryzin and Lavena 2014; Riccucci, Van Ryzin and Li 2016). Such effects occur even when a specific policy area is unrelated to the identity of the representative bureaucrat (Riccucci, Van Ryzin and Li 2016). It follows that by making room for a specific group within a given bureaucracy, organizations can signal inclusion to their clients. This inclusion is thought to have far reaching attitudinal and behavioral effects.

3 Can Representative Bureaucracies Reduce Prejudice?

While there is theory and evidence to suggest that representative bureaucracies have a positive effect on public goods provision for minorities, their effects on intergroup relations more broadly have been neglected. Identifying the effects of representation on majority attitudes is of key importance, since exposure to minorities within a given bureaucratic organization can engender threat perceptions. If that is the case, specific forms of representation may engender prejudice and animosity which can harm minorities. Alternatively, under certain conditions, representation may foster positive interactions between members of different groups, serving to build bridges in divided societies.

While scholars of representative bureaucracy have yet to consider the aforementioned question, 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. Con-

currently, Hajnal (2001) demonstrates that elected black mayors in the U.S. can reduce fears towards the black community, mainly amongst democratic constituents. Similarly, Beaman et al. (2009) provide evidence that exposure to female village council leaders can reduce implicit gender bias.

The evidence regarding minority elected politicians and intergroup attitudes enhances one's intuition regarding the importance of representation 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 bureaucracies with intergroup attitudes, they currently remain under-theorized and lacking rigorous empirical evidence.

3.1 Representation, Contact, and Prejudice Reduction

The study of representative bureaucracies mainly focuses on street-level bureaucrats which enjoy a considerable amount of discretion in policy implementation, while engaging face-to-face with ordinary citizens (Lipsky 1980). Given the frequency of direct engagement between street-level bureaucrats and ordinary citizens, intergroup contact is a natural channel through which representation may alter majority attitudes. Indeed, since the early twentieth century, hundreds of empirical investigations have associated intergroup contact with prejudice reduction (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 hy-*

pothesis, is based on Allport's empirical evidence and thorough review of pioneering social science research from the first half of the twentieth century. On the basis of his review of existing studies and a detailed classification of six ideal-type forms of contact, Allport suggested 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). Scholars 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). Following critical reviews of the existing literature which raised important reservations regarding the methodological rigor of existing evidence (Paluck 2006; Paluck, Green and Green 2017), a recent wave of field-experiments has demonstrated the positive causal effects of intergroup contact on both behaviors and attitudes (Ditlmann and Samii 2016; Broockman and Kalla 2016; Scacco and Warren 2018; Rao 2019; Mousa 2018; Barnhardt 2009).[3] 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. 2017; Enos 2017; Bazzi et al. 2019).

Building on the existing literature, I expect representative bureaucracies which directly engage with citizens to reduce prejudice through a mechanism of intergroup contact. As

[3]In 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 minorties.

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 (1) institutional support, (2) equal status interactions, (3) a common goal, and (4) potential to facilitate conceptions of common humanity between members of opposing groups.[4]

If so, the effects of embedding minorities within diverse institutions may have heterogenous effects. For example, representation within an extractive tax collecting agency or a police station, may have different consequences than representation among a hospital staff or a school. This is since contact between a policeman and a citizen may often be non-cooperative, 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 bureaucracies 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. Inspired by Allport's four conditions for effective contact, I consider positive interactions as *a direct face-to-face encounter, entailing cooperation around a shared goal or mutual assistance, through which an in-group can learn about and develop a sense of appreciation towards a member of her out-group*.[5]

Contact promoted by a representative bureaucracy is unique in that minorities engage with majority group members from an elevated social position. As acknowledged by Kteily and McClanahan (2019) most studies of intergroup contact are agnostic to social hierarchies, which likely play a central roll in the way contact leads to attitudinal and behavioral change. The unique hierarchy innate to contact between majority group members

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

[5]To 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.

and representative bureaucracies can be expected to amplify the effects of contact. Thus by promoting positive contact, directed towards a shared goal, under a recognized institution, majority members receiving service from a minority bureaucrat can "learn about the other" and develop more favorable attitudes towards out-groups. This effect will likely manifest even after very brief interactions, since the type of engagement facilitated by a representative bureaucracy is often novel and meaningful for majority members receiving service from a minority bureaucrat.

It is important to note that as a mechanism linking representation with intergroup attitudes, contact may often operate as a bundled treatment. This is since clients engaging with a minority bureaucrat experience intergroup contact, as well as general exposure to the presence of minorities within a bureaucracy, resembling a unique social hierarchy (Kteily and McClanahan 2019), or a normative institutional signal (Tankard and Paluck 2016). As further detailed in the following sections, I focus on a case where the share of minorities within a bureaucracy is rather high and salient. Therefore once engaging with the bureaucracy under investigation, awareness to minority representation is commonly acknowledged by clients regardless of the bureaucrat with whom they engage. It follows that attributing attitudinal change to contact per se is rather reasonable. This is since variation across clients in my study pertains to the experience of contact, while general exposure and awareness to a unique social hierarchies and norms is quite reasonably assumed to remain constant.

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 (Bar-Tal 1998). Like in many divided societies, prejudice, animosity, and preferences for 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). It follows that the Israeli case serves as a crucial and hard test for theories of representative bureaucracy and prejudice reduction.

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 diversity 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.[6] 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 facilitated by a representative bureaucracy is hierarchal in nature, I expect the positive engagement with an Arab superior to amplify the effects of contact in this case. Fourth, given the widely documented segregation in Israeli society (Bhavnani et al. 2014; Enos and Gidron 2016, 2018; Rokem, Weiss and Miodownik 2018), hospitals serve as a central hub where Jews and Arabs interact,[7] and qualitative policy reports suggest that such interactions can bridge gaps between Jews and Arabs (Rosner 2016).

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

[7]Hospitals 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 market (Rosner 2016).

4.2 Experimental Context

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.[8] 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.[9] Services include emergency treatment as well as other routine checkups.[10]

Dispersed across Israel, clinic X provides services to a diverse crowd of patients ranging from 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.[11] As indicated in the clinic's webpage, and further corroborated in qualitative 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.[12] Given the prominence of clinic X,[13] and given the intense intergroup animosity between

[8] I use the term *clinic X* since the medical clinics which I am collaborating with requested to keep their identity anonymous at this stage.

[9] For 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.

[10] Emergency medical services are subsidizes 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.

[11] An elaborate depiction of conditions treated by doctors in clinic X is portrayed in Figure A1 of my supplementary materials.

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

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

Jews and Arabs in Israel, this site serves as a crucial case for my theory of representative bureaucracy and prejudice reduction. More so, the assignment mechanism of patients to doctors which I now turn to discuss, allows for a credible identification of one channel through which representative bureaucracies may shape intergroup attitudes – intergroup contact.

4.3 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.[14] 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.[15] 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).[16] Wait times for treatment vary by clinic and day, and can range between several minutes to close 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

[14]Unfortunately, I do not have data regarding the prevalence of such requests. Nonetheless, qualitative interviews with clinic administrators suggest that that such requests are rare and not accommodated.

[15]It 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.

[16]Prior 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

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 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.[17]

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 pa-

[17]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. In addition, I consider an alternative logit specification. As evident from table A2 in my supplementary material, 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.

tients and doctors are aware of 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.

4.4 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.[18] The main items relating to intergroup attitudes which I use as outcomes in my analyses are presented in Table 1.

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

[18]The 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.

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). A general schema of my research design and data collection procedure is depicted in Figure 2.

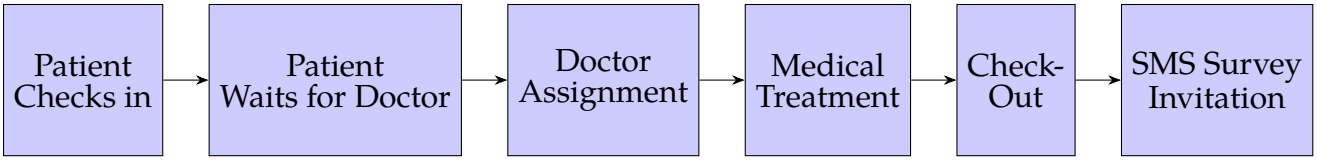


Figure 2: Research Design

4.5 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 treatment within 21 clinics over a period of three months,[19] by incorporating clinic, and day fixed effects. I cluster standard errors by doctor.[20] Equation 1 depicts my main model, a linear regression, where Y_{icd} denotes a survey response (as depicted in Table 1) 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.

[19]Data were collected between October 30, 2018 and January 30, 2019.

[20]In 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.

ϵ_{icd} represents my model's error term.

