MINNESOTA POPULATION CENTER

UNIVERSITY OF MINNESOTA

Analyzing the Demographic, Spatial, and Temporal Factors Influencing Social Contact Patterns in the U.S. and Implications for Infectious Disease Spread

Audrey Dorèlien†
University of Minnesota

Aparna Ramen
University of Minnesota

Isabella Swanson
University of Minnesota

August 2020

Working Paper No. 2020-05
DOI: https://doi.org/10.18128/MPC2020-05

†Address correspondence to Audrey Dorèlien, University of Minnesota, Twin Cities, Minnesota Population Center, 50 Willey Hall, 225 19th Ave S., Minneapolis, MN 55455 (email: dorelien@umn.edu). We thank Rachelle Hill for her help throughout the development of this paper. We would also like to thank Sarah Petterson and Ayesha Mahmud for their comments to early drafts of this paper. We thank Jennifer Hook for sharing the O*NET/Census and 2010 Standard Occupational Classification crosswalk code with us. Dorélien, Ramen, and Swanson were supported by the Minnesota Population Center P2C HD041023 grant. This research was also funded by the University of Minnesota's Grant-in-Aid of Research, Artistry and Scholarship. Finally, we would like to acknowledge the grant that funds the American Time Use Survey: Data Access System R01HD053654.

Analyzing the demographic, spatial, and temporal factors influencing social contact patterns in the U.S. and implications for infectious disease spread

Authors:
Audrey M Dorélien, Aparna Ramen, Isabella Swanson
University of Minnesota

(This paper has not yet been peer reviewed)

Abstract

Background: We know diseases such as the 2019 Novel Coronavirus (COVID-19) are spread through social contact. Moreover, interventions to control social contacts such as stay-home orders are required to stop disease spread in pandemics for which vaccines have not yet been developed. However, existing data on social contact patterns in the United States (U.S.) is limited.

Method: Consequently, we use American Time Use Survey data from 2003-2018 to describe and quantify the number and duration of social contacts occurring at home and in non-household locations. For household locations we also estimate age contact matrices (who spends time with whom by age). This is the first study to describe variation in U.S. social contact patterns across space, time, and based on demographic characteristics.

Findings: We find that gender differences in social contact patterns exist. In the home, they appear to be driven by caretaking responsibilities. Non-Hispanic Blacks have a shorter duration and fewer social contacts than non-Hispanic Whites. However, they are more likely to work in jobs that require close physical proximity, therefore the nature of their contacts is riskier. Hispanics have the highest number of household contacts and are also more likely to work in jobs requiring close physical proximity compared to non-Hispanic whites. Seasonal differences in contact patterns are not large; they appear to be driven by the school term and therefore are chiefly present in school-aged respondents. Consequently, if the main mechanism driving infectious disease seasonality is seasonality in contact patterns, then we should not expect to see large seasonal differences in disease incidence when the young are not very susceptible or infectious. Spatial differences in contact patterns are small.

Conclusion: In addition to age**,** demographic characteristics, particularly race and ethnicity and gender, are associated with differences in social contact patterns. In contrast, seasonal differences seem to be associated with school participation and therefore less consequential for older respondents.

1 Introduction

Emerging infectious diseases, such as SARS-CoV-2, which causes the 2019 novel coronavirus disease (COVID-19), pose a substantial challenge to global and US public health. SARS-CoV-2, a respiratory pathogen, spreads primarily through in-person social contacts, and time spent in locations such as schools and the workplace greatly influence the number and durations of these contacts. This challenge of effectively responding to respiratory pathogens is greater when vaccines are not yet available. Therefore, in the period between the onset of a pandemic and the creation of a vaccine, stopping the spread of infectious disease becomes a question of promoting non-pharmaceutical interventions (NPI). NPIs such as school closures and physical distancing measures require a majority of individuals (those not classified as essential workers) to stay at home except for taking essential trips to get food or medicine. Using such interventions correctly requires a better understanding of social contact patterns, which are a critical factor in the transmission and control of infectious diseases such as coronavirus and influenza.

Social contact patterns vary by population, so it is important to get context-specific estimates to tailor interventions to the country or region of interest. Unfortunately, there is a paucity of empirical data on social contacts from the US. As a result, current US interventions are often difficult to target, leading to suboptimal outcomes. Moreover, when capacity for testing is limited early in an epidemic, as with COVID-19, it can be targeted towards regions, populations, and settings most likely to have high community spread when social contact patterns are known, facilitating identification of both mild symptomatic and asymptomatic carriers and preventing transmission by super-spreaders. When vaccines become available, the lack of accurate information impacts optimal vaccine distribution as well as vaccination booster schedules. In order to fill this information gap, we describe and quantify US social contact patterns using data from the Bureau of Labor Statistics' American Time Use Survey (ATUS), 2003-2018.

1.1 There are four different approaches for estimating age contact patterns

Social contact patterns can be summarized with age-contact matrices (Wallinga, Teunis, and Kretzschmar 2006). These matrices can capture many of the important components of social contact structures such as typical household composition, daily routines, and activities (e.g. school, work), and can be used as inputs into infectious disease models.

In particular, four main approaches to measure contacts directly from social data have been proposed (Wallinga, Teunis, and Kretzschmar 2006; Zagheni et al. 2008; Iozzi et al. 2010). The first and most common approach relies on contact surveys in which the respondent self - reports the number of contacts he or she had during a randomly sampled day (Wallinga, Teunis, and Kretzschmar 2006; Beutels et al. 2006; Edmunds, O'callaghan, and Nokes 1997; Mossong et al. 2008). Additional information captured in the survey includes age/sex of contacted persons, type of contact, duration, location, and frequency of contacts. The best-known study to use this approach is the Mossong et al. (2008) study, which collected contact information from 7,290 participants in 2006 for eight different European countries as part of the POLYMOD (Improving Public Health Policy in Europe through Modelling and Economic Evaluation of Interventions for the Control of Infectious Diseases) project funded by the European Commission. They recorded contacts over a 24-hour period using paper diaries in which information on the demographics of contacted persons, the location, frequency, duration, and type of contact (physical or non-physical) were collected. They found that age-specific social contact patterns do vary by country and that the differences are epidemiologically meaningful. Over the past ten years there have been additional POLYMOD-like studies conducted in Vietnam, Zimbabwe, Russia, and a few other developed countries (Horby et al. 2011; Melegaro et al. 2017; Ajelli and Litvinova 2017; Hoang et al. 2019). No such equivalent study exists for the US as a whole. Thus many researchers have used the POLYMOD data from UK and Germany, which are already more than 10 years old, as a substitute (Ewing et al. 2017; Medlock and Galvani 2009).

In a second approach, contact matrices are estimated from the simulation outputs of individual-based models, appropriately calibrated to socio-demographic and time use data, to generate the underlying contact network structure of the population (Del Valle et al. 2007; Iozzi et al. 2010).

The most recently developed and third approach creates a model that simulates individual-level contact based on POLYMOD data but uses different inputs from surveys such as the Demographic and Health Surveys (DHS) and International Labor Organization. This approach lacks individual-level contact data, but includes data on household age structure, population age composition, labor force participation and other factors that strongly influence individual- level contacts (Prem, Cook, and Jit 2017; Mistry et al. 2020).

The fourth approach relies on time use data and generates “time-of-exposure” age matrices (matrices of "who spends time with whom") by age. The age of the respondents’ contacts are generated by assuming that for single activity/locations and relatively small time intervals, people mix with each other proportionally to the relative presence of their age group in the location (Zagheni et al. 2008).

1.2 Existing information on US social contact patterns

As mentioned above, relatively little is known about US contact patterns. What is known has either been based on geographically small populations which may not be generalizable (DeStefano et al. 2011) or does not describe variation in social contacts across time and space (Zagheni et al. 2008). DeStefano et al. (2011) conducted a study of social contact patterns in four small North Carolina counties during the 2007-08 influenza season and found that the number of contacts varied with age and was lower on weekends than weekdays. They also found that for adults, the number of contacts increased during times of peak influenza activity but that this was not the case for children. There was also evidence of seasonal variation in mean daily contacts, but since this data was limited to one year they could not be certain if the pattern repeats every year (DeStefano et al. 2011).

Since there are no national US surveys of contact structures, Zagheni et al. (2008) used a single year of ATUS data to summarize one aspect of contact patterns—the duration of time people spend with other people of different ages. They found that people tend to spend more time with individuals of the same age and with individuals one generation apart, such as parents' interactions with children and vice versa. They illustrated that a model of age-specific immunity to varicella that incorporated the contact matrices from the time use survey was able to predict US varicella seroprevalence well. Zagheni et al.'s (2008) results have three important limitations that are addressed by our study. First, because they only use a single year of ATUS data, Zagheni et al. are unable to study how contact patterns may have changed over time. Second, they do not examine seasonal variations in contact patterns, which are well-known to be important drivers of the spread of close-contact diseases. Third, they do not examine spatial variation in contact patterns across the US. We are working with and building on this earlier work by taking advantage of the multiple years of data now available in the ATUS to identify meaningful sources of variation in contact patterns over time and across seasons and space.

1.3 Paper contributions

Time use diaries such as the ATUS contain some of the same information that is present in social contact surveys. Specifically, the number and duration of contacts in the respondents' households as well as the age and sex of the respondents' household members. Therefore, in this paper, we use the social contact survey approach to generate empirically based estimates of the age pattern of **duration** (in minutes), **mean number** of contacts, as well as **age contact matrices** (who spends time with whom by age) for household contacts in the US. We analyze household/home contacts in two ways—with household members only and with anyone present in the respondent's home or yard. We also estimate the age pattern of duration of social contacts for other locations, but we are not able to generate age contact matrices directly from the data. In a follow-up paper, we will use proportionate time mixing assumptions to create age specific contact matrices (showing duration of contacts) for non-household locations.

Because we have a large dataset that spans the years 2003-2018, **this is one of the first studies to describe these social contact patterns for different spatial (e.g., metro vs non-metro, regions) and temporal scales (e.g., day of the week, seasons, year, holidays), while**

incorporating socio-demographic characteristics of the respondents. We hope that this information can be used to parameterize models for the spread of close-contact infectious diseases, such as coronaviruses, influenza, and measles, while helping identify groups and settings to target for testing and interventions. The information generated in this paper can also be used to evaluate the effectiveness of physical distancing measures as a result of the COVID-19 pandemic (Jarvis et al. 2020). Future work will identify the main social and demographic determinants of US contact patterns.

2 Data and Methods

2.1 ATUS Data

The Bureau of Labor Statistics' American Time Use Survey (ATUS) has been fielded continuously from 2003 to present and focuses on time use in the United States. The goal of the survey is to measure how people divide their time among the various activities of daily life. Survey participants are asked to recall all of the activities that took place in the 24 hours on the day preceding the survey. Data are collected in the form of diaries in which respondents describe their daily activities chronologically, in increments as small as one minute. Diaries cover all seasons, days of the week, and holidays. ATUS data are publicly available and respondents are representative of all residents living in households in the United States that are at least 15 years old. ATUS does not include active military personnel and people residing in institutions such as nursing homes and prisons. ATUS respondents are randomly selected from households that recently completed their participation in the Current Population Survey. The ATUS has information on adults spending time with children but lacks information on the time children under the age of 15 engage in other activities or with other children.

Our sample consists of data from 200,136 individuals from 2003 to 2018 using IPUMS Time Use (Hofferth, Flood, and Sobek 2018). The sample consists of 112,286 females (56%) and 87,850 males (44%) who reside in the contiguous 48 US states. For every individual, we have information on his or her age, sex, race, marital status, state of residence, education level, labor force status, and occupation (see Table 1 for more details). The time diaries include information on activities and the location of activities, as well as information on whom the respondent is conducting the activity with. Information detailing the location and presence of others in the room was collected for most activities with the exception of sleeping, grooming, and other personal activities. Respondents exclusively reporting activities that did not include information on who was present or the location of the activity or refused to respond were also excluded from our final dataset (see Figure 1). There is no information on the age of individuals the respondents have contact with when they are not household members.

[Figure 1 about here]

Table 1. Counts illustrating some of the sociodemographic, spatial, and temporal variation available in the ATUS data set.

Total Sample N= 200,136 (excluding AK & HI)

Sex		**Presence of kids in HH**		**Race/Ethnicity**	
Male	87,850	No	109,616	Non-Hispanic white	136,252
Female	112,286	Yes	90,520	Hispanic	27,273
				Non-Hispanic black	26,775
				Non-Hispanic other	9,836
Economy		**Weekday vs. weekend**		**Employment status**	
Recession (12/01/2007 - 06/30/2009)	20,464	Weekday	99,711	Employed	124,356
Non recession	179,672	Weekend	100,425	Not in the labor force	66,503
				Unemployed	9,277
Education status		**Seasonal**		**Age**	
Less than high school	30,801	Winter	51,874	15 - 19	12,173
High school degree	52,038	Spring	50,621	20 - 24	8,734
Some college	53,695	Summer	49,276	25 - 29	13,748
Bachelors degree	39,692	Fall	48,365	30 - 34	18,459
Advanced degree	23,910			35 - 39	20,225
				40 - 44	20,488
				45 - 49	18,912
Climatic region				50 - 54	17,486
Northeast (CT, DE, ME, MD, MA, NH, NJ, NY, PA, RI, VT)			33,785	55 - 59	16,357
Southeast (AL, FL, GA, NC, SC, VA, DC)			22,708	60 - 64	14,410
Central (IL, IN, KY, MO, OH, TN, WV)			39,692	65 - 69	12,521
South (AR, KS, LA, MS, OK, TX)			8,771	70 - 74	9,522
North Central (IA, MI, MN, WI, MT, NE, ND, SD, WY)			26,768	75+	17,101
West (CA, NV)			35,721		
Southwest (AZ, CO, NM, UT)			11,035		
Northwest (ID, OR, WA)			21,656		

2.2 Methods

2.2.1 Defining social contacts in the home

We are interested in identifying social contacts that can influence the transmission of respiratory pathogens. We establish rules for defining social contacts first by location, then activity type (work versus non work), and then based on information on “with whom” the activity was done.

