Title
Essays in Applied Microeconomics of Decision Making Under Uncertainty

Author
Uruci, Xhulio

Publication Date
2024

University of California
Santa Barbara

Essays in Applied Microeconomics of Decision Making Under Uncertainty

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Xhulio Uruci

Committee in charge:

Professor Peter Kuhn, Chair
Professor Richard Startz
Professor Alisa Tazhitdinova

June 2024

The Dissertation of Xhulio Uruci is approved.

Professor Richard Startz

Professor Alisa Tazhitdinova

Professor Peter Kuhn, Committee Chair

May 2024

Essays in Applied Microeconomics of

Decision Making Under Uncertainty

by

Xhulio Uruci

Acknowledgements

I am grateful to a number of people who have supported me in making this dissertation possible. Peter Kuhn, who has offered endless encouragement and many emails and pages of detailed feedback on my work; Dick Startz, who has provided valuable guidance, both on my research and on navigating several challenges throughout my time in graduate school; and Alisa Tazhitdinova, who has provided many hours of feedback in meetings, and challenged me to be a better and bolder researcher. I am very grateful to have been advised by this committee.

I would like to thank Mark Patterson, who has been a tremendous help in navigating many administrative obstacles throughout my time at UCSB. I would like to thank the economics faculty members, colleagues, and friends at UCSB for their support.

Most of all, I would like to thank my family. My parents, Alfred and Ariana Uruci, and my brother, Adnan Uruci, whose love, support, and guidance have been my rock my entire life. My parents have taken enormous risks and made many sacrifices to have provided me with the opportunity to pursue a Ph.D. and I am forever grateful.

Curriculum Vitæ

Xhulio Uruci

Education

2024	Ph.D. in Economics (Expected), University of California, Santa Barbara.
2019	M.A. in Economics, University of California, Santa Barbara.
2014	B.S. in Mathematics, University of Massachusetts, Amherst.

Interests Applied Microeconomics: Labor, Migration, Discrimination.

Research

Animus or Beliefs? Consumer Discrimination In Uncertain Times.

Consumer Discrimination and Social Media In Uncertain Times.

Isolating the Effect of Grant Rates on Asylum Applications: A Relative Discrete Choice Approach.

Conferences and Presentations

2023	All California Labor Economics Conference, UC Santa Barbara.
2023	Western Economic Association International, San Diego.
2023	Applied Microeconomics Lunch, UC Santa Barbara.
2022	Young Economists Symposium, Yale University.

Honors and Awards

2023	Outstanding Undergraduate TA, UC Santa Barbara.
2023	Job Market Fellowship, UC Santa Barbara.
2022	Summer School in the Economics of Migration, UC Davis.
2021	Graduate Research Fellowship, UC Santa Barbara.
2020	Mortimer Andron Fellowship, UC Santa Barbara.
2014	Commonwealth Honors College, UMass Amherst.

Research Experience

2022	Research Assistant for Alisa Tazhitdinova, UC Santa Barbara.
2018	Research Assistant for Yun Kim, UMass Boston.

Teaching Experience

As Instructor:

3 Terms	Financial Management, UC Santa Barbara.
1 Term	Math Camp (Ph.D.), UC Santa Barbara.

As Assistant:

6 Terms	Personnel Economics, UC Santa Barbara.
3 Terms	Financial Management, UC Santa Barbara.
3 Terms	Intermediate Microeconomics, UC Santa Barbara.
3 Terms	Introductory Macroeconomics, UC Santa Barbara.

Coding R Programming Language, Python, Stata, SQL, Matlab, Latex.

Languages Albanian (native); English (fluent).

Dissertation Committee

Chair	Peter Kuhn, Distinguished Professor of Economics.
Member	Richard Startz, Professor of Economics.
Member	Alisa Tazhitdinova, Assistant Professor of Economics.

Abstract

Essays in Applied Microeconomics of

Decision Making Under Uncertainty

by

Xhulio Uruci

My dissertation consists of three chapters on discrimination, social media, and immigration, enriching the empirical literature of economic decision-making under risk and uncertainty.