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

I employ this model to study the effects of contact between Jewish patients and Palestinian 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 1, and Hebrew translations can be found in my pre-registered materials.[21]

Item	Question	Answers
Social Distance	What is the maximum level of proximity which you would accept with (Arabs / Jews / Foreign Workers / Tourists)?	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
Peace	Do you agree with the following statement? "Most Arabs want to live in peace"	1. Disagree 2. Tend to disagree 3. Tend to agree 4. Agree
Feeling Thermometer	Please place the following social groups on a feeling thermometer (Arabs / Jews / Foreign Workers / Tourists)	1. 0 - 20 2. 20 - 40 3. 40 - 60 4. 60 -80 5. 80 - 100
Trust	Do you agree with the following statement? "It is possible to trust most Arabs in Israel"	1. Disagree 2. Tend to disagree 3. Tend to agree 4. Agree
Composite Index	Average z-scores of all outcomes	

Table 1: Main Outcomes

Though these four items measure distinct concepts, I combine them into a general index of intergroup attitudes by averaging their z-scores.[22] This intergroup attitude index

[21] See https://egap.org/registration/5216.

[22] This index has a 0.8 Corenbach's Alpha.

serves as my fifth outcome.[23] I expect my treatment to affect all five outcomes in a similar fashion enhancing positive intergroup attitudes. 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 A4-A5 of my supplementary materials.

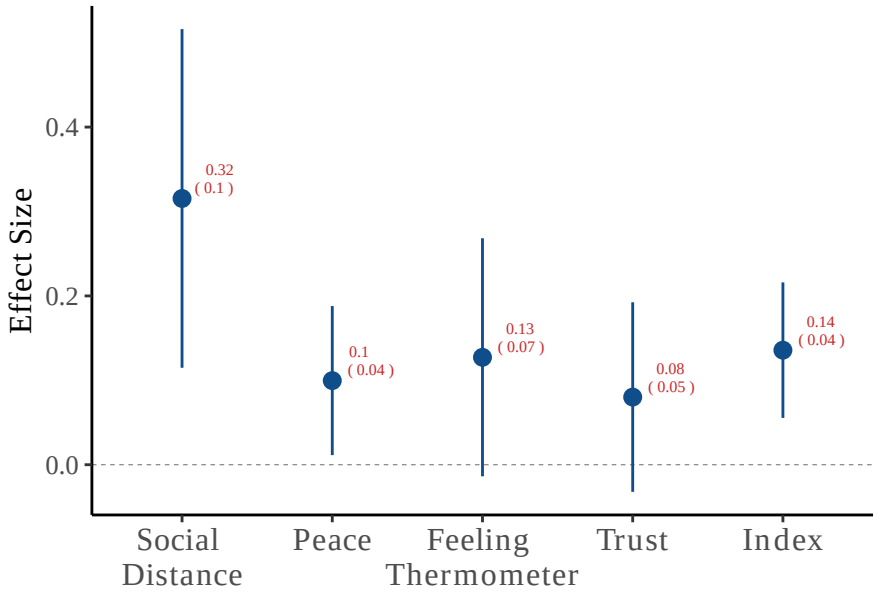
5 Analyses

5.1 Main Results

Results from my main analyses are depicted in Figure 3, as well as in Table A5 of the supplementary materials. Figure 3 plots the main coefficient of interest from five different models, representing the effect of contact on four distinct outcomes relating to intergroup attitudes, as well as a composite index of all four items. It suggests that receiving treatment from an Arab doctor has a positive effect on all outcomes. 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.

Thus Jewish patients who are treated by an Arab doctor report increased willingness to exist in proximity to Arabs. More specifically, 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 so-

[23] Note that I did not register this index in my pre-analysis plan.



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

Figure 3: Effects of Contact with Arab Doctors on Jewish Attitudes

cial distance, are equivalent to the effects of a one unit shift on conventional religiosity scales, moving from ultra-orthodox to religious Jewish practices. This comparison is remarkable given previously documented exclusionary attitudes towards Arabs amongst ultra-orthodox Jews, and given the strong links between exclusionary attitudes 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 unit increase in responses to a standard 4 item question. This effect is slightly larger than a one unit change in religiosity, as measured in the data.

The effects of contact on intergroup feelings do not meet conventional levels of statistical significance. However, since they are marginally significant ($p < 0.08$), one can carefully interpret them as further suggestive evidence that contact improves intergroup

attitudes as measured by commonly used feeling thermometers.[24] That said, unlike other forms of contact which have been shown to foster intergroup trust (Mousa 2018), brief interactions with an out-group doctor do not seem to build trust between Jews and Palestinians 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 treatment evaluation surveys either one or ten-days following their visit in a local clinic. My main results pool over both conditions. However, Figure A6 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.[25]

The null effects across all models, and my inability to detect a significant difference across timing conditions are likely driven by the power decrease resulting from splitting my sample into two timing conditions. In addition, though assignment to timing conditions was randomized, individuals opting into the survey when approached immediately after treatment, may differ from individuals opting into the survey ten days following treatment.[26] 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 A6 controlling for standard observed demographics.[27] Results presented in figure A7 of my supplementary materials still fail to detect a difference between survey responses in the immediate and longer time-period following treatment. Therefore, I am currently in the process of fielding a follow up with all respondents (6-8 months post-treatment), to provide a decisive understanding of the decay-effects of con-

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

[25] In 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.

[26] Indeed, it seems that the ten day condition includes more religious subjects.

[27] gender, religiosity, age, and education.

tact.

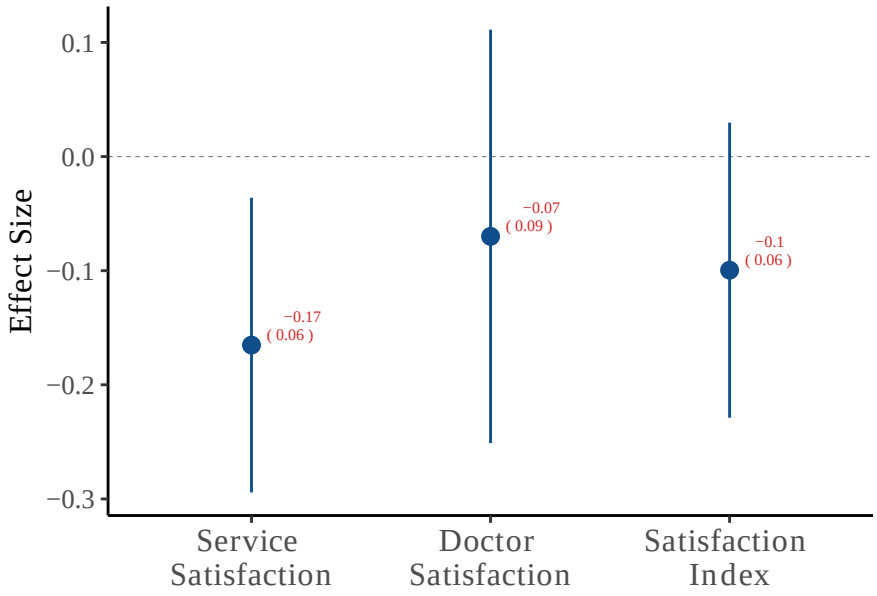
5.3 Does Attitudinal Change Come at the cost of Service Satisfaction?

The findings above suggest that the presence of minority bureaucrats 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 2019), or even provide the general public with net-benefits (Nanes 2018).[28]

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 a Palestinian doctor does indeed reduce general satisfaction from the service received in the clinic by about a tenth of a standard deviation. Nonetheless, similar effects do not seem to register with regards to satisfaction from the 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 while embedding minorities within a bureaucracy has positive externalities for intergroup attitudes, such externalities may come at the cost of decrease in general satisfaction from the

[28] Meier, Wrinkle and Polinard (1999) make a similar argument examining Latino student performance as a function of their teachers' identity. Nonetheless, a more recent replication has pointed to several specification mistakes in their original data analyses (Nielsen and Wolf 2001).

organization, but not from the specific street-level bureaucrat. 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.



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

Figure 4: Effects of Arab Doctors on Satisfaction from Service

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,[29] Previous exposure to Arabs, Ideology,[30] and experience of contact during a period of escalated violence in the Gaza strip.[31]

[29] As measured by aggregate doctor ratings collected before the implementation of the study.

[30] Though these models should be taken with a grain of salt, since ideology may be effected by contact, introducing post-treatment bias.