In the case of social contacts in the home, we consider any activity done in the respondent's home or yard with someone else a social contact. We exclude activities for which location and "who was present" were not collected from the respondents. Similarly, we drop those activities where questions about who was present were asked but the respondent refused to answer. Sleeping and other personal activities in the home are not included because they do not have information on "with whom."

2.2.2 Analyzing contacts in the home

The ATUS oversampled participants based on state, day of the week, race/ethnicity, age of children, presence of children, and number of adults in adults-only households. Therefore all of our analyses include sampling weights to account for this oversampling, as well as differential response rates, and the complex survey sampling design. Information about the complex survey sample design can be found in the ("American Time Use Survey User's Guide - Understanding ATUS 2003 to 2018" 2019). Controls are not used because the purpose of this analysis is descriptive, and only sample weights are necessary to make the analysis representative of the target population—the civilian, non-institutionalized US population within the contiguous 48 states (Solon, Haider, and Wooldridge 2015).

We group the respondents into the following five-year age groups: from 15-19 to 70-74, and all respondents 75 and over into a single group. For each respondent's age group we find d_j the **mean duration** (in minutes) and n_j the **mean number** of household members and all contacts in the home/yard.

We have age information for respondents' household members and use that to create two different **age contact matrices (m_{ij}, who spends time with whom by age)** showing the mean number (n_{ji}) and mean duration (d_{ji}) of contacts for respondents in age group *j* and household member contacts in age group *i*. We use 10-year age groups for the age contact matrices. Since we have age information for the respondents' household members, we also group the respondent's household members into the following age groups (0-5, 6-14, 15-24,…,60-74, 75+). The current version of the analysis does not correct for reciprocity (the fact that total contacts between respondents and contacts should be reciprocal) (Melegaro et al. 2017).

2.3 Social contacts outside the home (all other contacts)

We also estimate the age pattern of duration of social contacts for other locations (work, school, public locations, etc.), but we are not able to generate age contact matrices directly from the data.

2.3.1 Defining social contacts outside the home

We assume all activities involved social contact for the following locations: someone else's home; restaurant or bar; place of worship; grocery store; other store, mall; school; library; bank; gym/health club; post office; bus; subway, train; taxi, limousine service; and airplane. Most of these are public locations, so we assume that someone else would always be present (e.g. someone else, even if only staff, would always be present at a restaurant or bar). For the rest of the locations (except for work activities in the respondent's workplace pre-2010 and personal activities) we assume a social contact is present only if another person is recorded as present (under the variable relatew) during the activity. We exclude activities for which location and who was present were not collected from the respondents, except for personal activities described as kissing, cuddling, etc. In those instances, we assume social contact is always present. Similarly, we drop those activities where questions about who was present was asked but the respondent refused to answer. See Figure 1 for more details on exclusions from the analytic sample.

Prior to 2010, work-related social contacts (e.g., boss or manager; people whom the respondent supervises; co-workers; customers) were not recorded in the ATUS. To address this omission, we calculate the percent of time spent in social contact by detailed occupation (occ) categories for post-2010 work activities in the workplace (see appendix A for more information). We use that calculation to impute the percent of time spent in social contact for pre-2010 work activities in the workplace. We multiply the duration of work activities pre-2010 by the percent of work time spent in social contact for that person's occupation category post-2010 to get an adjusted duration of time spent in social contact at work. For those detailed occupation codes that did not match between pre- and post- 2010 work activities, we either substitute post-2010 occupation codes - by using a cross-walk where available - or use the percentages from the broader occupation categories (occ2). Work activities in locations other than the workplace, post-2010 work activities in the workplace, and all non-work activities in the workplace are not included in this imputation, and they are treated the same as other activities in those locations (e.g. a work activity in a restaurant or bar would be treated as always involving social contact). Finally, we sum the duration of activities with social contact for each individual using the rules described above and merge this dataset with individual and household characteristics for our analyses.

3 Results and Discussion

In this section we present the results for socio-demographic (gender and race/ethnicity), temporal (type of day and seasons), and spatial (climatic regions and metro/non-metro) variation in social contact patterns. For each of these categories, we describe the patterns of social contacts taking place in the home, then we describe patterns occurring across all locations. The figures in this section describe the means; in the supplementary materials we include corresponding tables with 95 percent confidence intervals.

3.1 General Age Patterns and Gender Differences

Contacts in the home

The mean number of social contacts in the home dips in the twenties and peaks in the thirties (Figure 2). The mean duration of contacts also peaks in the thirties, but there is no corresponding dip in the twenties. This aligns with the average age of first marriage in the US which is 28. People in their twenties are more likely to be single and thus live alone or have few contacts with roommates. As they enter their thirties and grow their families, they tend to have a greater number of household member contacts and spend more time with children (Dukhovnov and Zagheni 2015). Interestingly, while elderly males (75+) have the fewest average number of household contacts, they on average spend the second highest duration of time with other household members, second only to women in their thirties. While elderly women have a similar peak duration at ages 65-74, unlike elderly men, the time they spend with others declines again after age 75. This decline is likely due to the fact that a higher proportion of elderly women are widowed. Women have higher durations of household social contact before age 60, but after that age, men spend increasingly more time with others at home (Glauber 2017).

The age patterns shown here are very similar to recently published age patterns for the UK (Klepac et al. 2020). Specifically, both the data from the US and the UK document a peak in the mean number of contacts for young and middle aged adults (35-39 for the ATUS data versus 40-44 for UK adults). This discrepancy in the timing of the middle age adult peak may reflect the fact that our figures are based on data from the past 15 years. We may find similar results if we restrict our sample to more recent years. On average, the mean number of household contacts in our ATUS-based estimates are smaller than the UK based estimates. We will need to conduct further analyses to see if this finding can be explained by differences in household composition or other factors.

[Figure 2 about here]

Results for all locations

The age pattern of duration of social contacts across all locations is similar to the age pattern of the duration of household contacts in that women younger than 45 had a higher duration of contact than men, while men over 60 had a higher duration of contact than women. The peak in mean duration of social contact occurs in the 30-34 age group. Presumably also due to increased childrearing and childcare responsibilities, but it is much less steep than the peak of household contact duration. Duration of social contact is closely aligned for men and women ages 45-60 and steadily decreases through middle age. This makes sense as childrearing responsibilities are reduced or absent in middle age and men and women would have similar work and household contact patterns. After age 60 duration of social contacts continues to decline for both men and women (due to declining labor force participation) but declines at a faster rate for women. Consequently, elderly men have longer duration of social contacts compared to women likely due the higher proportion of elderly women who are widowed and do not remarry.

[Figure 3 about here]

If the duration of social contacts is strongly correlated with the number of contacts in the US as in other settings(Melegaro et al. 2017), then although individuals below the age of 50 are at lower risk of COVID-19 than the elderly, they may be responsible for the majority of the spread through interpersonal contacts.

3.2 Racial/Ethnic Differences

Contacts in the home

There are large differences in age patterns of social contacts across different racial/ethnic groups in the United States. At home, Non-Hispanic Blacks have the fewest number and shortest duration of contacts, while Hispanics on average have the highest number and duration of contacts particularly at younger ages (Figure 4). Non-Hispanic others and Non-Hispanic whites have similar numbers of contacts below age 35, but at older ages the mean number of contacts for Non-Hispanic others was similar to those of Hispanics. Non-Hispanic white and Non-Hispanic others have a shorter duration of social contacts than Hispanics at younger ages but tracks the time use pattern of Hispanics closely starting at age 40.

[Figure 4 about here]

Results for all locations

We find similar patterns when we sum the mean duration of social contacts across all locations. Non-Hispanic Blacks report substantially shorter durations of social contact than other racial/ethnic groups, and the gap (which ranges from about 50-100 minutes) between Non-Hispanic Blacks and Hispanics/Non-Hispanic whites persists at all age groups. Hispanics and Non-Hispanic whites have nearly identical durations of social contact at all ages. Non-Hispanic others have significantly lower durations of social contact than Hispanics and Non-Hispanic whites below age 30 and age 50-54, but higher durations than them after age 64.

[Figure 5 about here]

Our preliminary findings indicate that Non-Hispanic Blacks have fewer contacts and shorter duration of contacts compared to other groups. Though these results are based on data from a period of time before the COVID-19 pandemic and stay-at-home orders were issued, these results were unexpected because Non-Hispanic Blacks appear to have a higher risk of contracting and dying from COVID-19 (Chowkwanyun and Reed 2020; Gross et al. 2020). One might expect that this disparity could be partially explained by higher number of household contacts or higher duration of social contacts. However, we find that the number and duration of social contacts are not likely responsible for COVID-19 black-white racial disparities.

Results for O*NET occupational analysis

We merge the ATUS data with data from Occupational Information Network (O*NET), to see if there are racial differences in the types of social contacts at work by occupation. The O*NET data contains a “physical proximity” variable that is an average of sampled workers’ responses within an occupation regarding the level of physical proximity experienced in the workplace. The responses are based on a five-point Likert scale, ranging from 1 - “I don’t work near other people (beyond 100 ft)” to 5 – “Very close (near touching)”. The additional analysis reveals that Non-Hispanic Blacks are more likely to work in occupations with higher levels of physical proximity which increases disease risk (Figure 6) (see appendix B for more details on this analysis).

We divide the physical proximity responses into four categories (with equal ranges) including high (e.g. dental hygenists), mid to high (e.g. cooks), mid to low (e.g engineers), and low levels of physical proximity (e.g. loggers). We then compare the composition of these groups by racial and and ethnic categories in the ATUS. Figure 6A shows that Non-Hispanic Blacks have the highest proportion (38%) of occupations with high levels of physical proximity and the lowest proportion of occupations with mid to low levels of physical proximity. The racial differences are even more prominent in panel B, where the bars displaying the racial composition of physical proximity categories show that Non-Hispanic Whites make up a decreasing share and Non-Hispanic Blacks make up a correspondingly increasing share of workers in occupations with the highest levels of physical proximity. The proportion of Hispanics also increases across the top three physical proximity categories, while the proportion of Non-Hispanic Others stays essentially the same. The lowest physical proximity category contains a very small number of observations, and only consists of Non-Hispanic Whites and Hispanics.

The mean O*NET physical proximity score is highest for Non-Hispanic Blacks (65) followed by Hispanics (64) which are statistically significantly different than Non-Hispanic whites (60) ($p<0.0001$). So Non-Hispanic Blacks may on average have higher risk/intensity contacts due to the physical proximity levels associated with their occupations, even though they may have an overall lower number and duration of contacts. This may help explain the COVID-19 racial disparities in infection and mortality rates particularly if these workers are more likely to be deemed essential.

[Figure 6 about here]

3.3 Weekday versus Weekend Differences

Contacts in the home

When restricting our analysis to household member contacts, we did not find that the number of weekend contacts were meaningfully higher than weekday contacts (Figure 7). But the number of all contacts occurring in the respondents' home/yard is higher for weekends versus weekdays (a difference of 0.11 contacts). This difference is less than in recent findings for the UK (Klepac et al. 2020). We did find that the duration of contacts with household members were longer during the weekend (see the bottom panels of Figure 7).

[Figure 7 about here]

Results for all locations

We expect that if we had data on number of contacts occurring outside the home, we would find that the mean number of contacts during the weekday exceeds that of the weekend (DeStefano et al. 2011; Klepac et al. 2020; Béraud et al. 2015). Unlike the age pattern of duration of household contacts, which is always higher during the weekend compared to the weekday, the duration of social contact in all locations is higher during the weekday for those under age 25 and is the highest for those 15-19. This is likely explained by the large amount of time spent at school and college at those ages. From ages 25-59, the duration of social contact is nearly identical when comparing weekend days and weekdays. After age 60, the duration of social contact is greater during the weekend, which corresponds to the reduction in work hours and associated social contacts after retirement age.

[Figure 8 about here]

3.4 Seasonal Differences

Contacts in the home

Seasonal differences in mean number of household contacts do not appear to be meaningful; however mean duration of household contacts are on average a little longer during the winter months (Figure 9). These findings are similar to a study that did not find that seasons had meaningful impact on the number of contacts or the mixing patterns in France (Béraud et al. 2015); but contrast with the DeStefano et al. (2011) paper that found evidence of seasonal variation in mean daily contacts.

[Figure 9 about here]

Results for all locations

When summing contacts across all locations, we find no meaningful seasonal differences in the duration of contacts except for respondents under age 25, who have shorter durations of contacts in the summer compared to the other seasons. This is likely explained by seasonality of the school term. Young people who are of school or college age have higher rates of social contact when school is in session compared to the summer (supplemental analysis available from authors upon request).

[Figure 10 about here]

These results imply that if the main mechanism driving infectious disease seasonality are seasonality in contact patterns, then we should not expect to see large seasonal differences in disease incidence when the young are not very susceptible or infectious, as is the case for COVID-19 (Zhang et al. 2020). Disease seasonality would have to be driven by other factors such as seasonality of pathogen survival outside the host or seasonal changes in host immunity (Grassly and Fraser 2006)

3.5 Spatial Differences in Contact Patterns

We did not find large differences in contact patterns when comparing metro with non-metro areas or when comparing across the nine climatically consistent regions (Karl and Koss 1984) (see figures in appendix C for more details).

3.6 Age contact matrices

Recall that the ATUS only has data on the age of respondents' contacts if the contacts are household members. Therefore, we can only create age contact matrices between household members. The age contact matrix showing the mean number of household contacts (Figure 6) displays many of the same features found in the matrices from the POLYMOD and a recent UK survey (Mossong et al. 2008; Klepac et al. 2020). Our household contact matrix was not similar to the Melegaro et al (2017) matrix for the Manicaland Province of Zimbabwe; where household sizes are larger and extended families are more common. We document assortative contacts with age (siblings spending time with siblings, and similar aged couples or roommates spending time with each other); there is also evidence of people one generation apart spending time with each other (parents and their children). We will take a closer look at the data to identify which POLYMOD country the US data is most similar to in future work.

The age contact matrix showing the duration of contacts (duration of exposure matrix) is qualitatively similar to the Zagheni et al. (2008) household member duration of exposure matrix (Figure 11 and see Zagheni Figure 3). Both of these time-use based duration of contact matrices were similar to the matrices with mean number of contacts. The line graphs are consistent with the patterns in the age contact matrices, in which you can see that the total number of contacts is decreasing overall with respondent age, while the total duration is high for younger and older age groups but dips in the middle.