Chapter 1 explores the relative roles of beliefs and prejudice in the early weeks of the COVID-19 pandemic. Neither beliefs nor preferences are observed in most studies of discrimination, making it hard to discern the underlying motive. This study overcomes the challenge by testing two plausible motives in consumer discrimination following the first case of COVID-19. Did consumers avoid Chinese restaurants due to belief-based concerns over virus risks or due to anti-Chinese sentiment? Using foot-traffic data on the universe of US restaurants, I exploit variation in service type to capture perceived health risks and variation in the ethnic mix of customers to capture anti-Chinese sentiment. I find strong evidence consistent with belief-based discrimination and inconsistent with taste-based discrimination.

Chapter 2 is closely related to Chapter 1 and examines the role of social media within the same observational setting. Specifically, this paper explores the role of social media in demand changes for restaurants following the first case of COVID-19. Using data on foot-traffic to the universe of US restaurants and Twitter data, I develop a novel identification

strategy and exploit plausibly exogenous spatial variation in pre-covid Twitter usage to find that counties where Covid was a more salient topic of conversation on Twitter saw greater reduction in demand for Chinese restaurants relative to other restaurants. Placebo tests support the results, indicating a causal effect of social media on off-platform market choices.

Chapter 3 studies the destination choices of asylum applicants. I use the destination choices of migrants who arrive under different statuses as a plausible counterfactual to estimate the effect of grant rates (recognition rates) on the share of applications for asylum received by a destination country. Using a relative discrete choice model, I find a robust estimate of a 6.5 percent increase in the share of applications received by destination countries in response to a 10 percentage point increase from the mean grant rate. This approach may provide applied researchers with a simple and convenient method to attain estimates in other settings.

Contents

Chapter 1

Animus or Beliefs? Consumer Discrimination In Uncertain Times

1.1 Introduction

A growing strand of the discrimination literature highlights the importance of belief accuracy when distinguishing taste (preference-based) vs. statistical (belief-based) discrimination (Arnold, Dobbie and Yang 2018; Bohren et al. 2022; Bordalo et al. 2016, 2019; Coffman, Exley and Niederle 2021; Hedegaard and Tyran 2018). In the conventional dichotomy between taste- and *accurate* statistical-discrimination, the researcher compares differences in outcomes with differences in productivity across groups to determine the extent of each type. However, allowing for discriminators to hold inaccurate beliefs leads to an identification challenge. If, for example, the set of tastes and set of beliefs lie in two different continuums, then many combinations of tastes and false beliefs can result in the same outcome, preventing the researcher from discerning the underlying motive. Discerning between the two is important because belief-based discrimination can be mitigated with the provision of correct information whereas taste-based discrimination generally cannot. Researchers have therefore relied on experimental studies that manipulate the provision of information to tease out the type (Bertrand and Duflo 2017). Evidence from natural and more representative settings, however, remains scarce.[1]

In this paper, I present evidence on the relative roles of beliefs and tastes from a natural experiment that affected the entire restaurant-going US population. Using GPS foot-traffic data, I examine the change in demand for restaurants immediately following the first case of COVID-19 (covid) in the US. I use the covid announcement as the exogenous event that made restaurant-goers susceptible to the belief that Chinese restaurants carry a greater risk of infection, prompting greater aversion to visiting Chinese restaurants than to visiting other restaurants. An identification challenge arises because consumers may have also developed anti-Chinese sentiment (animus) that is unrelated to health concerns, such as blaming Chinese people for the virus, and neither health-related concerns nor anti-Chinese sentiment is observed.[2]

To overcome this challenge, I adopt an identification strategy based on a simple intuition of how motives translate into restaurant choice. A belief-driven discriminator, motivated by health concerns, discriminates only against Chinese restaurants with an elcoevated perceived risk of infection, whereas a taste-driven discriminator, motivated by anti-Chinese sentiment, discriminates against all Chinese restaurants regardless of perceived risk. Accordingly, if discrimination is belief-based, it would be strongest among

[1]For example, few if any resume audit studies construct a broadly representative sample of resumes and also manipulate the accuracy of information in the resumes.