[31] To 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 implement by Hjort (2014)

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 subsets of the Israeli population. Given the prevalence of segregation 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

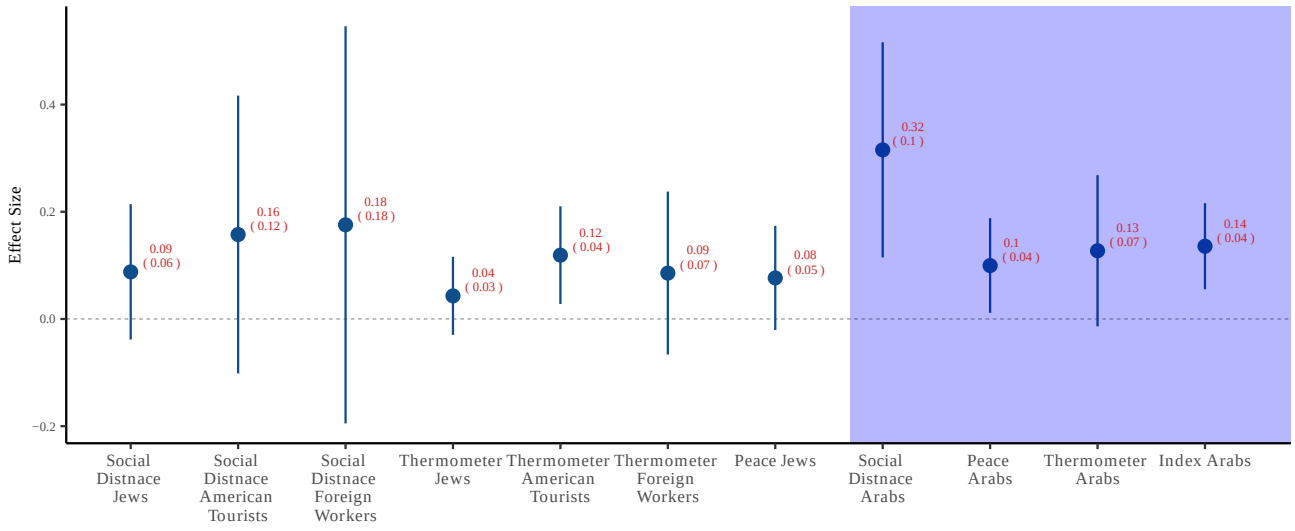
A skeptical interpretation of the findings presented above may raise concerns that doctor specific qualities which correlate with Arab identity may drive my main results. This interpretation would suggest that unobserved doctor characteristics and not intergroup contact per-se account for my main effects. To relax such concerns I implement a battery of 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 feelings (measured through feeling thermometers) towards Jews, American tourists, and foreign workers, as well as a question regarding Jewish willingness to live in peace. As depicted in Figure 5, unlike attitudes towards Arabs (shaded in light blue), 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

when examining the way a cycle of violence moderated the effects of ethnic diversity on firm productivity in Kenya.

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

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



*Arab doctor coefficients from linear regression models with clinic and date fixed effects. Standard errors are clustered by clinic and presented in parentheses. Shaded region represents outcomes that should be effected by treatment.

Figure 5: Placebo Test of the Effects of Contact with an Arab Doctor on Alternative Outcomes

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.

6.2 Unbalanced Covariates

As mentioned in Section 4.3, 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 Section B of my supplementary materials I test this assumption with several pre-treatment covariates. Indeed, across different specifications, it seems that 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 A5 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 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 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 2017).

While all patients are invited to participate in evaluation surveys, only a minority of them opt-in to report outcomes of interest.[35] 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 rule out 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.[36]

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.[37] Selective non-

[34] Models incorporating covariates were not specified in my pre-analysis plan.
[35] Approximately 1% of all patients fully participate in the evaluation survey.
[36] These models were not included in my original pre-analysis plan.
[37] More so, forcing responses may potentially lead to lower-quality responses and more severe attrition.

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 towards Arabs, the effects identified throughout the paper may be over-exaggerated. Luckily, I relax these concerns in Section B through a series of models which show that treatment assignment does not predict non responses to my main outcome measures.

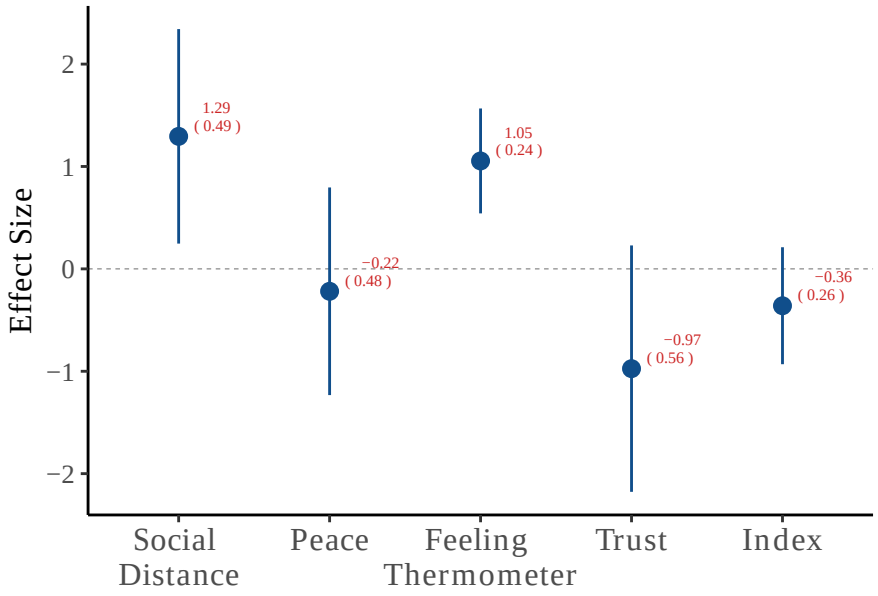
6.4 Effects on Arab Patients

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.[38]

With these caveats in mind, Figure 6 presents 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.[39] These results suggest that contact with out-groups embedded within a bureaucracy *may* have positive effects not only for minority, but also for majority group members. However,

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

[39]However, 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.



*Jewish doctor effects coefficients from linear regression models with clinic and date fixed effects. Standard errors are clustered by clinic and presented in parentheses.

Figure 6: Effects of Jewish Doctors on Palestinian Attitudes

future research leveraging a more appropriate setting is required to further establish this claim.

7 Future Extensions

I plan to elaborate on my existing evidence through three additional extensions relating to the measurement of behavioral outcomes, the implementation of semi-structured qualitative interviews, and an exploration of long-term effects. More specifically, I seek to better understand whether brief intergroup contact can shape costly political behavior. Recent evidence regarding the links between attitudinal and behavioral effects of contact are inconclusive. Some interventions have been shown to affect implicit and explicit attitudes (Barnhardt 2009), whereas others mainly shaped behaviors (Scacco and Warren 2018). Therefore, I hope to collaborate with clinic X to target patients with SMS based invitations to sign petitions protesting the prevailing hostile references towards Arabs in election campaigns. In line with my findings that contact with an Arab doctor can re-

duce social distance amongst Jewish patients, I expect brief interactions between Jewish patients and Arab doctors to encourage costly political behavior in the forms of signing a petition promoting intergroup tolerance.

As discussed throughout this paper, intergroup contact serves as one channel linking representative bureaucracies with improved intergroup attitudes. However, in order to better understand the mechanisms through which contact leads to prejudice reduction, I will conduct semi-structured qualitative interviews with Clinic X doctors. In doing so, I seek to learn about the nature of interactions between patients and doctors, and the ways through which contact reduces prejudice. I consider this empirical portion as an inductive endeavor through which I will articulate specific hypotheses to be tested in future research. Doing so will contribute to the intergroup contact literature which has thus far under-theorized the mechanism linking contact with reduced prejudice (Pettigrew and Tropp 2006).

Lastly, clinic X is currently working on re-approaching all survey respondents in order to test the longer-term effects of contact (i.e. 6-8 months post-treatment). Doing so will enable me to test the persistence of originally identified effects. While it is unlikely that a brief interaction with a doctor can cause an effect which persists for as long as 6-8 months, it is possible that contact encourages Jews to engage with Arabs more often in their daily routine. Such behavioral adaptations may sustain (or even amplify) the original effect of contact with an Arab doctor. To test this proposition, clinic X will add an additional question to the follow up survey, asking individuals to report how often they have engaged with Arabs over the past 6 months.

8 Conclusion

In this paper I consider a crucial, yet under-investigate question: How do representative bureaucracies shape intergroup relations? In doing so, I develop a theory regarding one particular mechanism through which representative bureaucracies may improve intergroup relations – positive intergroup contact. I develop my theory by synthesizing

insights from the representative bureaucracy literature (Lipsky 1980; Mosher 1968; Meier 1975; Selden, Brudney and Kellough 1998), together with theory and evidence regarding the contact hypothesis (Allport 1954; Pettigrew and Tropp 2006; Paluck, Green and Green 2017). Doing so, I propose that bureaucracies employing 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.

Leveraging a natural experiment in Israeli medical clinics, I demonstrate that bureaucracies that promote positive contact can improve intergroup relations on a large scale. More specifically, exploiting the random assignment of patients to doctors, I demonstrate that Jewish patients treated by Palestinian 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 representative bureaucracies 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 within the ranks of a bureaucracy may cut both ways. Thus not all contact with state bureaucrats is positive. More so, institutional diversity may engender threat perceptions and ignite concerns regarding affirmative action and labor market competition. Therefore, future research should more directly identify the conditions under which representation promotes positive intergroup relations. Doing so is especially important, in order to ensure that representation which promotes enhanced provision, does not come at the cost of intergroup animosity.

I also make four contributions to the prejudice reduction and intergroup contact liter-

ature. 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.

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). Nonetheless, I demonstrate that under certain conditions even very brief interactions can reduce prejudice and improve intergroup relations. Fourth, by linking the contact hypothesis with an institutional framework of representative bureaucracy, I take the contact literature a step forward. Thus, my evidence suggests that scholars of prejudice reduction should further consider how institutions, organizations, and bureaucracies that facilitate innumerable and repeated instances of intergroup contact may serve to amplify the effects of contact, and promote prejudice reduction in a scalable and sustainable way.

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 Palestinians 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

Allport, Gordon W. 1954. *The nature of prejudice*. Unabridged, 25th anniversary ed ed. Reading, Mass.: Addison-Wesley Pub. Co.