[Figure 11 about here]

In Figure 12, which breaks down the age contact matrices by sex, we can see that women ages 25-44 spend much more time with children and have higher numbers of contacts with them than men. Women spend more time on average on child-rearing in dual parent households (Bianchi, Robinson, and Milke 2006), and a higher proportion of women are single parents than men, and thus would have more contacts with children. Women have higher numbers of contacts and longer durations of contacts with those in the 5-year age groups directly above theirs, while men have the exact opposite pattern, with higher numbers and durations of contact in the 5-year age groups directly below theirs. This is likely explained by the age gap in married couples, where women’s partners skew older, with 45% falling between 2-9 years older. As in the line graphs, elderly men report higher numbers and durations of contact with their age groups (and those directly below theirs) because they are less likely to be widowed than elderly women.

[Figure 12 about here]

4 Conclusion

We find the largest variation in social contacts based on demographic factors such as age, gender, and race/ethnicity. We also find some temporal differences in social contact patterns; the largest differences are in the duration of social contacts as opposed to number of social contacts. Despite the large size of the U.S., we do not document meaningful differences in social contact patterns across regions or when comparing metro and non-metro areas.

There are several important limitations to our results. The ATUS data do not include respondents below age 15. We only know the age and sex of the respondents' contacts if they are household members and therefore are not able to create age-specific contact matrices for non-household locations without making some assumptions about proportionate time mixing, which may not always be good approximations (Zagheni et al. 2008). Also, the data do not specify the type of social contact (conversational versus physical). However, some of that may be inferred based on the description of the activity. Additionally, ATUS does not include active military personnel and people residing in institutions such as nursing homes and prisons, while the latter may be some of the most vulnerable to infectious disease spread and impact.

Nevertheless, the ATUS data do allow us to estimate social contacts before substantial physical distancing measures were implemented to control the COVID-19 pandemic. **For the US this may be one of the only sources of pre-pandemic social contact data.** This data can help us analyze the effectiveness of physical distancing measures by comparing the pre- pandemic social mixing patterns and matrices with changing contact patterns under different mitigation strategies. For instance, do we see changes in contacts occurring in the respondents' house/yard such as increased contacts with household members and neighborhood kids or decreased contacts with grandparents who are in higher risk age groups? We can also check whether the total duration of contacts in places like restaurants, workplaces, schools, public transportation, and grocery stores has declined as a result of physical distancing measures.

Another advantage of the ATUS data is that we are able to disaggregate the social mixing data by geographic region and respondents' sociodemographic characteristics. Though our analytic sample is strengthened by pooling multiple years together, we are careful to ensure that there are sufficient observations within each category in order for the results to be meaningful. This disaggregation is important because the US is large and heterogeneous, and physical distancing measures have not been uniformly enacted or embraced. Social determinants (e.g. socioeconomic status, metropolitan vs non-metropolitan areas, and occupation) can impact both baseline social contact patterns and the ability to physically distance. Moreover, disaggregation can identify who remains most at risk and where testing and interventions should be targeted to prevent spread.

5 Acknowledgements

We thank Rachelle Hill for her help throughout the development of this paper. We would also like to thank Sarah Petterson and Ayesha Mahmud for their comments to early drafts of this paper. We thank Jennifer Hook for sharing the O*NET / Census and 2010 Standard Occupational Classification crosswalk code with us. Dorélien, Ramen, and Swanson were supported by the Minnesota Population Center P2C HD041023 grant. This research was also funded by University of

Minnesota's Grant- in-Aid of Research, Artistry and Scholarship. Finally, we would like to acknowledge the grant that funded the American Time Use Survey: Data Access System R01HD053654.

References

Ajelli, Marco, and Maria Litvinova. 2017. "Estimating Contact Patterns Relevant to the Spread of Infectious Diseases in Russia." *Journal of Theoretical Biology* 419 (April): 1–7. https://doi.org/10.1016/j.jtbi.2017.01.041.

"American Time Use Survey User's Guide - Understanding ATUS 2003 to 2018." 2019, 118.

Béraud, Guillaume, Sabine Kazmercziak, Philippe Beutels, Daniel Levy-Bruhl, Xavier Lenne, Nathalie Mielcarek, Yazdan Yazdanpanah, Pierre-Yves Boëlle, Niel Hens, and Benoit Dervaux. 2015. "The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases." *PLoS ONE* 10 (7). https://doi.org/10.1371/journal.pone.0133203.

Beutels, P., Z. Shkedy, M. Aerts, and P. Van Damme. 2006. "Social Mixing Patterns for Transmission Models of Close Contact Infections: Exploring Self-Evaluation and Diary-Based Data Collection through a Web-Based Interface." *Epidemiology & Infection* 134 (6): 1158–66. https://doi.org/10.1017/S0950268806006418.

Bianchi, Suzanne M., John P. Robinson, and Melissa A. Milke. 2006. *The Changing Rhythms of American Family Life*. Russell Sage Foundation.

Chowkwanyun, Merlin, and Adolph Reed. 2020. "Racial Health Disparities and Covid-19 — Caution and Context." *New England Journal of Medicine*, May. https://www.nejm.org/doi/full/10.1056/NEJMp2012910.

Del Valle, S. Y., J. M. Hyman, H. W. Hethcote, and S. G. Eubank. 2007. "Mixing Patterns between Age Groups in Social Networks." *Social Networks* 29 (4): 539–54. https://doi.org/10.1016/j.socnet.2007.04.005.

DeStefano, F., M. Haber, D. Currivan, T. Farris, B. Burrus, B. Stone-Wiggins, A. McCALLA, et al. 2011. "Factors Associated with Social Contacts in Four Communities during the 2007–2008 Influenza Season." *Epidemiology & Infection* 139 (8): 1181–90. https://doi.org/10.1017/S095026881000230X.

Dukhovnov, Denys, and Emilio Zagheni. 2015. "Who Takes Care of Whom in the United States? Time Transfers by Age and Sex." *Population and Development Review* 41 (2): 183–206. https://doi.org/10.1111/j.1728-4457.2015.00044.x.

Edmunds, W. J., C. J. O'callaghan, and D. J. Nokes. 1997. "Who Mixes with Whom? A Method to Determine the Contact Patterns of Adults That May Lead to the Spread of Airborne Infections." *Proceedings of the Royal Society of London B: Biological Sciences* 264 (1384): 949–57. https://doi.org/10.1098/rspb.1997.0131.

Ewing, Anne, Elizabeth C. Lee, Cécile Viboud, and Shweta Bansal. 2017. "Contact, Travel, and Transmission: The Impact of Winter Holidays on Influenza Dynamics in the United States." *The Journal of Infectious Diseases* 215 (5): 732–39. https://doi.org/10.1093/infdis/jiw642.

Glauber, Rebecca. 2017. “Gender Differences in Spousal Care Across the Later Life Course.” *Research on Aging* 39 (8): 934–59. https://doi.org/10.1177/0164027516644503.

Grassly, NC, and C Fraser. 2006. “Seasonal Infectious Disease Epidemiology.” *Proceedings of the Royal Society B: Biological Sciences* 273 (1600): 2541.

Gross, Cary P., Utibe R. Essien, Saamir Pasha, Jacob R. Gross, Shi-yi Wang, and Marcella Nunez-Smith. 2020. “Racial and Ethnic Disparities in Population Level Covid-19 Mortality.” *MedRxiv*, May, 2020.05.07.20094250. https://doi.org/10.1101/2020.05.07.20094250.

Hoang, Thang, Pietro Coletti, Alessia Melegaro, Jacco Wallinga, Carlos G. Grijalva, John W. Edmunds, Philippe Beutels, and Niel Hens. 2019. “A Systematic Review of Social Contact Surveys to Inform Transmission Models of Close-Contact Infections:” *Epidemiology* 30 (5): 723–36. https://doi.org/10.1097/EDE.0000000000001047.

Hofferth, S. L., S. M. Flood, and M. Sobek. 2018. “American Time Use Survey Data Extract Builder: Version 2.7 [Dataset].” *College Park, MD: University of Maryland and Minneapolis, MN: University of Minnesota. Doi* 10: D060. https://doi.org/10.18128/D060.V2.7.

Horby, Peter, Pham Quang Thai, Niel Hens, Nguyen Thi Thu Yen, Le Quynh Mai, Dang Dinh Thoang, Nguyen Manh Linh, et al. 2011. “Social Contact Patterns in Vietnam and Implications for the Control of Infectious Diseases.” *PLoS ONE* 6 (2): e16965. https://doi.org/10.1371/journal.pone.0016965.

Iozzi, Fabrizio, Francesco Trusiano, Matteo Chinazzi, Francesco C. Billari, Emilio Zagheni, Stefano Merler, Marco Ajelli, Emanuele Del Fava, and Piero Manfredi. 2010. “Little Italy: An Agent-Based Approach to the Estimation of Contact Patterns- Fitting Predicted Matrices to Serological Data.” *PLOS Computational Biology* 6 (12): e1001021. https://doi.org/10.1371/journal.pcbi.1001021.

Jarvis, Christopher I, Kevin Van Zandvoort, Amy Gimma, Kiesha Prem, CMMID COVID-19 working group, Petra Klepac, G James Rubin, and W John Edmunds. 2020. “Quantifying the Impact of Physical Distance Measures on the Transmission of COVID-19 in the UK.” Preprint. Epidemiology. https://doi.org/10.1101/2020.03.31.20049023.

Karl, Thomas R., and Walter J. Koss. 1984. “Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983.” Historical Climatology Series 4-3. Asheville, NC: National Climatic Data Center.

Klepac, Petra, Adam J Kucharski, Andrew JK Conlan, Stephen Kissler, Maria Tang, Hannah Fry, and Julia R Gog. 2020. “Contacts in Context: Large-Scale Setting-Specific Social Mixing Matrices from the BBC Pandemic Project.” Preprint. Epidemiology. https://doi.org/10.1101/2020.02.16.20023754.

Medlock, Jan, and Alison P. Galvani. 2009. “Optimizing Influenza Vaccine Distribution.” *Science* 325 (5948): 1705–8. https://doi.org/10.1126/science.1175570.

Melegaro, Alessia, Emanuele Del Fava, Piero Poletti, Stefano Merler, Constance Nyamukapa, John Williams, Simon Gregson, and Piero Manfredi. 2017. “Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe.” *PLOS ONE* 12 (1): e0170459. https://doi.org/10.1371/journal.pone.0170459.

Mistry, Dina, Maria Litvinova, Ana Pastore y Piontti, Matteo Chinazzi, Laura Fumanelli, Marcelo F. C. Gomes, Syed A. Haque, et al. 2020. “Inferring High-Resolution Human Mixing Patterns for Disease Modeling.” *ArXiv:2003.01214 [Physics, q-Bio]*, February. http://arxiv.org/abs/2003.01214.

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, et al. 2008. “Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases.” *PLOS Medicine* 5 (3): e74. https://doi.org/10.1371/journal.pmed.0050074.

Prem, Kiesha, Alex R. Cook, and Mark Jit. 2017. “Projecting Social Contact Matrices in 152 Countries Using Contact Surveys and Demographic Data.” *PLOS Computational Biology* 13 (9): e1005697. https://doi.org/10.1371/journal.pcbi.1005697.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. “What Are We Weighting For?” *Journal of Human Resources* 50 (2): 301–16. https://doi.org/10.3368/jhr.50.2.301.

Wallinga, Jacco, Peter Teunis, and Mirjam Kretzschmar. 2006. “Using Data on Social Contacts to Estimate Age-Specific Transmission Parameters for Respiratory-Spread Infectious Agents.” *American Journal of Epidemiology* 164 (10): 936–44. https://doi.org/10.1093/aje/kwj317.

Zagheni, Emilio, Francesco C. Billari, Piero Manfredi, Alessia Melegaro, Joel Mossong, and W. John Edmunds. 2008. “Using Time-Use Data to Parameterize Models for the Spread of Close-Contact Infectious Diseases.” *American Journal of Epidemiology* 168 (9): 1082–90. https://doi.org/10.1093/aje/kwn220.

Zhang, Juanjuan, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, et al. 2020. “Changes in Contact Patterns Shape the Dynamics of the COVID-19 Outbreak in China.” *Science*, April. https://doi.org/10.1126/science.abb8001.

Figure 1. Flowchart illustrating our sample selection and how we define social contacts.