[2]Several reports and studies document a rise in anti-Asian hate crimes and online sentiment shortly after the covid announcement (Cable News Network 2020; Department of Justice 2021; Dipoppa, Grossman and Zonszein 2021; Huang et al. 2023; Time Magazine 2020).

restaurants with the highest *perceived* risk. On the other hand, discrimination among restaurants with the lowest *perceived* risk is more likely to be taste-based.

To capture perceived health risk, I use variation in restaurant service type, which is measured by the pre-covid share of visitors with a prolonged duration of stay. Whereas concerns over food-borne transmission (beliefs) and anti-Chinese sentiment (tastes) may be common to all Chinese restaurants, concerns over air-borne transmission (beliefs) would fall only on dine-ins, which require a prolonged stay in proximity to restaurant staff and customers. For this reason, the covid announcement made dine-ins appear riskier than take-outs. Consistent with a belief-based motive and inconsistent with an taste-based motive, I find discrimination against Chinese dine-ins but *not* take-outs, indicating that restaurant-goers were worried about air-borne transmission at Chinese restaurants.

One may be concerned that consumers who avoided Chinese dine-ins were acting on anti-Chinese sentiment because the covid announcement hardened their predisposition to "like the food but not the people." I provide three results that show this is unlikely. The first result exploits variation in the ethnic mix of a restaurant's pre-covid clientele. Following the covid announcement, Chinese customers became *more* reluctant to visit restaurants that typically have a large Chinese clientele, whereas non-Chinese customers did *not*—a result that holds at both Chinese and non-Chinese restaurants. This suggests that Chinese customers developed broadly-based beliefs over health risks from interacting with other Chinese customers, whereas non-Chinese customers did not develop broad beliefs nor taste-based aversion to interacting with Chinese customers. The second result comes from variation by service type. Following the covid announcement, Chinese customers became *just as* likely as non-Chinese customers to avoid Chinese dine-ins, indicating both customer types developed heightened concerns over air-borne risks.

The third result exploits variation by cuisine and highlights a local salience effect. Following new cases in a county, consumers discriminated against Chinese, other Asian, and European restaurants, but not against Latin American restaurants. Since reported cases were rising in China, other Asian countries, and Europe, but not in Latin America during the period I study, this response is consistent with concerns that foreigners were bringing the virus to own-group restaurants in the US. As before, this finding aligns with beliefs over virus risks rather than taste-based motives, which would be directed solely at Chinese or other Asian restaurants.

I use SafeGraph's GPS foot-traffic data on over 40 million smart-phone users to measure restaurant-goer's response to the January 20, 2020 confirmation of the first US case (SafeGraph 2022). The baseline specification employs a triple-differences (DDD) estimation strategy where the consumer response to the covid announcement is estimated as the change in log visits from the seven weeks before January 20 to the seven weeks after—a unique period during which consumers became aware of the virus but no policies had yet been announced that would restrict their restaurant choices. In the second difference, I compare visits in the current year, December 2019 – March 2020, to the prior year, December 2018 – March 2019, which differences out seasonal variation in cuisine preferences. I then take the third difference across cuisines to estimate consumer discrimination.

To measure service and clientele, I rely on the breadth and detail of the data, which includes weekly detail on the count, duration, and home origin of visits to the universe of full-service and limited-service restaurants. I use the share of pre-covid visits that last between 21 minutes and 2 hours, defined as the *dine-share*, as a continuous measure of service type, and this is interacted with DDD indicator variables to estimate the differential response by service type. To measure the ethnic composition of restaurant clientele, I allocate pre-covid visits to customer types based on the ethnic composition of visitor home census tracts, and construct a restaurant's typical *Chinese-share* of customers.