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.

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, Raghabendra 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. 2019. "Does Affirmative Action Worsen Bureacratic Performance? Evidence from the India Administrative Service.".

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.

Chattopadhyay, Raghabendra 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 Xefteris 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.

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. 2017. "Does Exposure to the Refugee Crisis Make Natives More Hostile?".

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. 2017. "Restoring confidence in post-conflict security sectors: Survey evidence from Liberia on female ratio balancing reforms." *British Journal of Political Science* pp. 1–23.

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.

Key, Valdimer Orlando and Alexander Heard. 1949. *Southern politics in state and nation*. Vintage Books New York.

Kingsley, John Donald. 1944. "Representative bureaucracy.".

Kopstein, Jeffrey S and Jason Wittenberg. 2009. "Does familiarity breed contempt? Interethnic 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.

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 and Jill Nicholson-Crotty. 2006. "Gender, representative bureaucracy, and law enforcement: The case of sexual assault." *Public Administration Review* 66(6):850–860.

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. 2018. "Overcoming the Trust Deficit: Intergroup Contact and Associational Life in post-ISIS Iraq.".

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.

Nielsen, Laura B and Patrick J Wolf. 2001. "Representative bureaucracy and harder questions: A response to Meier, Wrinkle, and Polinard." *The Journal of Politics* 63(2):598–615.

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 Green and Donald P Green. 2017. "The Contact Hypothesis Revisited.".

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 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.

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., Jeruslaem, Israel: .

Samooha, 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.

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.

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.

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.

Supplementary Materials

A Survey Instrument, Data Collection Procedures, 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.[40] As noted in section 4.4 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 A2-A3. The distribution of my main outcome variables are portrayed in Figure A4.

[40] Link: https://egap.org/registration/5216

Table A1: Descriptive Statistics - Jewish Patients

Statistic	N	Mean	St. Dev.	Min	Max
Age	2164	32.477	24.693	0.110	96.096
Male	2164	0.463	0.499	0	1
Religiosity	2084	1.042	1.317	0	9
Right Left	1914	2.503	1.220	0	4
Education	1349	2.135	0.865	0	3
Previous Contact Arabs	1346	2.958	0.268	0	3
Previous Contact Jews	2164	0.360	0.480	0	1
10 Day Survey	2164	0.033	0.179	0	1
Violence Dummy	2164	0.798	0.402	0	1
Arab Doc	1521	2.404	1.215	0	4
Thermometer Arab	1516	3.514	0.734	0	4
Thermometer Jewish	1485	2.292	1.183	0	4
Thermometer Worker	1473	3.081	0.936	0	4
Thermometer American	1285	3.202	1.911	0	6
Soc Dis Arabs	1273	2.577	2.001	0	6
Soc Dis Workers	1268	3.830	1.923	0	6
Soc Dis Americans	1286	5.547	1.094	0	6
Soc Dis Jews	1461	1.582	0.925	0	3
Trust Jews	1472	2.062	0.698	0	3
Trust Arabs	1456	1.795	0.888	0	3
Peace Arabs	1464	2.297	0.683	0	3
Peace Jews	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. **One can trust most Arabs in Israel**
9. **One can trust most Jews in Israel**
10. **Most Arabs want to live in peace with Israel**

11. **Most Jews want to live in peace with Israel**

12. **Social Distance**: (Jews, Arabs, Foreign workers, Tourists)

13. **In your life, how often do you interact with Arabs**

14. **In your life, how often do you interact with Arabs**

A.4 Additional Information from Medical Records

- Date
- Clinic name
- Doctor name
- Nurse name
- The general quality ranking of the treating doctor[41]
- Patients overall time in clinic

[41] This is a cumulative ranking score which doctors receive over time.

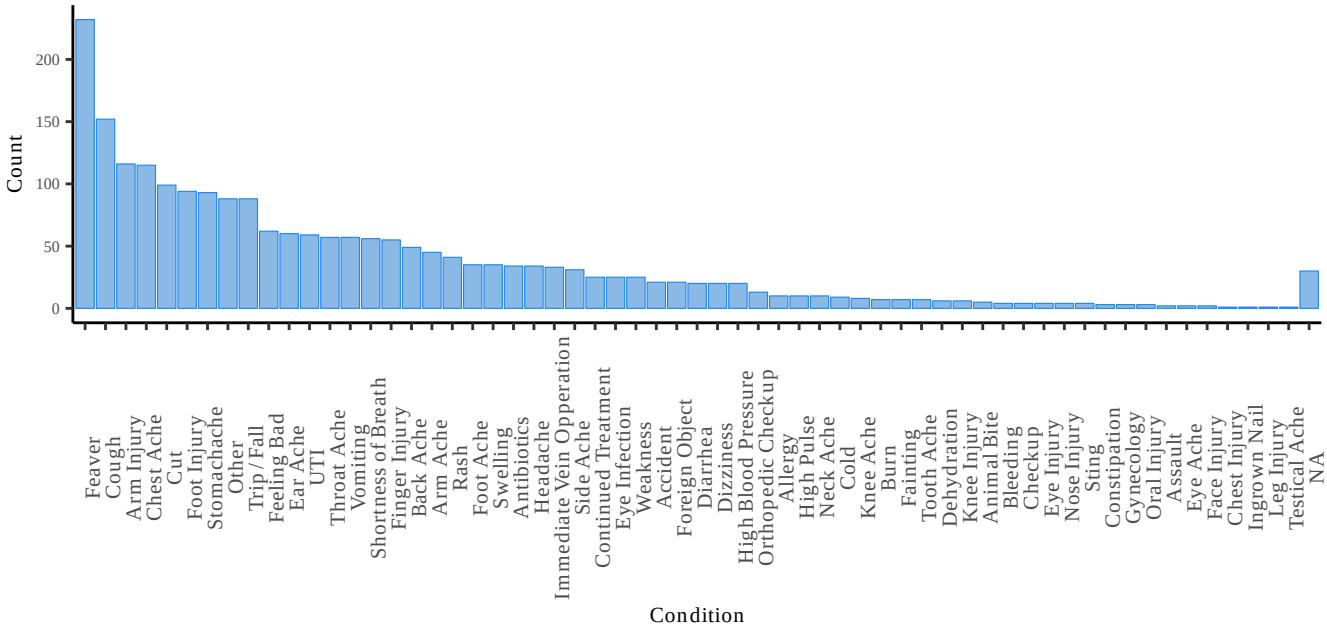


Figure A1: Illnesses in Clinics

B Design Diagnoses

B.1 Treatment Assignment and Selection into Survey

Figure A5 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 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 A2 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 ethnicity, education and age do not predict assignment to an Arabt doctor. Nonetheless, gender seems to have a small, yet statistically significant correlation with assignment to a Palestinian doctor. Therefore, as detailed in the main text of my paper, I consider variations to all pre-registered models, controlling for gender (as well as other

