README.txt

This folder contains the main, build, analysis, and table/figure creation code. The analysis was run with Stata18 and R version 4.4.0.

The analysis uses confidential data which cannot be shared. The data was obtained via a partnership with an e-commerce company with whom the authors signed a non-disclosure agreement to obtain and analysis the data. For information regarding how others might attempt obtain similar data, please reach out to the authors directly at laneg@uchicago.edu.

---------- Code

******* REQUIRED PROGRAMS

* R
chron
cld2
countrycode
data.table
dplyr
foreign
googleLanguageR
foreign
Haven
kableExtra
knitr
magrittr
vader
viridian
readxl
readstata13
rgeolocate
sentimentr
scales
stringr
textstem
tidyverse

tidytext
tm

* Stata
reghdfe
estout

******* INSTRUCTION TO RUN CODE

To produce all the figures and tables in the paper, the code was run in the follow order:

STEP 1 - RUN master.R - MUST UPDATE PATH TO THE REPLICATION FOLDER ON LINE 19

STEP 2 - RUN master.DO - MUST UPDATE PATH TO THE REPLICATION FOLDER ON LINE 12

Note: Chat translation may be required and is called in master.R. This step requires a google key and may have associated fees from google depending on sample size. Therefore researchers may be required to set up an account with google, generate a key, and provide funds.

******** FILE STRUCTURE

* MAIN FOLDER - The code files in the main folder are

- master.do

* CONSTRUCTS VARIABLES AROUND ETHNICITY AND PRODUCES ALL TABLES IN FIGURES FOUND IN PAPER

- master.R

* READS IN THE RAW CHAT DATA AND MERGES IT WITH SALES AND AGENTS DATA TO CONSTRUCT KEY VARIABLES FOR ANALYSIS

** SUB FOLDERS - The files in the sub folders are

*CODE sub folder

- 0.randomization

0_treatment_assignment.R

* PRODUCES RANDOM ASSIGNMENT OF MALE AND FEMALE NAMES TO CHAT AGENTS

- 1.cleaning

1_chatcleaning_raw.R

* CONVERTS CHAT DATA TO 1-MESSAGE PER ROW
* TRANSLATES ALL CHAT MESSAGES TO ENGLISH (THIS STEP OFF BY DEFAULT - REQUIRES GOOGLE KEY AND $ TO RUN)
* CALCULATES SENTIMENT SCORE OF EACH MESSAGE

2_chatcleaning_handcoding.R

* CODE QUESTIONS AND VARIABLES IN THE HAND-CODED CHAT DATA INTO A STANDARD FORMAT

- 2.construct

1_building_intermediate.R

* CONSTRUCT AGENT-DAY LEVEL DATA
* CONSTRUCT AGENT-CHAT LEVEL DATA
* MERGE AGENT-CHAT DATA WITH SALES DATA TO CREATE SALES OUTCOMES

2_building_final.R

* FURTHER PREP THE DATASETS CREATED IN THE STEP ABOVE FOR ANALYSIS

3_build_ethnicity_assignment.do

* MERGES NAME ETHNICITIES WITH NAME ASSIGNMENTS

4_build_english_name_assignment.do

* MERGES NAME ENGLISH SOUNDING HAND CODING WITH NAME ASSIGNMENTS

5_build_country_locations.R

* GEO CODES IP ADDRESS TO COUNTRIES

- 3.analysis

sales_analysis.do

* CREATES ALL TABLES AND FIGURES

- ado

balanceTable

* CREATES TABLE OF SINGLE AND MULTIPLE VARIABLE RANDOMIZATION BALANCE CHECKS

bysmeanw

* HELPER FILE USED IN BALANCE CHECKS; USED FOR DEMEANING

effectXtime

* CREATES FIGURES FOR TREATMENT EFFECTS OVER TIME

flag missing

* HELPER FILE USED IN BALANCE CHECKS; CHECKS VARIABLE MISSINGNESS

MultiPartTabEnd; MultiPartTabPanelEnd; MultiPartTabPanelStart; MultiPartTabPanelStartMoved; MultiPartTabStart

* SUITE OF FUNCTIONS TO CREATE SPECIALITY MULTI-PANEL TABLES

tabnoteWidth

* CHANGES TABLE NOTE WIDTHS

weightControlMean

* CALCULATES WEIGHTED CONTROL MEANS ACCOUNTING FOR REGRESSION FIXED EFFECTS

* Data

CANNOT BE SHARED

---------- Table Folders

The "table" folder is a shell folder that will populate with figures and tables produced by the code

---------- List of main tables and figures ----------

Figures

	Program	Line Number
Figure 1: Effects of female name assignment over time	sales_analysis.do	160

Tables

Table 1: Balance tests for female assignment	sales_analysis.do	186
Table 2: Effect of female assignment on purchase outcomes	sales_analysis.do	261
Table 3: Effect of female name assignment on chat response and purpose	sales_analysis.do	303
Table 4: Correlational relationship between female agent and administrative outcomes	sales_analysis.do	434
Table 5: Effect of female and non-English name assignment on purchase outcomes	sales_analysis.do	560

---------- List of appendix tables and figures ----------

Tables

	Program	Line Number
Table A1: Example names used in name assignment, by gender and non-English status	sales_analysis.do	700
Table A2: Outcome Variable Descriptions	NOT DATA OUTPUT	N/A
Table A3: Effect of female assignment on purchase prices	sales_analysis.do	791
Table A4: Effect of female assignment on initial chat purpose		

sales_analysis.do 862
Table A5: Effect of female assignment on purchase outcomes (customer-day level) sales_analysis.do 915
Table A6: Effect of female assignment on purchases, by outcome source and non-gender sample sales_analysis.do 958
Table A7: Effect of female assignment on purchase outcomes, alternative specifications sales_analysis.do 1024
Table A8: Name features and interactions on purchase outcomes sales_analysis.do 1210
Table A9: Effect of female name assignment and customer gender sales_analysis.do 1382
Table A10: Effect of female assignment on any sales by agent (48 hours) sales_analysis.do 1456
Table A11: Effect of female assignment on chat tones sales_analysis.do 1562

------- Memory and Runtime Requirements ----------

Memory Requirements: 2GB

Runtime R: 1 hour
Runtime Stata: 6 minutes
Run on MacBook Pro M1 with 16gb RAM