The baseline framework finds no discrimination against Chinese take-outs following the covid announcement. In contrast, consumers reduced visits to Chinese dine-ins by 8.1 percentage points *more* than to other dine-ins. Among Chinese customers, this estimate rises to 12.8 percentage points. After allowing for heterogeneity in beliefs and anti-Chinese sentiment at different cuisines, I show in an extended framework that any scenario where non-Chinese customers exhibited taste-based discrimination would need to take a highly particular form that is not reflected in demand changes: it cannot come from a reduced willingness to do business with Chinese restaurants, nor from aversion to brief interactions with Chinese customers (e.g. a take-out order), nor from aversion to prolonged interactions with Chinese customers (e.g. dining-in).

The central contribution of this paper is to test the relative roles of belief-driven and taste-driven discrimination in a fully observational, natural setting involving millions of US consumers. Bohren et al. (2022) conduct a literature review of top 10 economics journals from 1990 to 2018 and find that although there were 105 papers published on discrimination broadly, seven studies measure beliefs and just one tests for belief accuracy in a quasi-experimental setting.[3] Furthermore, past studies generally examine static settings whereby outcomes and productivities are measured in *levels*, requiring a

[3] Arnold, Dobbie and Yang (2018) find suggestive evidence of inaccurate statistical discrimination in bail decisions. Other studies measuring beliefs include Agan and Starr (2018); Beaman et al. (2009); Fershtman and Gneezy (2001); Hedegaard and Tyran (2018); List (2004); Mobius and Rosenblat (2006). Related studies include Carlana (2019); Coffman, Exley and Niederle (2021); Kline, Rose and Walters (2022) and Mengel, Sauermann and Zolitz (2019).

test of belief accuracy to determine the extent of each motive.[4] This paper differs from most prior studies by examining a highly dynamic and uncertain setting whereby an exogenous shock associated with a particular ethnic group prompted a discriminatory *change* in behavior. As such, this study shows that examining changes in behavior does not necessarily require testing belief accuracy in order to identify the underlying motive.

To the extent that beliefs over cuisine-specific risks were exaggerated, such beliefs may have formed due to cognitive biases such as representative heuristics, whereby restaurant-goers overestimated the additional probability of infection at Chinese restaurants (Bordalo et al. 2016). Another possible bias is representative signal distortion, whereby greater salience of covid made safety concerns more important when evaluating Chinese restaurants than when evaluating other restaurants (Esponda, Oprea and Yuksel 2023). Although effectively providing accurate information can be challenging in uncertain times, the results in this study suggest that doing so has the potential to mitigate discriminatory responses in highly dynamic settings like the one I study.

I also extend the literature on consumer discrimination in response to adverse events (Bartoš et al. 2021; Nardotto and Sequeira 2021; Pandya and Venkatesan 2016) and scapegoating during a crisis (Bursztyn et al. 2022*b*). Contemporaneous work by Yi (2023) also uses SafeGraph data to find discrimination against Chinese restaurants in the early weeks of the covid announcement, but focuses on a different research question relating to the economic consequences for Chinese restaurants and the association with local racial/ethnic diversity and political affiliation. Related, Huang et al. (2023) uses SafeGraph data but a longer time frame, until after the lockdowns and political influ-

[4]For example, past studies have provided evidence of consumer discrimination based on the race of salespersons (Doleac and Stein 2013), the race of front-end employees (Holzer and Ihlanfeldt 1998), and the race of the athletes (Kanazawa and Funk 2001).

ence engulfed the national dialogue, to examine the economic costs of discrimination.[5] Also related is Luca, Pronkina and Rossi (2022) who find declines in Airbnb ratings for listings associated with Asian named hosts relative to other hosts following the covid announcement.

The remainder of this paper is organized as follows. Section 1 describes the observational setting and data. Sections 2 and 3 outline the baseline and extended frameworks and provide the central results. Section 4 discusses belief accuracy. Section 5 concludes with a brief discussion of the results.