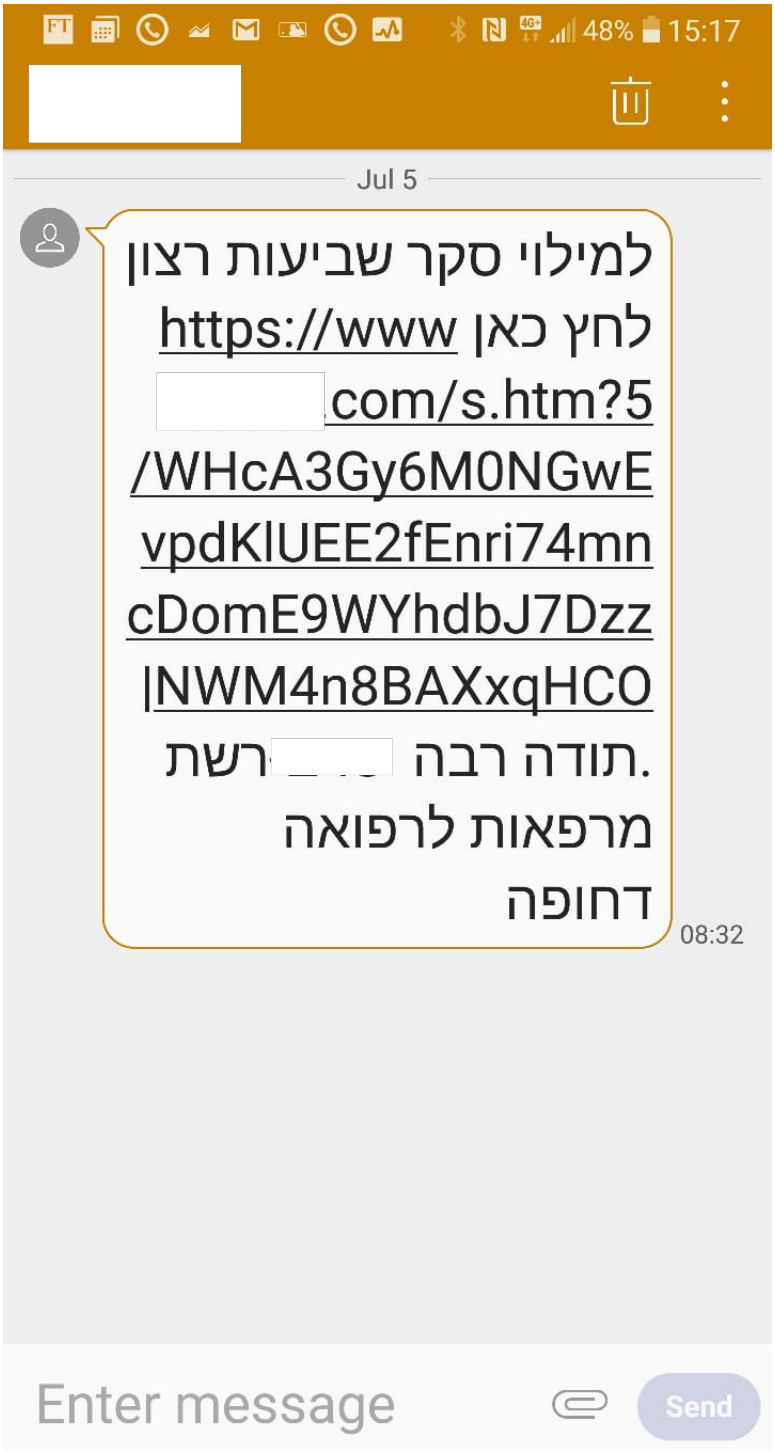


Figure A2: SMS Invitation to Participate in Survey – Sent to all Patients

pre-treatment covariates). Doing so does not change my main results.[42]

Table A2: Demographic Correlation with Assignment to Arab Doctor

	OLS	Logit
Religiosity	−0.01	−0.06
	(0.01)	(0.04)
Gender	−0.03*	−0.21
	(0.01)	(0.11)
Education	−0.01	−0.04
	(0.01)	(0.05)
Age	−0.00	−0.00
	(0.00)	(0.00)
Num. obs.	1903	1903
AIC		1954.34
BIC		1982.09
Log Likelihood		-972.17
Deviance		1944.34

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

As indicated in Section 6 of my paper, I consider the possibility that specific treatment assignments (i.e. contact with 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 A3 indicate that receiving treatment from an Arab doctor does not correlate with higher probabilities of participation in treatment evaluation surveys. More specifically, in both OLS (with clinic and date fixed effects) and Logit specifications where participation in evaluation surveys is a binary indicator, doctor assignment is substantively and significantly 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.

[42]Since my treatment indicator is binary, I further consider a Logit specification in Table A2. Results are similar, while gender does not seem to have a statistically effect in this model which does not include fixed effects.

Table A3: Selection into Survey by Treatment and Covariates

	OLS	OLS	Logit	Logit
Palestinian Doctor	0.00	−0.00	0.06	0.06
	(0.00)	(0.00)	(0.05)	(0.06)
Age		0.00*		0.00*
		(0.00)		(0.00)
Gender		−0.00*		−0.15*
		(0.00)		(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 model includes clinic and date fixed effects.

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 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 A4 serve to alleviate these concerns.

Indeed, it seems that non-responses is 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.

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 A5 I further consider the robustness of my main pre-

Table A4: Correlation of Missing Responses with Treatment and Covariates

	Distance	Distance	Peace	Peace	Trust	Trust	Thermometer	Thermometer
Arab Doc	−0.00	−0.01	0.03	0.01	0.03	0.01	0.02	−0.00
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Religiosity		−0.00		0.00		−0.00		−0.01
		(0.01)		(0.01)		(0.01)		(0.01)
Gender		−0.05*		−0.04*		−0.03		−0.04*
		(0.02)		(0.02)		(0.02)		(0.02)
Education		−0.00		0.01		0.01		−0.00
		(0.01)		(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	1903	2164	1903	2164	1903	2164	1903

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

registered models to alternative specifications controlling for the unbalanced gender indicator as well as additional pre-treatment covariates. Across all specification presented in Table A5 results remain consistent. In addition, I provide a table format results for Figure 4 regarding satisfaction from clinic service, which can be found in Table A6. Like in the case of my main models, Table A6 presents models according to my original pre-registered specification, as well as models considering pre-treatment covariates.

Table A5: 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*	0.33*	0.10*	0.10*	0.11*	0.08	0.08	0.08	0.13	0.13	0.12	0.14*	0.14*	0.14*
	(0.10)	(0.10)	(0.09)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)
Gender		0.08	0.12		0.01	0.04		−0.05	−0.02		−0.03	0.01		−0.00	0.03
		(0.05)	(0.07)		(0.04)	(0.05)		(0.04)	(0.04)		(0.05)	(0.05)		(0.03)	(0.04)
Religiosity			−0.35*			−0.07*			−0.10*			−0.10*			−0.12*
			(0.07)			(0.02)			(0.02)			(0.03)			(0.03)
Education			0.31*			0.11*			0.13*			0.13*			0.14*
			(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	1198	1456	1456	1359	1461	1461	1362	1521	1521	1415	1234	1234	1152

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

C.2 Survey Timing

In this section I graphically present results reported in Section 5.2 of the main article. These models and their interpertation are discussed in the main text. Figure A6 presents results

Table A6: 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	−0.06	−0.10	−0.10	−0.09
	(0.06)	(0.06)	(0.07)	(0.09)	(0.09)	(0.09)	(0.06)	(0.06)	(0.06)
Gender		−0.01	0.03		0.04	0.07		0.01	0.04
		(0.06)	(0.06)		(0.06)	(0.07)		(0.05)	(0.05)
Religiosity			−0.02			−0.00			−0.01
			(0.02)			(0.02)			(0.02)
Education			−0.04*			−0.03			−0.03
			(0.02)			(0.02)			(0.02)
Age			0.01*			0.00*			0.00*
			(0.00)			(0.00)			(0.00)
Num. obs.	2080	2080	1828	2091	2091	1839	2066	2066	1818

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

of pre-registered models on two different sub-samples of respondents receiving invitations to participate in the treatment evaluation survey either one or ten-day post treatment. Figure A7 presents similar models controlling for pre-treatment covariates.

C.3 Heterogenous Treatment Effects

Figures A8-A13 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 6 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 two different approaches of controlling for doctor quality, and ensuring that it is not driving my main results. First, in Table A7 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 A7 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 trust 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.[43]

Table A7: 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*	0.30*	0.09	0.09	0.11	0.06	0.06	0.07	0.12	0.11	0.10	0.12*	0.12*	0.13*
	(0.11)	(0.11)	(0.12)	(0.05)	(0.05)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)
Rating	0.15	0.15	0.18	0.06	0.06	0.07	0.10	0.10*	0.09*	0.05	0.05	0.05	0.06	0.06	0.08
	(0.10)	(0.10)	(0.09)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.08)	(0.08)	(0.07)	(0.04)	(0.04)	(0.04)
Gender		0.09	0.12		−0.01	0.02		−0.06	−0.02		−0.03	0.01		−0.01	0.01
		(0.06)	(0.06)		(0.04)	(0.05)		(0.04)	(0.04)		(0.06)	(0.07)		(0.04)	(0.04)
Religiosity			−0.31*			−0.07*			−0.10*			−0.10*			−0.11*
			(0.06)			(0.02)			(0.03)			(0.03)			(0.03)
Education			0.29*			0.10*			0.13*			0.12*			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	1046	1274	1274	1188	1279	1279	1190	1335	1335	1240	1077	1077	1004

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

A similar pattern is identified when adapting an alternative strategy to account for doctor quality. Thus, as reported in Table A8, 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. Concurrently, effects on other outcomes are directionally similar, though p-values increase when incorporating doctor random effects.[44]

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.

[43]These models were not pre-registered with EGAP.
[44]These models were not pre-registered with EGAP.

Table A8: Alternative Mixed Model Specification

	Distance	Distance	Peace	Peace	Therm	Therm	Trust	Trust	Index	Index
Arab Doc	0.32^*	0.33^*	0.10	0.11	0.14	0.13	0.08	0.09	0.14^*	0.15^*
	(0.15)	(0.15)	(0.07)	(0.07)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)
Gender		0.12		0.04		0.01		−0.02		0.03
		(0.11)		(0.05)		(0.06)		(0.05)		(0.05)
Religiosity		-0.35^*		-0.07^*		-0.10^*		-0.10^*		-0.12^*
		(0.04)		(0.02)		(0.02)		(0.02)		(0.02)
Education		0.31^*		0.11^*		0.13^*		0.13^*		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	4863.21	3961.94	3706.02	5008.45	4634.48	4073.13	3768.18	3156.87	2888.91
BIC	5893.89	5448.37	4553.69	4310.90	5599.76	5238.79	4665.26	4373.31	3724.97	3469.57
Log Likelihood	-2549.65	-2316.60	-1868.97	-1737.01	-2393.22	-2202.24	-1924.57	-1768.09	-1467.44	-1329.45
Num. groups: Doctor	209	205	214	210	216	212	217	213	206	202
Var: Doctor (Intercept)	0.07	0.08	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.03
Var: Residual	3.44	3.07	0.74	0.72	1.37	1.31	0.79	0.75	0.60	0.54