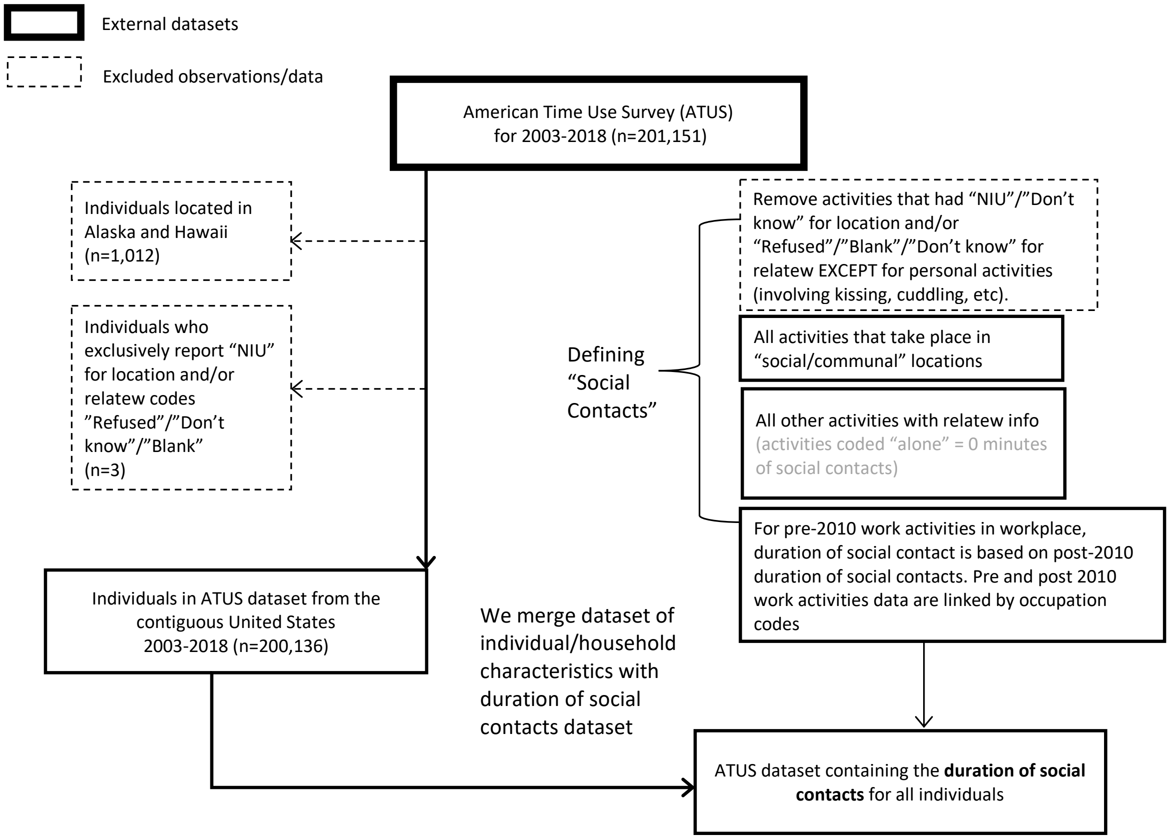
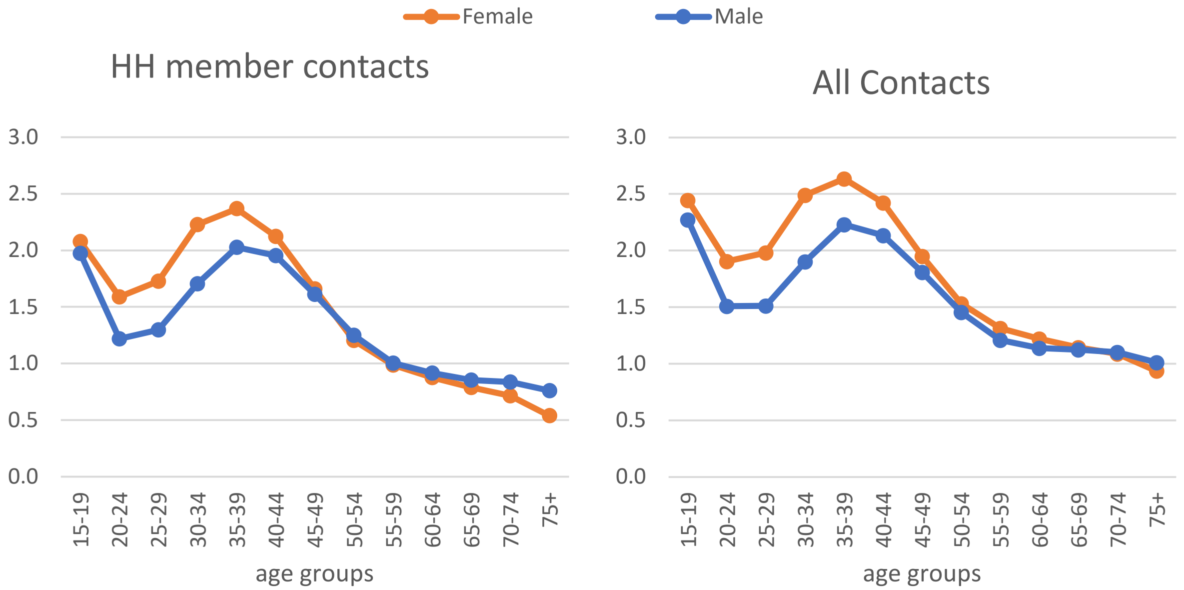


Figure 2. Sex and age pattern of mean number and duration of household social contacts. The number of household contacts is primarily composed of contacts between household members. The age patterns differ when we compare the mean number of contacts and the mean duration of contacts. For instance, while elderly males have the fewest average number of household contacts among all males, they on average spend the most time with other household members. Sleeping and other personal activities are not included.

Average **number** of contacts in home or yard by respondents age groups.



Average **duration** of contacts in home or yard by respondents age groups.

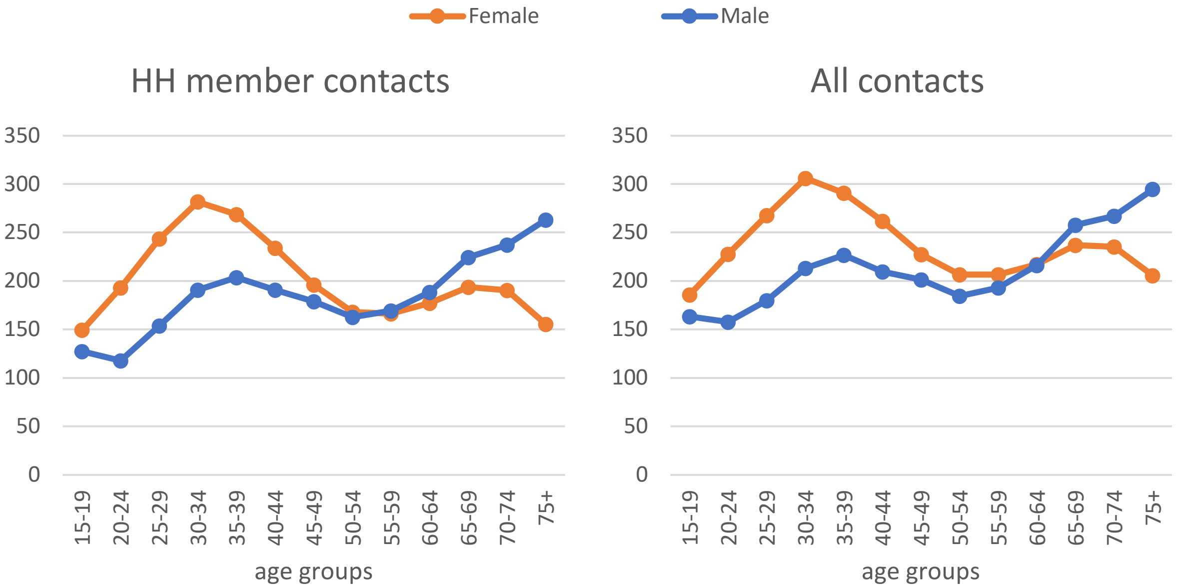


Figure 3. Sex and age pattern of duration of social contact for all locations. Younger women had a slightly higher duration of contact than men of the same age, while elderly men had a higher duration of contact than elderly women. The duration of contact declines for both men and women through middle and old age.

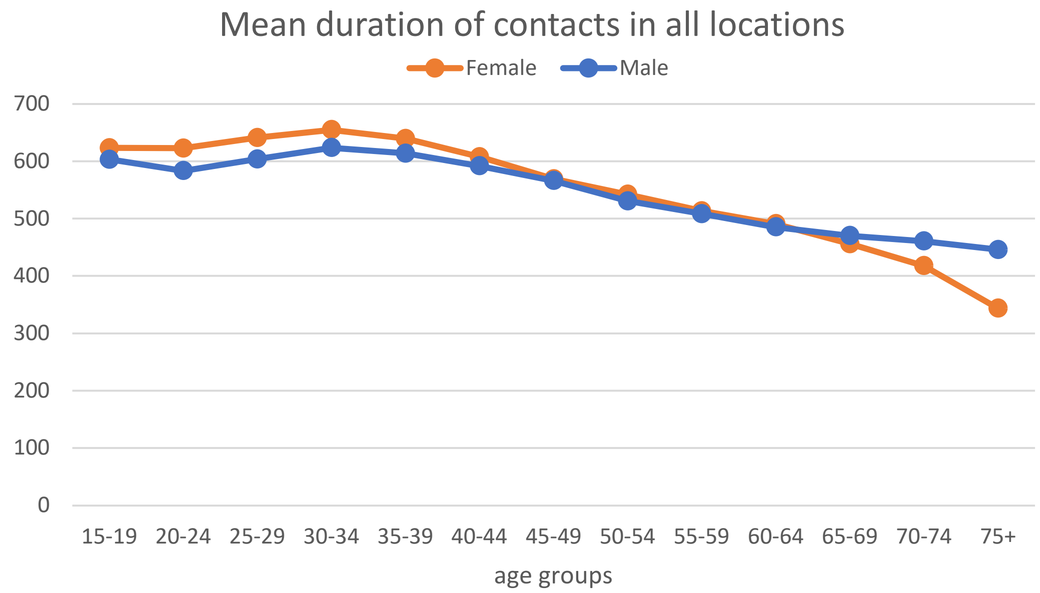
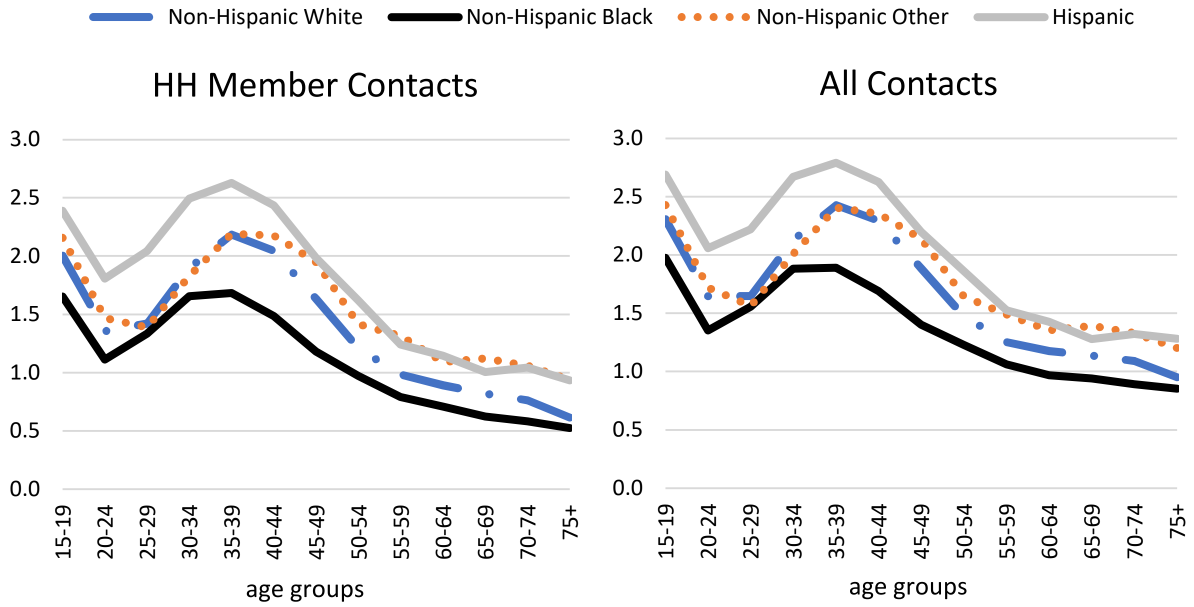


Figure 4. Differences in contact patterns by race/ethnicity. Non-Hispanic Blacks had the lowest number and shortest duration of household contacts, while Hispanics had the highest number and duration of household contacts. Non-Hispanic Others and Non-Hispanic Whites had similar patterns of contacts.

Average **number** of contacts in home or yard by respondents age groups.



Average **duration** of contacts in home or yard by respondents age groups.

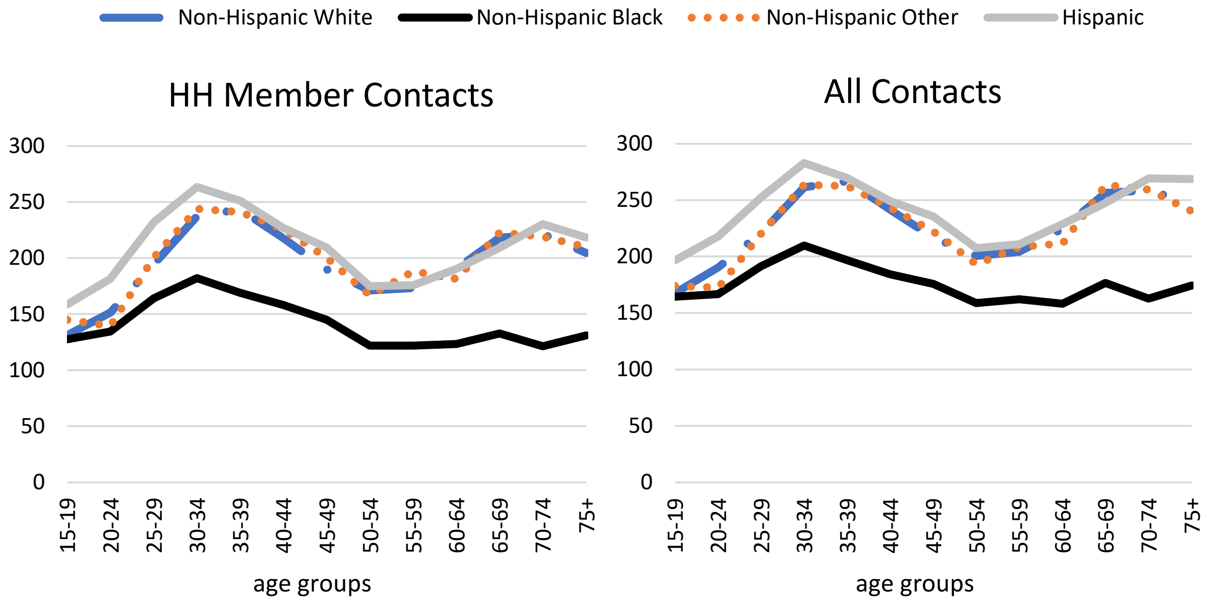


Figure 5. Race and age patterns of social contacts in all locations. Non-Hispanic Blacks had the shortest duration of contact, with a substantial and persistent gap between them and Non-Hispanic Whites/Hispanics at all age groups.