1.2 Observational Setting

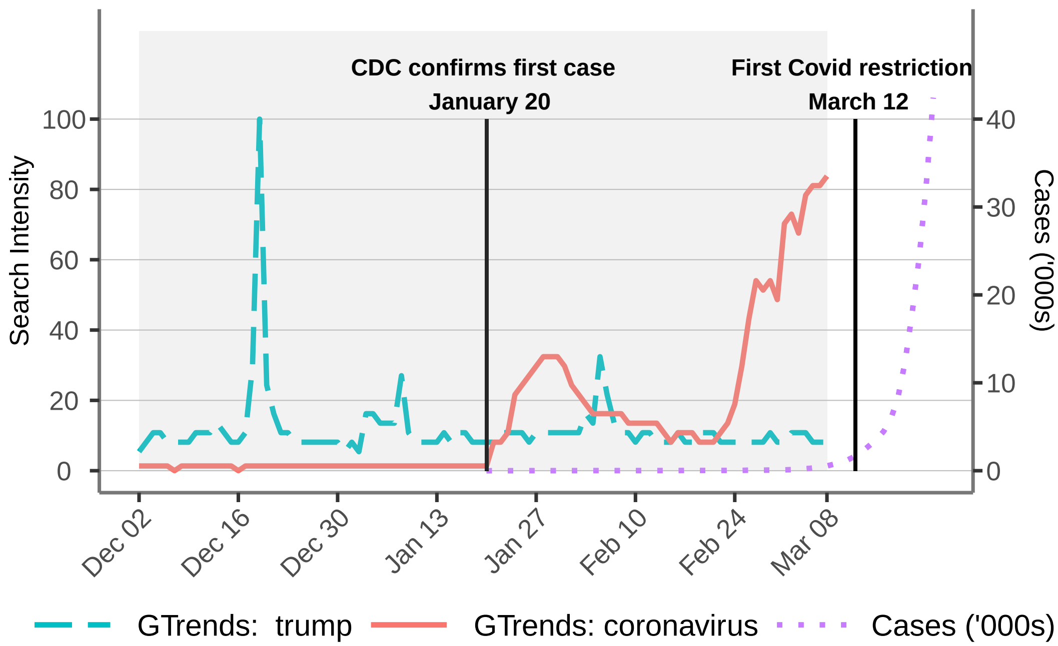
On January 20, 2020, the Center for Disease Control (CDC) confirmed the first case of covid in the US (Center for Disease Control 2020).[6] The following seven week period ending on March 8 exhibits several important features that allow us to test for alternate motives in demand changes for restaurants.

The first feature is that prior to March 8, no policies had yet been announced that would restrict restaurant choice. Consumers were therefore free to make unrestricted market choices but with new information on the presence of the virus. Second, the virus was a major topic of public concern. As shown in Figure I, "coronavirus" was the most searched word on Google during this period, surpassing "trump", who was undergoing an impeachment trial (Google 2022).

[5]While both studies discuss the different motives underlying consumer discrimination, they do not directly test for them using consumer outcomes, instead focusing on economic consequences or relying on survey data conducted in later phases of the pandemic.

[6]The first positive test was confirmed on Monday, January 20 and the announcement was made the following day. I consider the week starting on January 20th as the first treatment week.

Figure 1.1: Google Trends Search Intensity, December 2019 - March 2020



Note: Observational period in shaded region. The figure displays the Google Trends search intensity of "trump" and "coronavirus" between December 2019 - March 2020 (left axis) and the number of reported covid cases in the US, in thousands (right axis). The search intensity is normalized to the time period and scaled to a range of 0 to 100. The peak search for "trump" occurred on December 19, 2019, the day after Trump's impeachment by the House of Representatives. The maximum search intensity for "coronavirus", occurring on March 8, is approximately 85% of that of "trump" on December 19. Coronaviruses are a group of viruses that cause respiratory illnesses and are generally transmitted through close personal contact. Despite the official name of "COVID-19" announced on February 11, "coronavirus" was the top searched topic in the seven week period after the CDC announcement.