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

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 [clinics] does not consider the Jewish majority in the area, and chooses to promote an ideological agenda on the publics 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 women 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 nor-

mal 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 ethnical 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.	Addition	Table A2
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 A5
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, 5, and 6
5	I consider alternative models controlling for doctor quality, as well as doctor random-effects	Addition	Tables A7-A8
6	I consider an additional model controlling for patient covariates, when estimating the moderating effects of survey timing.	Addition	Figure A7
7	I model selection into survey, as well as the correlates of missingness in outcome data.	Addition	Tables A3-A4

Table A9: Deviations and Additions to Pre-Analysis Plan

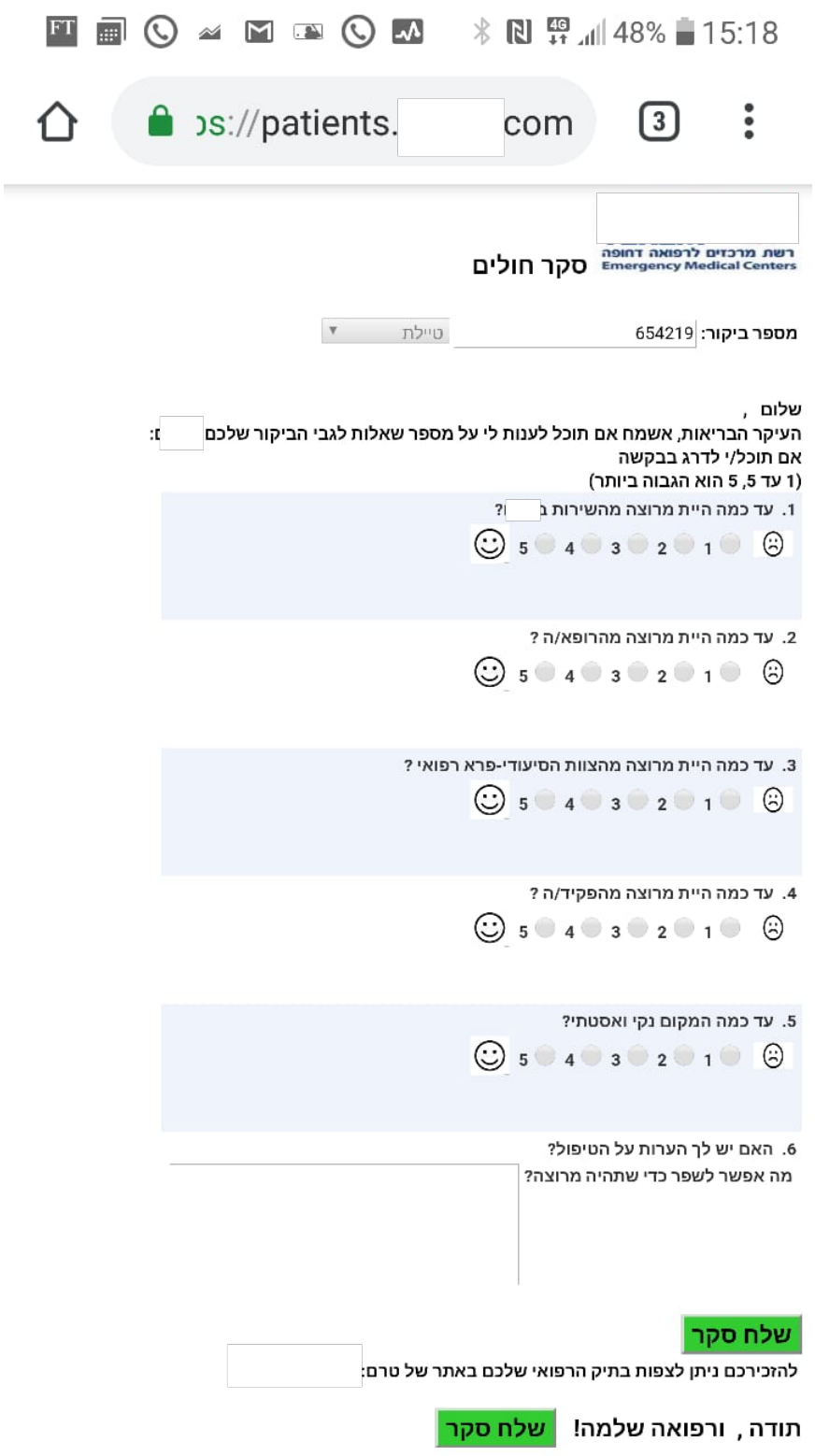
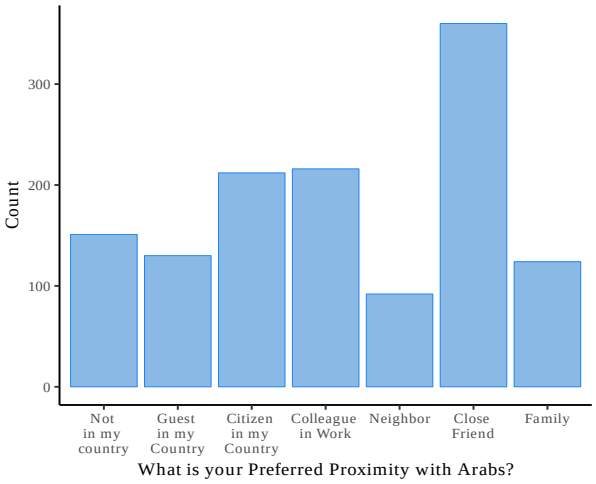
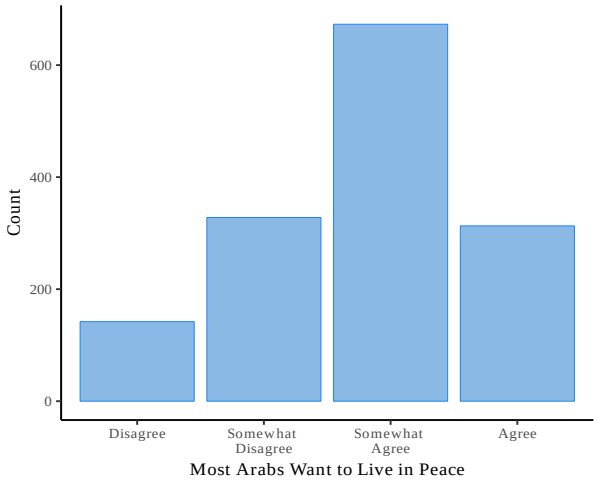