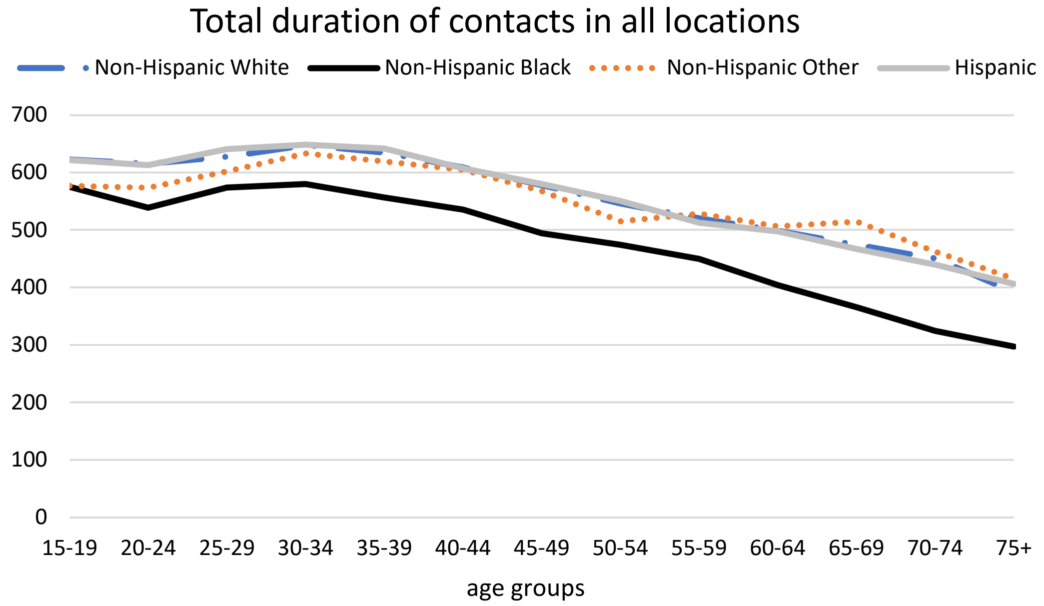


Figure 6. Racial differences in O*NET occupational physical proximity. **(A)** O*NET physical proximity category composition by race shows that Non-Hispanic Blacks have the highest proportion of jobs with the highest level of physical proximity. **(B)** Racial composition of O*NET physical proximity categories show that Non-Hispanic Whites make up a decreasing share and Non-Hispanic Blacks make up a correspondingly increasing share of the highest physical proximity categories.

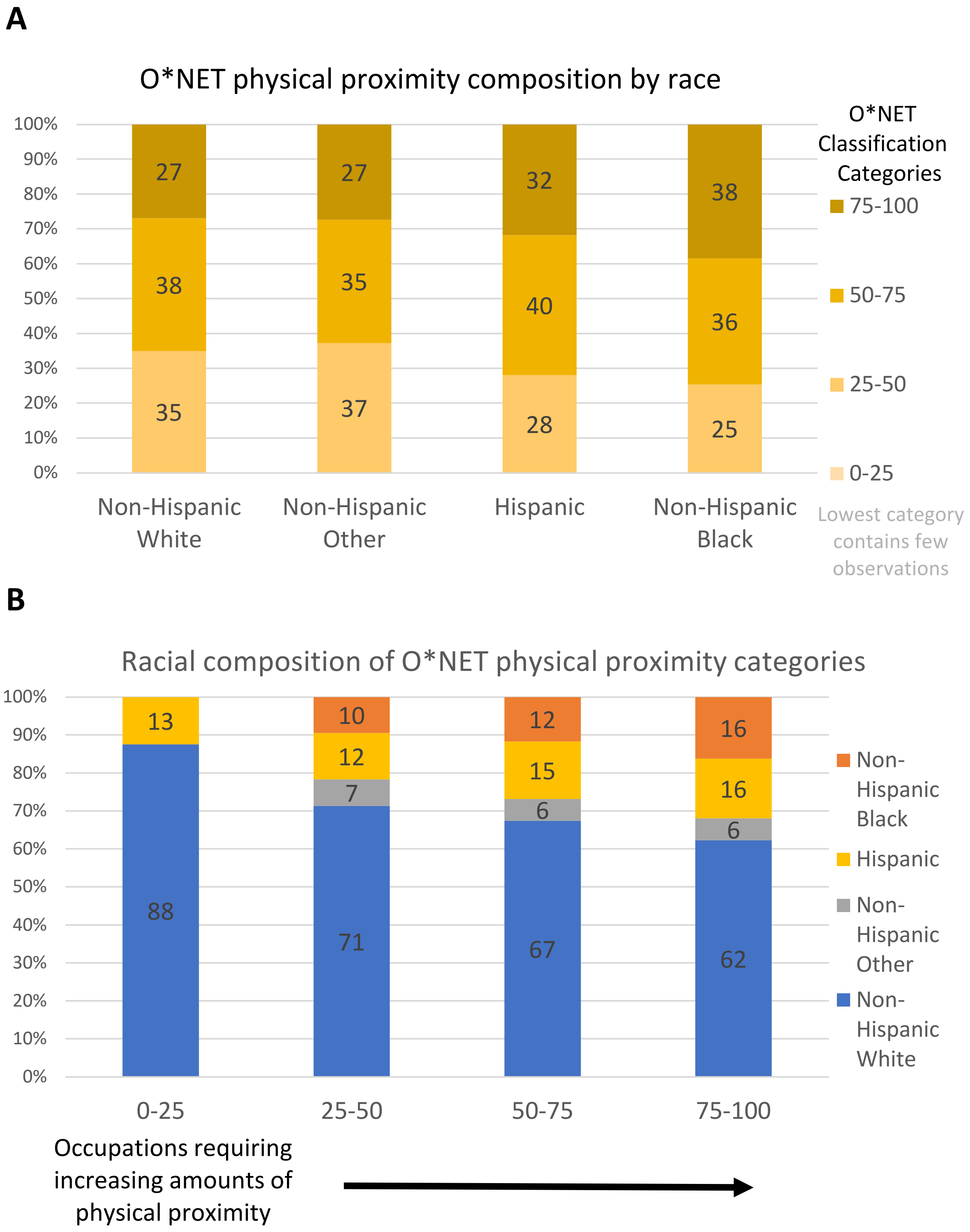
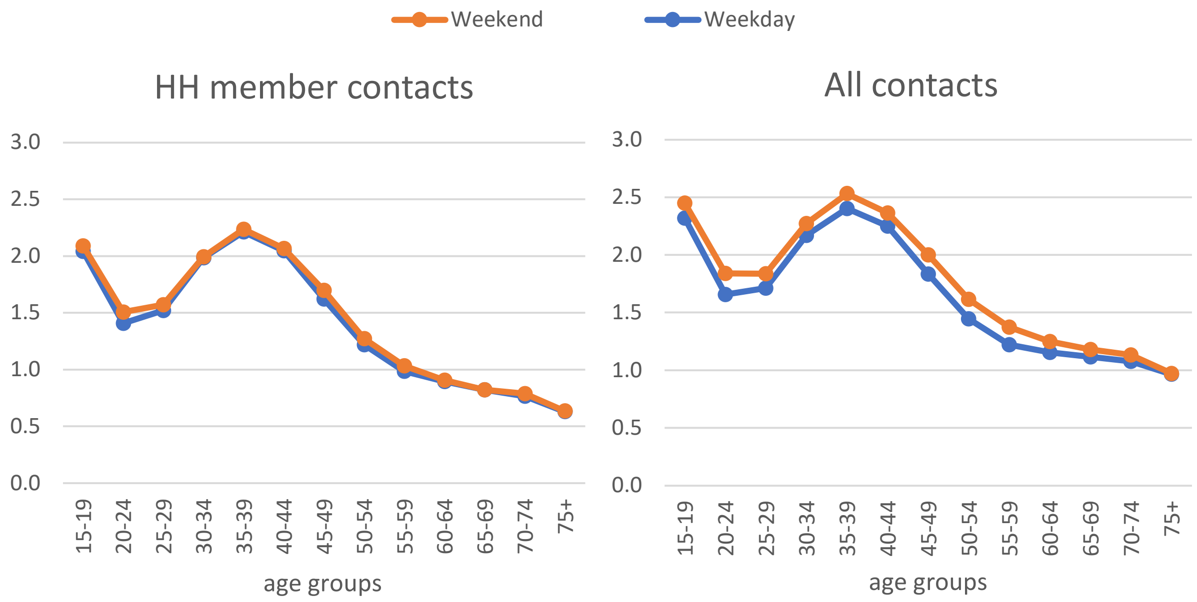


Figure 7. Differences in household contacts during Weekdays and Weekends. There does not appear to be a large difference between the average number of household member contacts between weekend and weekdays but during the weekend there does appear to be a small increase in the number of non-household member contacts taking place in the home. Across most respondents' age groups, the mean duration of contacts is longer during the weekend compared to weekdays.

Average **number** of contacts in home or yard by respondents age groups.



Average **duration** of contacts in home or yard by respondents age groups.

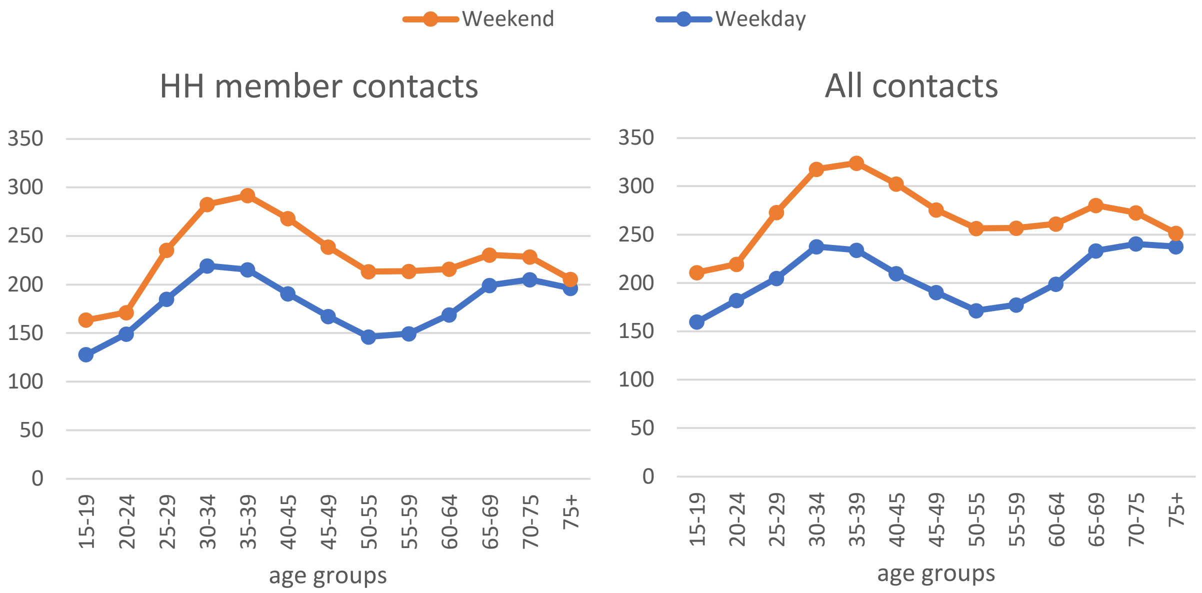


Figure 8. Differences in social contacts during weekdays and weekends in all locations. For most age groups, there is no significant difference in the duration of contact between weekdays and weekends. The duration of contact is slightly higher during the weekday for those under age 25, and slightly higher during the weekend for those 60 and older.