Third, this period lacks the influence of high-profile political figures using terms that pointedly associate the virus to Chinese people. For example, Donald Trump's use of "Chinese virus," "Wuhan virus," or "Kung Flu" did not first occur until March 16, and the first instance of any US politician to use such terms was on March 6 (Cao, Lindo and Zhong 2022; Huang et al. 2023). This suggests that the growing prominence of covid in news or on social media was a result of updates reported by health organizations that generated organic public dialogue rather than instigated by political figures. Nevertheless, it is crucial to acknowledge the rise in anti-Asian hate incidents during this time, indicating potential for such sentiment among consumers. In Appendix table A.1.1, I list key events during and immediately after the observational period.

Finally, covid cases remained low in the US during this time whereas cases in China had already spiked, as shown in Appendix table A.1.2, which fortifies the association of the virus with China. Cases were also rising in other parts of Asia and Europe, but not in Latin America, which provides cross-cuisine variation in perceived health risks of foreign restaurants.

Restaurants offer two important sources of variation that, during this unique period, allow us to directly map covid-induced beliefs and covid-induced anti-Chinese sentiment to consumer outcomes. First, restaurants provide varying degrees of service, so the product purchased may include only a good (food) in the case of take-outs or both a good *and* an in-person—and time-requiring—service in the case of dine-ins. Since the covid announcement had the immediate consequence of elevating the perceived health risks of in-person interactions, variation in services generates variation in the perceived health risks of restaurant visits, which can be used to detect a belief-based response. On the other hand, variation in the ethnic composition of the restaurant's clientele generates variation in potential for anti-Chinese sentiment.

In light of these defining characteristics, this brief window is uniquely suited to study a natural consumer reaction to an event that distinctly associates ethnically differentiated products and services with elevated health risks due to the geographic origin of a novel virus.

1.2.1 SafeGraph Data

The SafeGraph data comes in two sets called Places and Patterns for the universe of full-service and limited-service establishments. Places contains descriptive variables on each restaurant, including its name, location, open and close dates, and a description of the restaurant, with labels such as "Chinese", "Dinner", "Food Truck", and "Late Night". Using these labels, I classify restaurants into five broad cuisines: Chinese, Asian (non-Chinese), American, Latin, and European. I keep establishments that can be considered conventional restaurants and remove those with an ambiguous cuisine group, but otherwise classify restaurants as American if they do not suggest another major group. This restriction, which removes 22% of restaurants, is adopted because a strong and clear association between cuisine type and ethnicity is necessary to test for alternate motives.

The Patterns data provides a weekly panel of visit counts to each restaurant. I obtain data for the time periods December 3, 2018 - March 10, 2019 and December 2, 2019 - March 8, 2020.[7] For the remainder of the paper, "2019" will refer to the former period and "2020" will refer to the latter period. This way, there are three pre-treatment periods and one post-treatment period. In addition to visit counts, this data set also provides a seven-bucket distribution of dwell times and a distribution of visitor counts by home census tract.[8] I use the 2020 pre-treatment period to construct a continuous measure of

[7]Thanksgiving occurred one week prior in 2019 but two weeks prior in 2018, making the last two weeks of November less comparable between the two years.

[8]Throughout the analysis, I use state-weighted total visits provided by SafeGraph. This measure reweighs raw (GPS detected) visits based on the number of detected devices at the census block group level and SafeGraph's state sampling rate. More information on reweighting can be found in Chong (2021). For each observation, duration is provided as a count of visits that last for less than 5 minutes, 5-10 minutes, 11-20 minutes, 21-60 minutes, 61-120 minutes, 121-240 minutes, and 241+ minutes. Visitor origin is provided similarly for each home census tract. SafeGraph's data is highly representative of the overall population, as shown by Squire (2019).