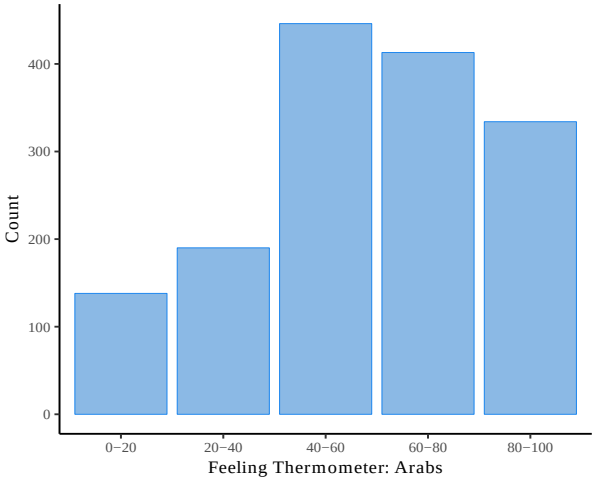
Figure A3: Survey Platform



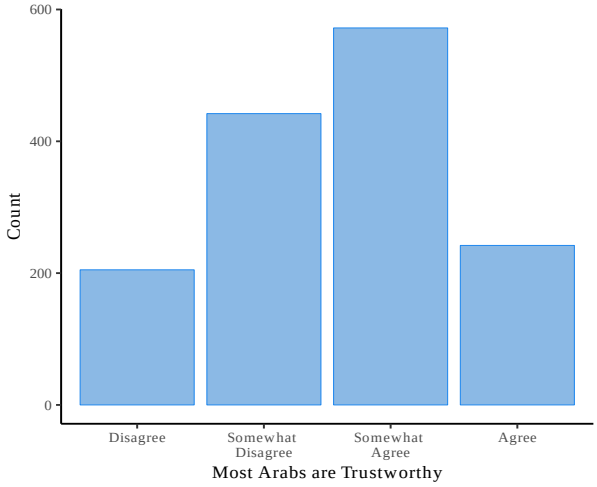
(a) Social Distance



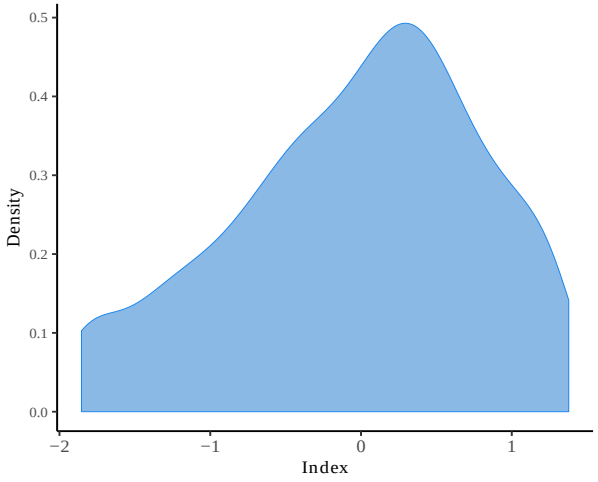
(b) Peace



(c) Thermometer



(d) Trust



(e) Index

Figure A4: Main Outcomes

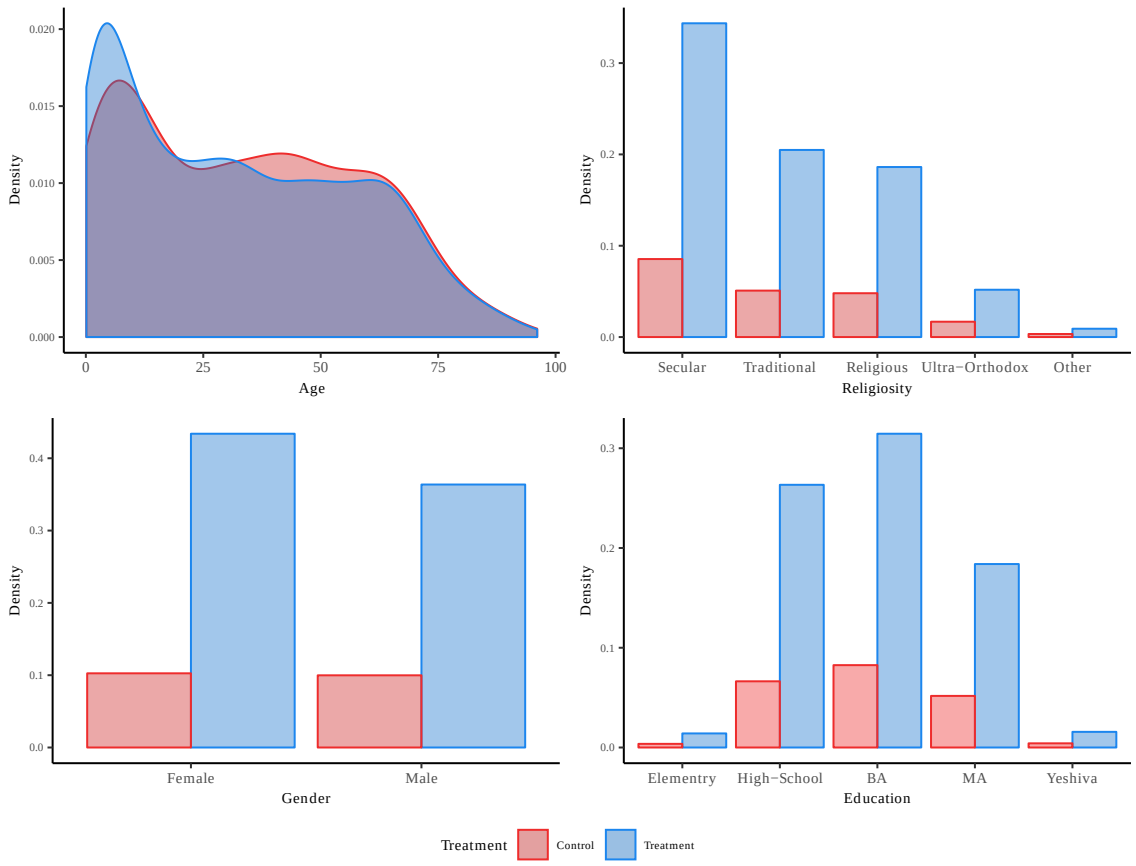
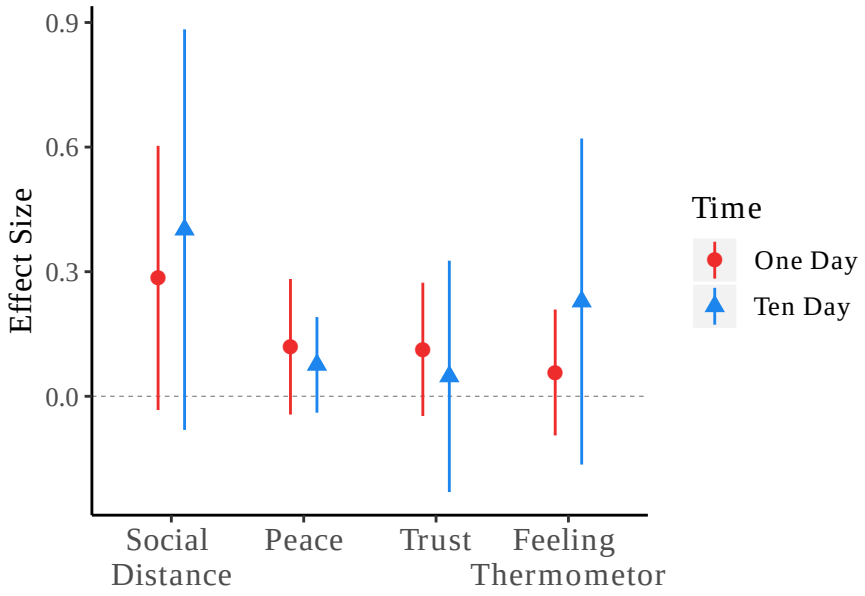
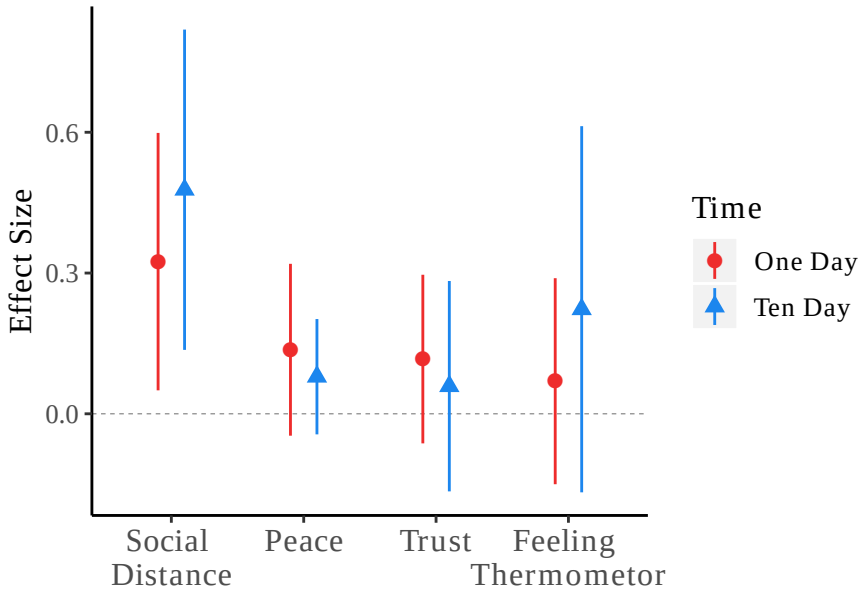