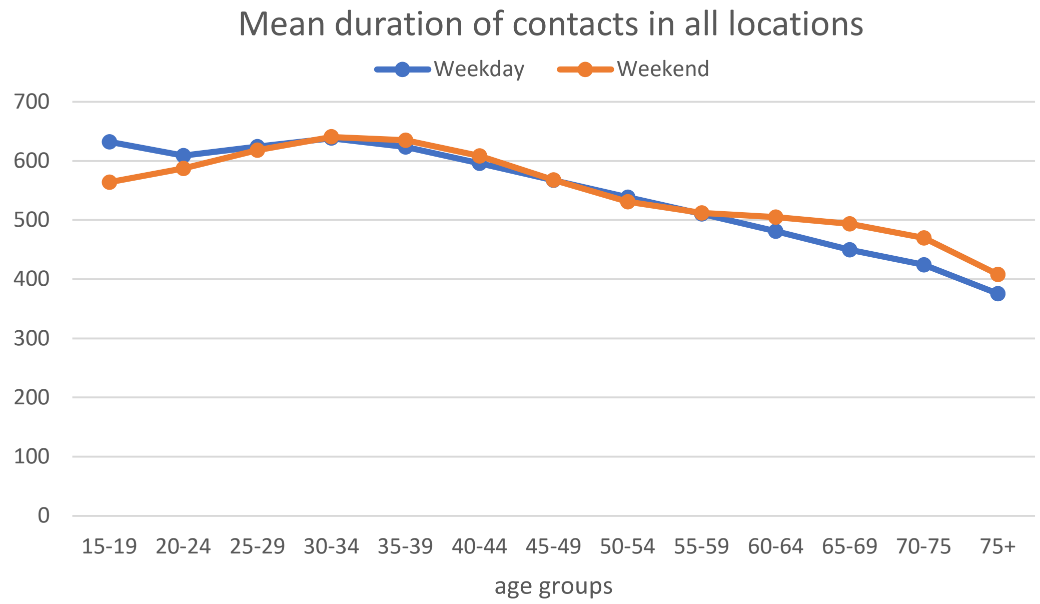
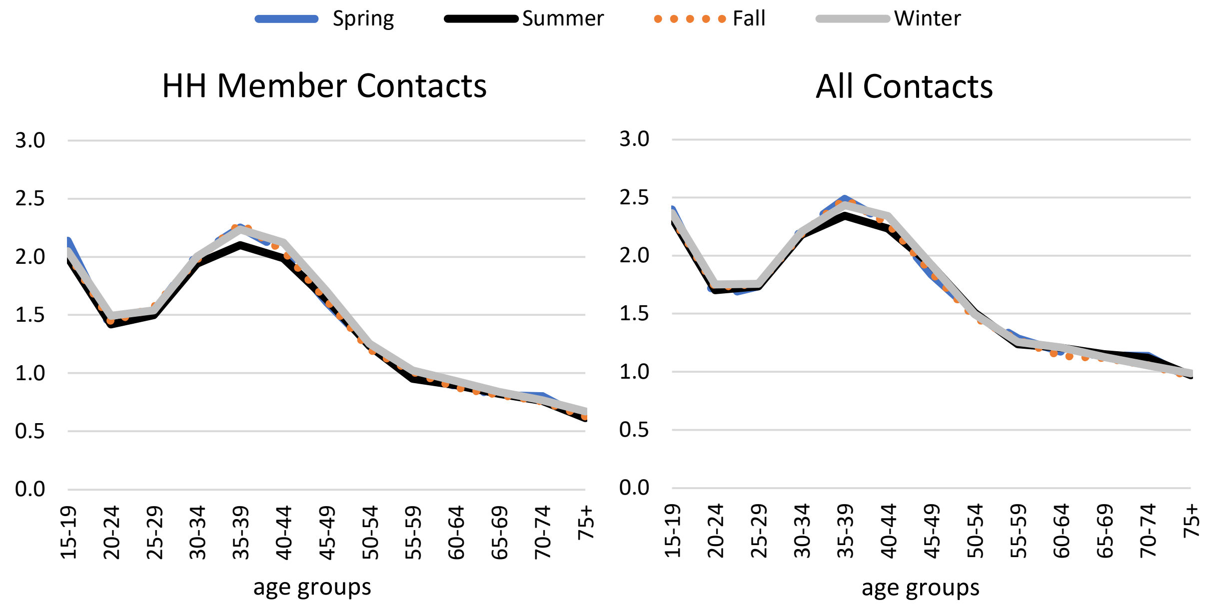
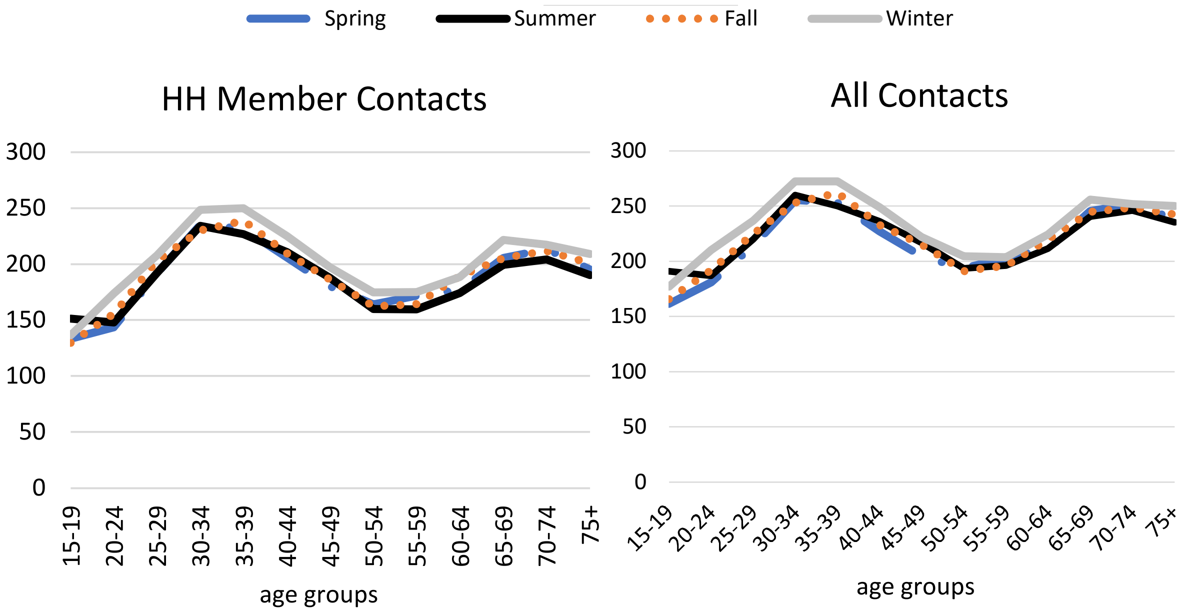


Figure 9. Seasonal and age pattern of social contacts in US. There do not appear to be large seasonal differences in the mean number of contacts; however, the duration of social contacts varies seasonally. Winter months have the highest duration of social contacts. Summer months have the lowest number of social contacts. Sleeping and other personal activities are not included.



Average **number** of contacts in home or yard by respondents age groups.



Average **duration** of contacts in home or yard by respondents age groups.

Figure 10. Seasonal and age patterns of social contacts in all locations. There do not appear to be meaningful seasonal differences in the duration of social contacts in all locations. However teens and young adults have lower durations of contacts in the summer than in the other seasons corresponding to school patterns.

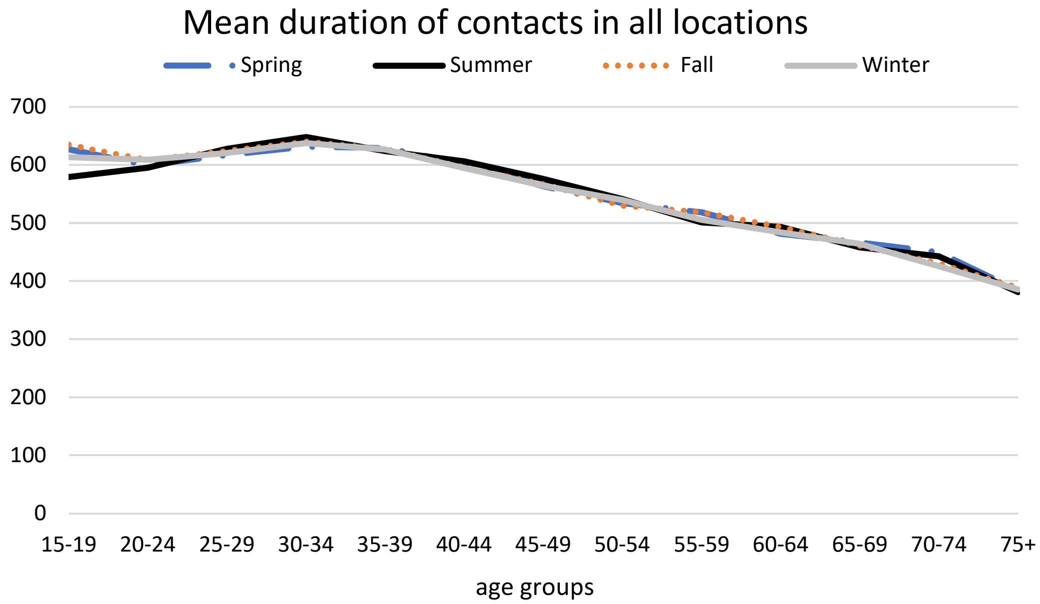
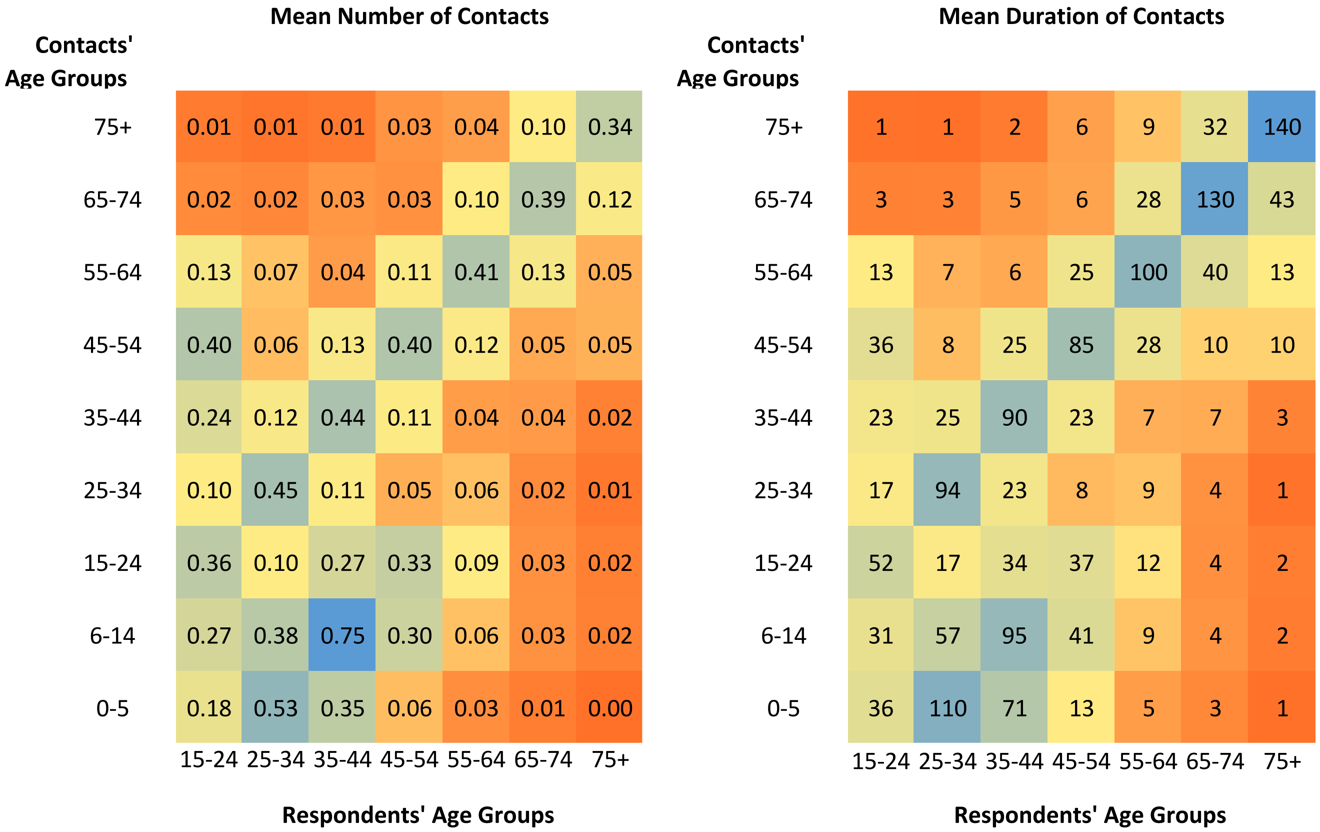


Figure 11. Age contact matrices showing mean number and mean duration of household member contacts. They display assortative contacts with age, and people one generation apart spending time with each other (parents and children). Based on ATUS 2003-2018 surveys. ATUS respondents had to be at least 15 years old.



Mean Number of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	0.01	0.01	0.01	0.03	0.04	0.10	0.34
65-74	0.02	0.02	0.03	0.03	0.10	0.39	0.12
55-64	0.13	0.07	0.04	0.11	0.41	0.13	0.05
45-54	0.40	0.06	0.13	0.40	0.12	0.05	0.05
35-44	0.24	0.12	0.44	0.11	0.04	0.04	0.02
25-34	0.10	0.45	0.11	0.05	0.06	0.02	0.01
15-24	0.36	0.10	0.27	0.33	0.09	0.03	0.02
6-14	0.27	0.38	0.75	0.30	0.06	0.03	0.02
0-5	0.18	0.53	0.35	0.06	0.03	0.01	0.00

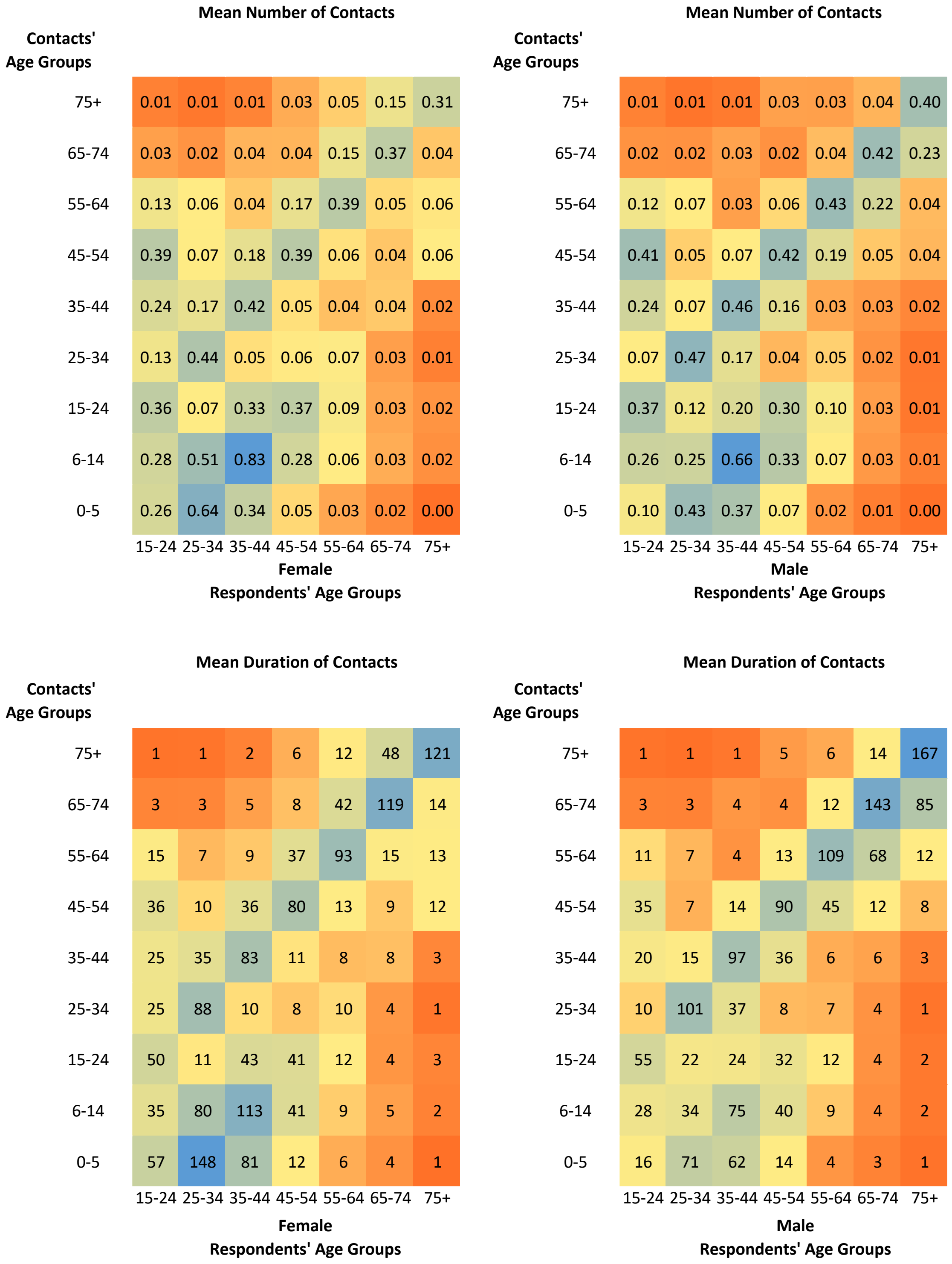
Respondents' Age Groups

Mean Duration of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	1	1	2	6	9	32	140
65-74	3	3	5	6	28	130	43
55-64	13	7	6	25	100	40	13
45-54	36	8	25	85	28	10	10
35-44	23	25	90	23	7	7	3
25-34	17	94	23	8	9	4	1
15-24	52	17	34	37	12	4	2
6-14	31	57	95	41	9	4	2
0-5	36	110	71	13	5	3	1

Respondents' Age Groups

Figure 12. Age contact matrices showing the mean number and duration of household member contacts by sex. They display assortative contacts with age and parents, especially women, spending more time with children. Women have more contact with the age group directly above theirs, while men have more contact with the age group directly below theirs. Based on ATUS 2003-2018 surveys. ATUS respondents had to be at least 15 years old.