Table 1.1: Summary Statistics of Restaurant Visits

Cuisine	Dine Share	Mean Visits 2019	Mean Visits 2020	Close Rate (%) Post-2020	Restaurants	Obs.
American	0-20	88.3	111.2	0.37	32,537	910,189
	20-40	78.6	96.8	0.54	97,207	2,718,777
	40-60	71.4	87.3	0.66	57,230	1,600,289
	60-100	103.4	127.0	0.30	31,932	893,311
Asian	0-20	34.4	38.2	1.00	1,596	44,580
	20-40	48.6	53.4	0.89	9,856	275,430
	40-60	53.3	59.9	0.69	14,101	394,297
	60-100	60.7	72.9	0.35	6,212	173,744
Chinese	0-20	40.1	42.0	1.27	1,962	54,812
	20-40	53.0	58.3	0.72	11,277	315,281
	40-60	51.7	61.0	0.44	8,188	228,986
	60-100	59.3	70.5	0.30	2,688	75,202
European	0-20	35.2	40.0	1.36	1,541	43,019
	20-40	49.3	55.0	1.04	8,979	250,913
	40-60	58.2	68.2	0.55	12,086	338,023
	60-100	102.3	125.8	0.29	4,887	136,685
Latin	0-20	63.6	78.1	0.72	8,519	238,130
	20-40	66.4	75.4	0.61	31,527	881,725
	40-60	64.2	73.8	0.69	20,528	574,034
	60-100	84.3	105.3	0.27	10,605	296,700
Sample	All	73.3	88.5	0.56	373,458	10,444,127

Note: The table provides summary statistics of the SafeGraph data, displaying for each cuisine by dine group the average weekly recorded raw visits, the percent of restaurants that closed between January 20 and March 8, 2020, counts of restaurants, and total restaurant-week observations. The close rate is calculated using the "closed on" date variable in the data. The higher visit count in 2020 is due to an increase in the pool of smartphone devices that SafeGraph tracked over the year, which is accounted for in the state-weighted visits and in all triple difference models.

service type: *dine-share* is defined as the share of visits lasting 21-120 minutes.[9] The Chinese-share variable is constructed similarly, using population data on the home census tracts of visitors in the 2020 pre-treatment period.[10] Finally, I keep restaurants with at least 40 raw visits in each of the three pre-treatment periods.[11]

The resulting sample includes 373,458 limited or full service establishments that were

[9] Visits lasting over two hours are likely employees.

[10] The home tract is based on the typical location of the phone during nighttime hours (6pm - 7am). I use the 2019 ACS 5-year census tract estimates of the Chinese and total population to calculate the Chinese share of visitors for each restaurant-week-tract over the entire seven week pre-treatment period.

[11] Restaurants with few detected visits are unlikely to accurately measure dine share and Chinese share or may be less precisely comparable across the three pre-treatment periods. This restriction removes 14% of restaurants before cuisine assignment.

open on January 20 and 1) are conventional restaurants, 2) are unambiguously classified into one of the five large cuisine groups, and 3) met the minimum 40 raw visits threshold in all three pre-treatment periods. This represents approximately 73% of all limited and full service restaurants reported by the US Census in 2019 (US Census 2019*b*).[12]

Table 1.1 provides summary statistics of the final restaurants sample, including mean visits, the percent of restaurants that closed post-announcement, counts of restaurants, and total observations. The low close rate indicates that restaurants did not preemptively close in anticipation of reduced demand or due to the virus, and shows that Chinese restaurants had similar close rates to other cuisines.

Control variables include cases in the last two weeks, the 2016 share of votes for the Democratic party, population, population density, and the share of the population of Chinese ancestry.[13]

[12]This also represents 67% of the unrestricted sample in the SafeGraph data, which includes limited or full service establishments that were open on January 20 but without the remaining three restrictions. I replicate the results using the unrestricted sample of 556,723 restaurants, and separately using raw visits rather than re-weighted visits and the results, discussed in Appendix subsection 1.4.2, remain broadly similar.

[13]Data on reported covid cases is provided by New York Times (2022) and includes the cumulative reported covid cases for each county-day. I take one plus the count of new cases in the prior two weeks, in logs, as the primary control variable. This data set does not distinguish the five counties of New York. I group these counties into one county representing New York City throughout the analysis. I obtain the county share of 2016 Democratic party votes from MIT Election Data and Science via Harvard Dataverse (2023), and this variable is centered at 0.5. I obtain data on the 2018 and 2019 county populations and population density from the US Census and ACS 2019 5-Year Sample (US Census 2019*a*; Ruggles et al. 2023). I use the 2015-2019 American Community Survey to construct the county level Chinese ancestry share of the population, which is limited to the population aged 16 and over.