Figure A5: Demographics



*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



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

Figure A7: Decay Effects of Contact – Controlling for Pre-Treatment Covariates

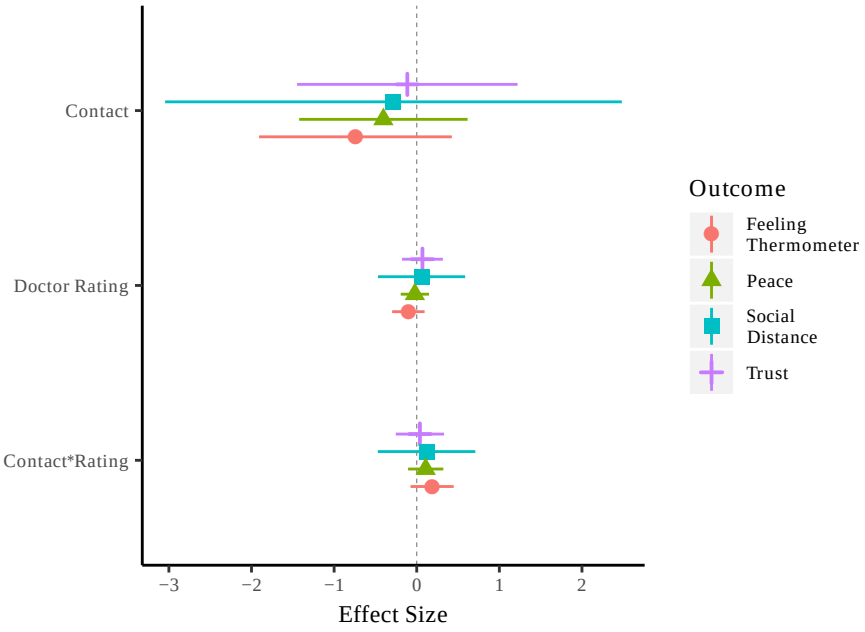


Figure A8: Heterogenous Effects of Contact Conditional on Doctor Quality

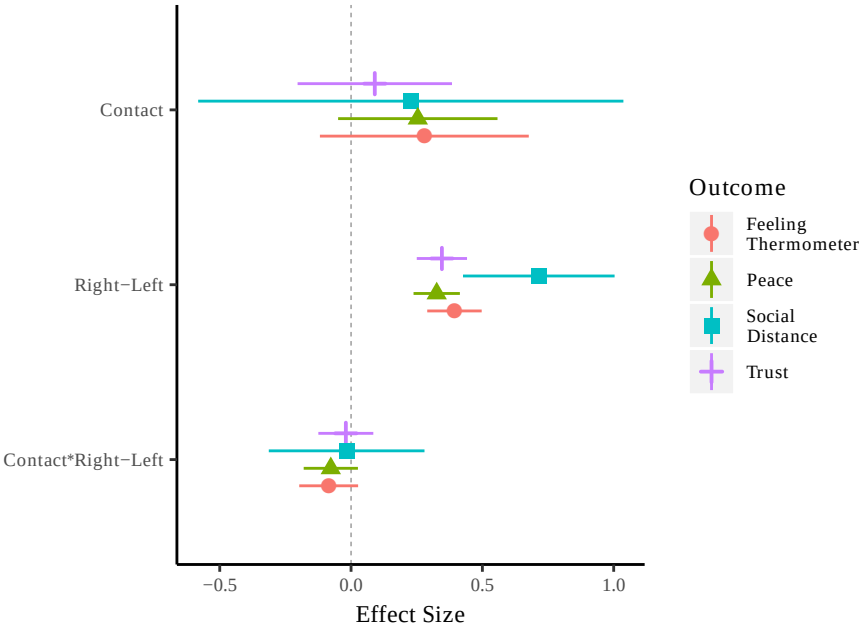


Figure A9: Heterogenous Effects of Contact Conditional on Patient Ideology

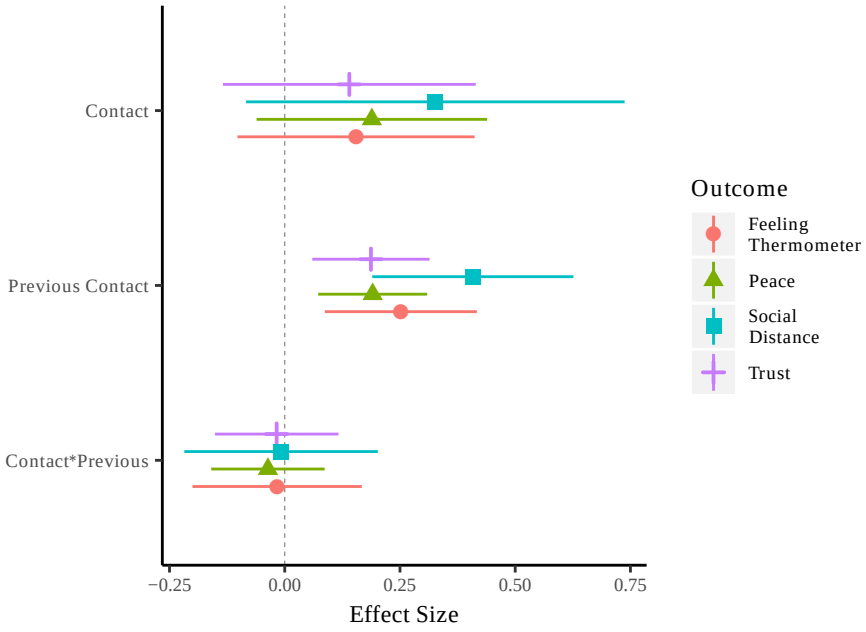


Figure A10: Heterogenous Effects of Contact Conditional on Previous Contact

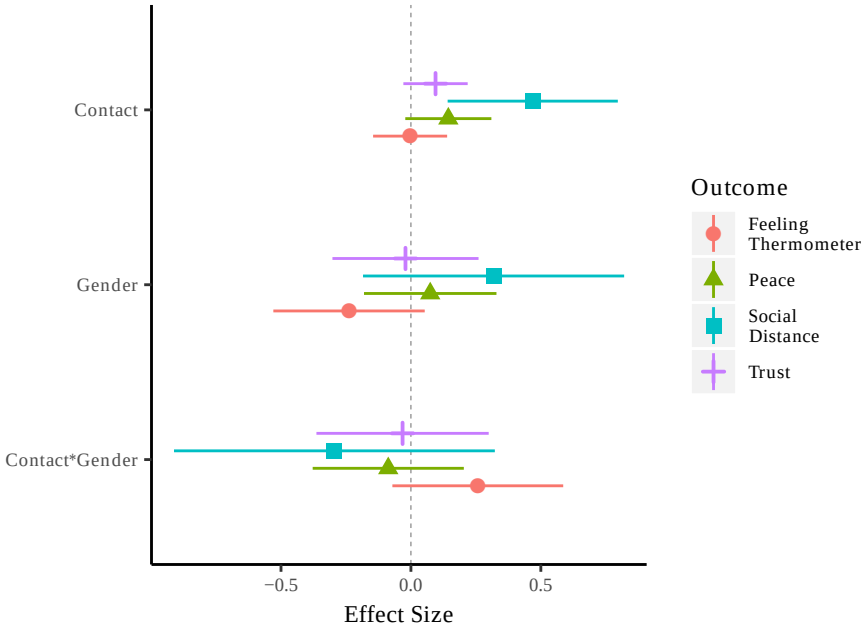


Figure A11: Heterogenous Effects of Contact Conditional on Gender

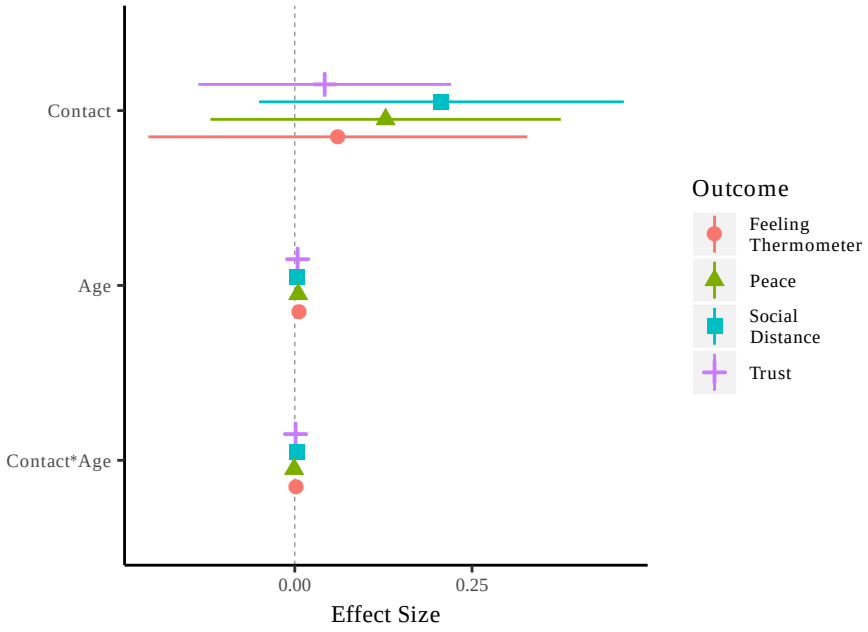
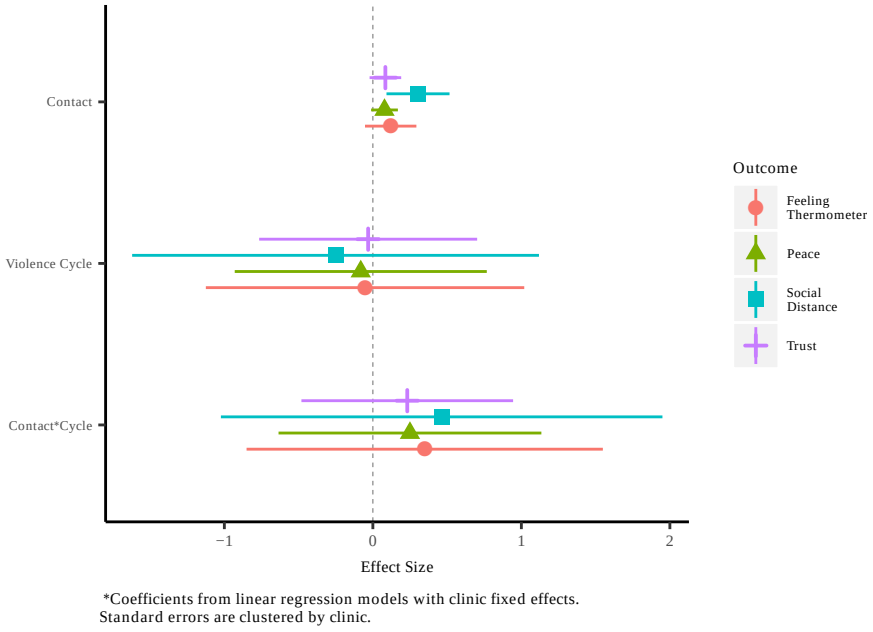


Figure A12: Heterogenous Effects of Contact Conditional on Age



*Coefficients from linear regression models with clinic fixed effects.
Standard errors are clustered by clinic.

Figure A13: Effects of Contact During Cycles of Violence