Mean Number of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	0.01	0.01	0.01	0.03	0.05	0.15	0.31
65-74	0.03	0.02	0.04	0.04	0.15	0.37	0.04
55-64	0.13	0.06	0.04	0.17	0.39	0.05	0.06
45-54	0.39	0.07	0.18	0.39	0.06	0.04	0.06
35-44	0.24	0.17	0.42	0.05	0.04	0.04	0.02
25-34	0.13	0.44	0.05	0.06	0.07	0.03	0.01
15-24	0.36	0.07	0.33	0.37	0.09	0.03	0.02
6-14	0.28	0.51	0.83	0.28	0.06	0.03	0.02
0-5	0.26	0.64	0.34	0.05	0.03	0.02	0.00

Female
Respondents' Age Groups

Mean Number of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	0.01	0.01	0.01	0.03	0.03	0.04	0.40
65-74	0.02	0.02	0.03	0.02	0.04	0.42	0.23
55-64	0.12	0.07	0.03	0.06	0.43	0.22	0.04
45-54	0.41	0.05	0.07	0.42	0.19	0.05	0.04
35-44	0.24	0.07	0.46	0.16	0.03	0.03	0.02
25-34	0.07	0.47	0.17	0.04	0.05	0.02	0.01
15-24	0.37	0.12	0.20	0.30	0.10	0.03	0.01
6-14	0.26	0.25	0.66	0.33	0.07	0.03	0.01
0-5	0.10	0.43	0.37	0.07	0.02	0.01	0.00

Male
Respondents' Age Groups

Mean Duration of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	1	1	2	6	12	48	121
65-74	3	3	5	8	42	119	14
55-64	15	7	9	37	93	15	13
45-54	36	10	36	80	13	9	12
35-44	25	35	83	11	8	8	3
25-34	25	88	10	8	10	4	1
15-24	50	11	43	41	12	4	3
6-14	35	80	113	41	9	5	2
0-5	57	148	81	12	6	4	1

Female
Respondents' Age Groups

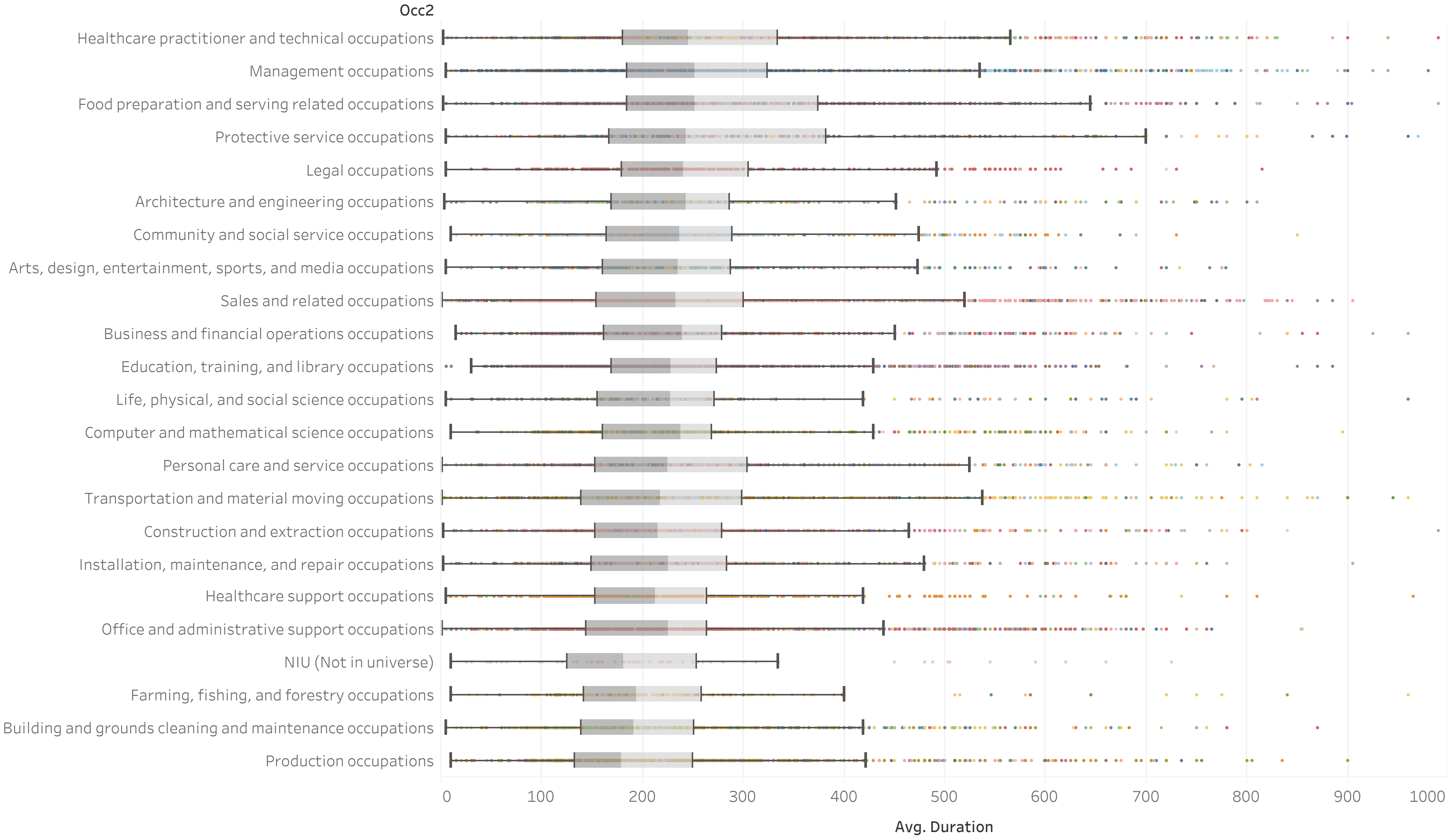
Mean Duration of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	1	1	1	5	6	14	167
65-74	3	3	4	4	12	143	85
55-64	11	7	4	13	109	68	12
45-54	35	7	14	90	45	12	8
35-44	20	15	97	36	6	6	3
25-34	10	101	37	8	7	4	1
15-24	55	22	24	32	12	4	2
6-14	28	34	75	40	9	4	2
0-5	16	71	62	14	4	3	1

Male
Respondents' Age Groups

Appendix A.

Figure A1. Average Duration (in minutes) of Social Contact at Work by Occupation Code 2010-2018



Average of Duration for each Occ2. Color shows details about Occ. Details are shown for Caseid.

Average Duration (in minutes) of Social Contact at Work by Occupation Code 2010-2018

Created by Audrey Dorélien (dorelien@umn.edu) based on ATUS data

Occ2	
Healthcare practitioner and technical occupations	247.96
Management occupations	245.71
Food preparation and serving related occupations	245.58
Protective service occupations	242.28
Legal occupations	235.46
Architecture and engineering occupations	224.93
Community and social service occupations	224.03
Arts, design, entertainment, sports, and media occupations	221.63
Sales and related occupations	220.08
Business and financial operations occupations	216.41
Education, training, and library occupations	216.04
Life, physical, and social science occupations	212.22
Computer and mathematical science occupations	212.04
Personal care and service occupations	209.84
Transportation and material moving occupations	206.44
Construction and extraction occupations	204.60
Installation, maintenance, and repair occupations	202.80
Healthcare support occupations	201.58
Office and administrative support occupations	196.37
NIU (Not in universe)	194.63
Farming, fishing, and forestry occupations	189.51
Building and grounds cleaning and maintenance occupations	185.77
Production occupations	179.08

Average of Duration broken down by Occ2.

Appendix B: O*NET Data and Methods

Description of O*NET Data:

The Occupational Information Network (O*NET) database is the nation's primary source of occupational information and is developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration. The O*NET database contains hundreds of standardized and occupation-specific descriptors across six domains on 974 occupations covering the entire range of the U.S. economy. The O*NET-SOC taxonomy is based on the Standard Occupational Classification (SOC) and defines the set of occupations across the world of work. The O*NET database is free and publicly available and is continually updated from input by a broad range of workers in each occupation ("About O*NET," n.d.).

Data collection began in 2001 and incorporates a two-stage design in which first, a statistically random sample of businesses expected to employ workers in the targeted occupations are identified, and second, a random sample of workers in those occupations within those businesses are selected. New data is collected by surveying job incumbents using standardized questionnaires ("O*NET Data Collection Overview," n.d.). Physical proximity is collected under the domain of occupational requirements and work context and asks respondents to answer the question "How physically close to other people are you when you perform your current job?" on a 5-point scale of increasing physical proximity, from 1 - "I don't work near other people (beyond 100 ft)" to 5 – "Very close (near touching)" ("Work Context: Physical Proximity,", n.d.).

Creating crosswalk:

To integrate the O*NET physical proximity data with ATUS race and occupation, we used a O*NET/Census and 2010 Standard Occupational Classification crosswalk generated by Jennifer Hook. O*NET occupation codes are a more detailed variant of the 2010 Standard Occupational Classification (SOC) codes. O*NET's SOC scheme consists of 974 occupation codes, while SOC 2010 uses only 749 codes. There is no prevalence data for O*NET's more detailed occupation codes. 85 percent of SOC codes link directly to one O*NET code. For the rest, the majority of the variation (57-85%) is found to be between SOC codes and not within detailed sub-occupations within SOC codes. Therefore, the first O*NET match was selected to trim O*NET to 749 SOC codes, taking the first value and dropping the other detailed variants.

Then the crosswalk was generated by merging the Census occ codes and the 2010 SOC codes. This created 882 records (matched to allow cases from each dataset). The file was then manually edited to reduce the crosswalk to Census codes. Although many finer detailed SOC codes were deleted, some SOC/O*NET codes had to be applied to multiple Census codes (e.g. the Census codes are more detailed for types of nurses). The SOC/O*NET codes that most closely matched the Census description were selected, with a preference to codes that had O*NET data (there were many "remainder" categories - i.e. "all other" occupation categories that did not have O*NET data).

Merging crosswalk with O*NET and ATUS data:

The crosswalk was then merged with the O*NET physical proximity dataset, resulting in 514 merged records. The resultant dataset was merged with an ATUS dataset containing respondents with work activities and sociodemographic information (including race and occupation) for years 2011-2018, excluding Alaska and Hawaii. 31,069 records (95%) matched. The majority (72%) of ATUS records that didn't match were NIU for occupation. The remaining unmatched ATUS records were distributed across only a handful of occ codes, which were in "all other" or "miscellaneous" categories. Physical proximity was then collapsed and tabulated by respondent race to obtain mean estimates and percentages.

References

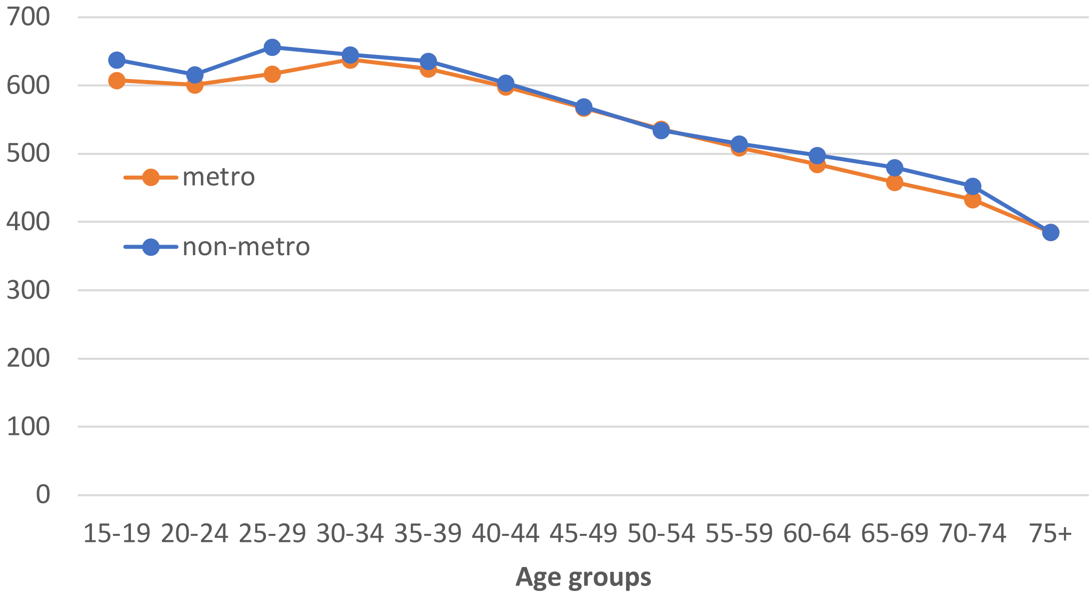
"About O*NET." *O*NET Resource Center*, National Center for O*NET Development, www.onetcenter.org/overview.html. Accessed 15 July 2020.

"O*NET® Data Collection Overview." *O*NET Resource Center*, National Center for O*NET Development, www.onetcenter.org/dataCollection.html. Accessed 15 July 2020.

"Work Context: Physical Proximity." *O*NET OnLine*, National Center for O*NET Development, www.onetonline.org/find/descriptor/result/4.C.2.a.3. Accessed 15 July 2020.

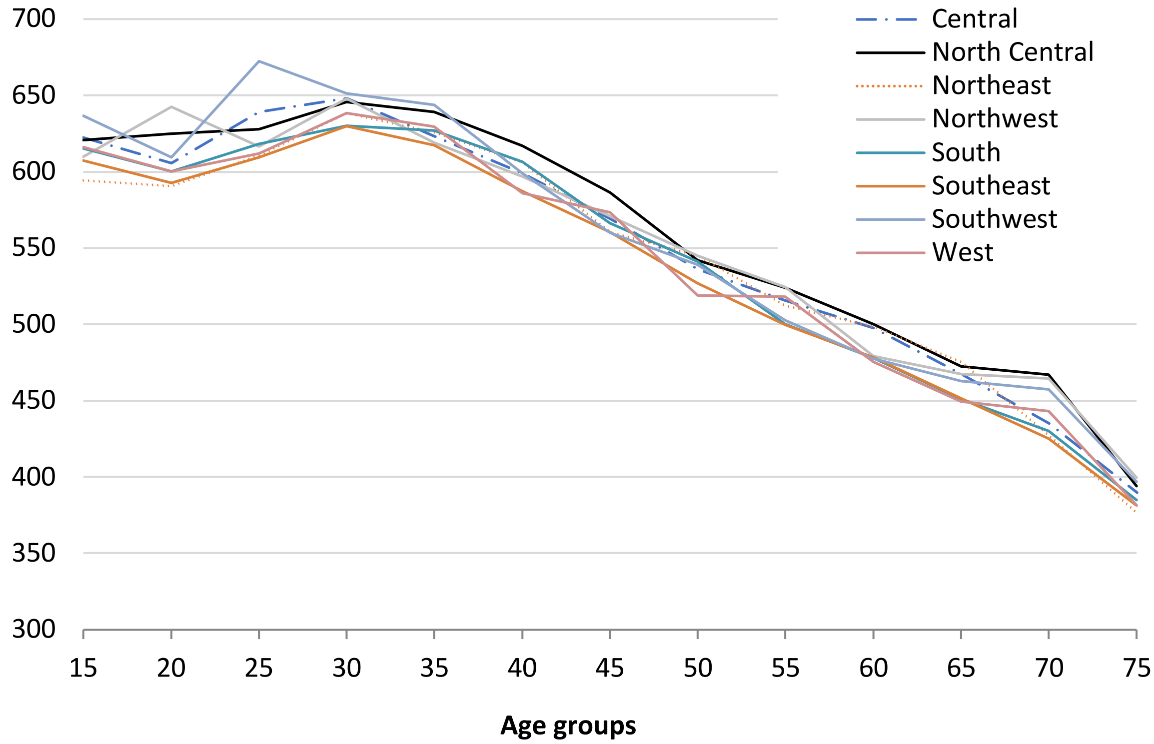
Appendix C. Spatial differences in contact patterns.

Figure C1. Total **duration** (minutes) of contacts for metro versus non-metro respondents by age groups.



The age patterns for total duration of contacts for metro versus non-metro respondents throughout the contiguous United States are similar. The overall shape is the same, with a slight increase for respondents at ages 20-29, followed by a gradual decrease for all respondents ages 30 and above. The total duration of social contacts for metro respondents is consistently the same or lower than the total duration for non-metro respondents.

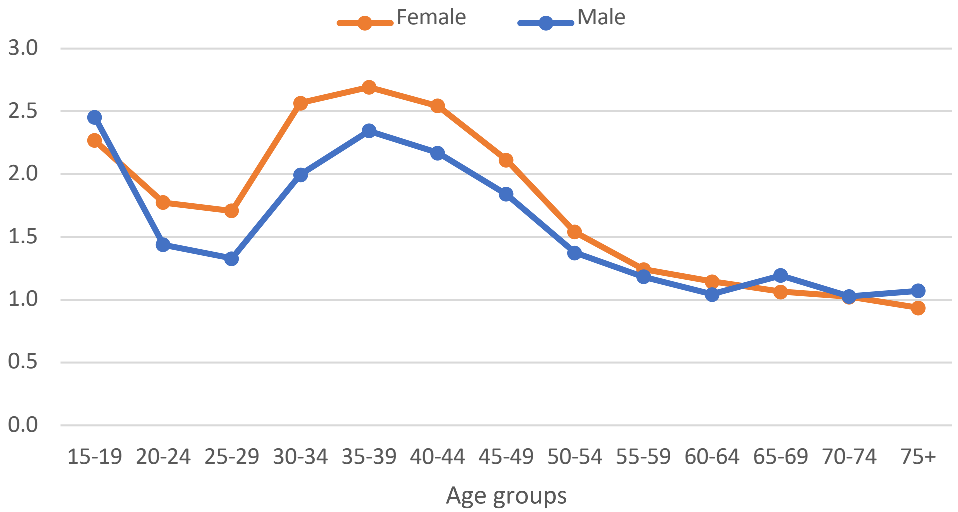
Figure C2. The age pattern of total **duration** (minutes) of contacts by climatic region.



The age patterns for total duration of contacts by climatic regions follow similar trends. There is a peak between ages 25 to 35, followed by a gradual decline. Out of all the climatic regions listed, we see that the South features an obvious peak at age 25, rising above all other regions only to join their fold as the age increases.

Figure C3. Age pattern of social contacts in respondent's home/yard for residents of the North Central region. Although there are some differences, the patterns are similar to those found in Figure 2 which depict the age pattern for the contiguous US overall.

Panel A. Average **number** of all contacts in respondent's home or yard.



Panel B. Average **duration** (minutes) of all contacts in respondent's home or yard.

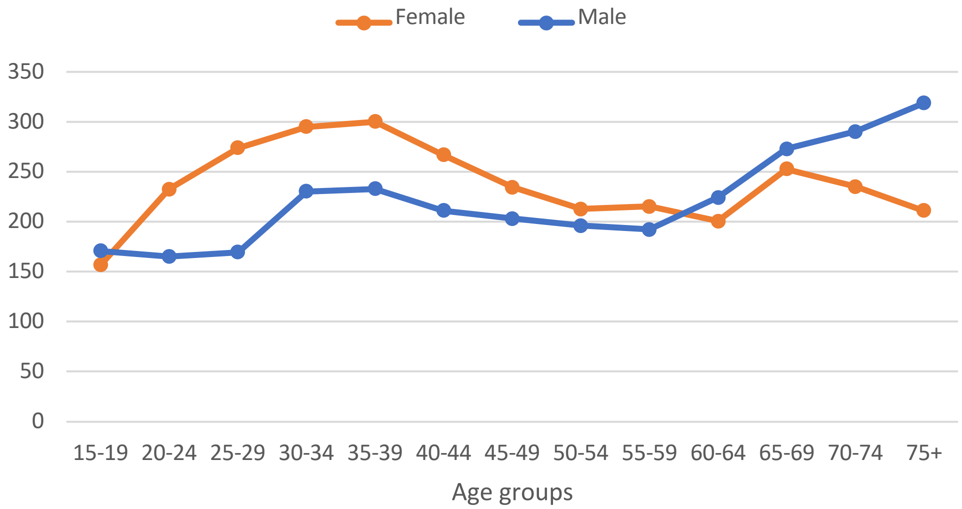
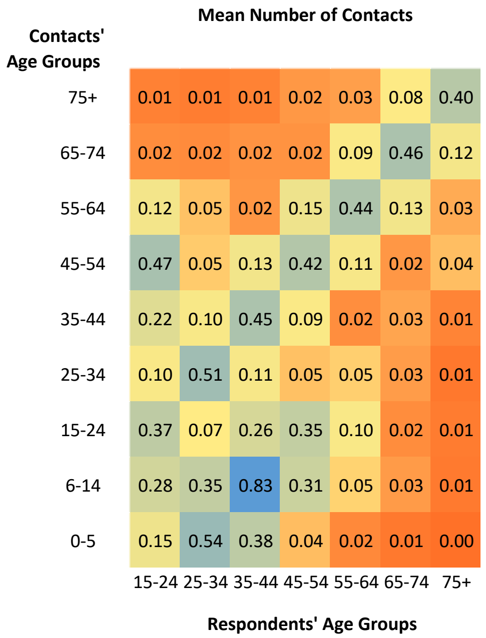


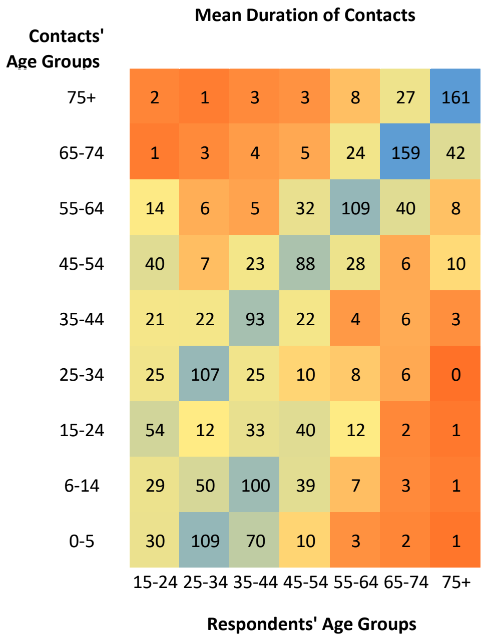
Figure C4. Comparing the age-contact matrices for North Central region with US.

North Central Region


Mean Number of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	0.01	0.01	0.01	0.02	0.03	0.08	0.40
65-74	0.02	0.02	0.02	0.02	0.09	0.46	0.12
55-64	0.12	0.05	0.02	0.15	0.44	0.13	0.03
45-54	0.47	0.05	0.13	0.42	0.11	0.02	0.04
35-44	0.22	0.10	0.45	0.09	0.02	0.03	0.01
25-34	0.10	0.51	0.11	0.05	0.05	0.03	0.01
15-24	0.37	0.07	0.26	0.35	0.10	0.02	0.01
6-14	0.28	0.35	0.83	0.31	0.05	0.03	0.01
0-5	0.15	0.54	0.38	0.04	0.02	0.01	0.00

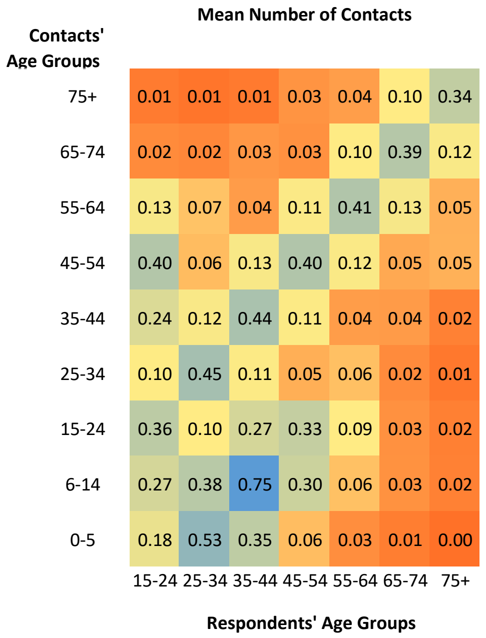
Respondents' Age Groups


Mean Duration of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	2	1	3	3	8	27	161
65-74	1	3	4	5	24	159	42
55-64	14	6	5	32	109	40	8
45-54	40	7	23	88	28	6	10
35-44	21	22	93	22	4	6	3
25-34	25	107	25	10	8	6	0
15-24	54	12	33	40	12	2	1
6-14	29	50	100	39	7	3	1
0-5	30	109	70	10	3	2	1

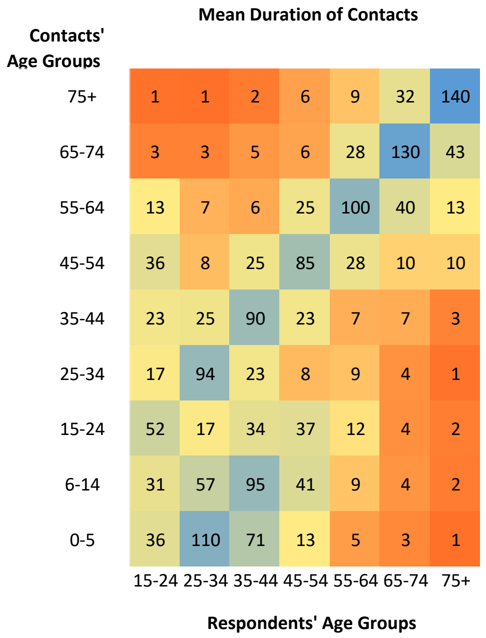
Respondents' Age Groups

Contiguous US


Mean Number of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	0.01	0.01	0.01	0.03	0.04	0.10	0.34
65-74	0.02	0.02	0.03	0.03	0.10	0.39	0.12
55-64	0.13	0.07	0.04	0.11	0.41	0.13	0.05
45-54	0.40	0.06	0.13	0.40	0.12	0.05	0.05
35-44	0.24	0.12	0.44	0.11	0.04	0.04	0.02
25-34	0.10	0.45	0.11	0.05	0.06	0.02	0.01
15-24	0.36	0.10	0.27	0.33	0.09	0.03	0.02
6-14	0.27	0.38	0.75	0.30	0.06	0.03	0.02
0-5	0.18	0.53	0.35	0.06	0.03	0.01	0.00

Respondents' Age Groups


Mean Duration of Contacts

Contacts' Age Groups	15-24	25-34	35-44	45-54	55-64	65-74	75+
75+	1	1	2	6	9	32	140
65-74	3	3	5	6	28	130	43
55-64	13	7	6	25	100	40	13
45-54	36	8	25	85	28	10	10
35-44	23	25	90	23	7	7	3
25-34	17	94	23	8	9	4	1
15-24	52	17	34	37	12	4	2
6-14	31	57	95	41	9	4	2
0-5	36	110	71	13	5	3	1

Respondents' Age Groups