1.3 Baseline Framework

1.3.1 Model

Belief-based discrimination due to the covid announcement is defined as differential changes in demand for cuisine types due to differences in perceived risk of infection. This may arise due to cognitive biases such as representative heuristics or representative signal distortion. As the consumer expects a longer stay from a visit, a belief-based discriminator assigns an even greater risk of infection if it were a Chinese restaurant due to additional concern over air-borne transmission. If they expect to spend little time, then concerns are limited to food-borne transmission. Conditional on the consumer's pre-covid cuisine preferences, a belief-based discriminator has relatively lower expected utility from a potential visit to a Chinese restaurants as the perceived risk of infection rises.

Taste-based discrimination due to the covid announcement is defined as differential changes in demand for cuisine types due to the associated ethnicity, independently of perceived risk of infection. This is distinct from the consumer's pre-covid cuisine preferences. Instead, this represents new anti-Chinese sentiment induced by the covid announcement that reduces the consumer's utility from a restaurant visit.[14] This may come in the form of animus, such as reduced willingness to do business with any Chinese restaurants, or in the form of a reluctance to interact with Chinese people. The taste-based discriminator

[14]Generally, cuisine preferences change gradually and are unlikely to have changed meaningfully in the absence of the covid announcement within a seven week period. As such, sharp declines in restaurant demand immediately following the covid announcement can be explained by concerns over virus transmission or by anti-Chinese sentiment, but not due to declines in the utility that consumers typically derive from a particular cuisine.

therefore avoids all Chinese restaurants or avoids any restaurant with a greater chance of interacting with Chinese people, including at non-Chinese restaurants.

The decline in demand for non-Chinese take-outs will represent aversion to *restaurant-cooked food* caused by the announcement due to concerns over food-borne transmission, which I denote F. While it was eventually determined that covid was not significantly transmitted through food, this was not known during the observational period, meaning this effect can be interpreted as *ex-post inaccurate beliefs* about the health risks of eating restaurant-cooked food. Second is the *additional* decline in demand for dine-ins, which represents aversion to *sitting in a restaurant* due to concerns over air-borne transmission and is denoted A. Since it was eventually determined that covid can be transmitted this way, this parameter constitutes *ex-post accurate beliefs* about the health risks of sitting in a restaurant for a prolonged period of time. As many restaurants offer a mix of take-out and dine-in options, A depends on the restaurant's service type, which is proxied with the continuous measure of dine-share, $d \in [0, 1]$.

I define an additional decline in demand for Chinese restaurants as discrimination of some form. Discrimination against take-outs can be explained by a combination of 1) *ex-post inaccurate beliefs* that Chinese food was more likely to transmit the virus, denoted F^C, and 2) a taste-based decline in the willingness to do business with Chinese restaurants, denoted δ_0^C.[15] Although the data does not allow me to distinguish between these two, together they make up a third parameter of interest. Finally, the additional covid-induced aversion to *sitting in a Chinese restaurant* for a prolonged period of time can

[15]Past studies have found evidence of discrimination in the form of lower willingness to do business. Dulleck, Fooken and He (2020) find experimental evidence of preference-based discrimination arising as a reduction in the willingness to compensate labor. Hedegaard and Tyran (2018) directly test for the "price of prejudice" in a field experiment that manipulates the monetary cost of discriminatory preferences.

be explained by 1) concerns over air-borne transmission, denoted A^C, and 2) reluctance to share space with Chinese people, denoted δ_d^C. Expected utility of a restaurant visit after the covid announcement is a function of utility prior to the announcement and the outlined response parameters, and is captured in equations 1.1a and 1.1b for non-Chinese (N) and Chinese (C) restaurants, respectively. The corresponding model of discrimination in response to the covid announcement translates to a difference-in-differences in expected utility.

$$E(U)^N_{\text{post}} = U^N_{\text{pre}} + F + A \times d \tag{1.1a}$$

$$E(U)^C_{\text{post}} = U^C_{\text{pre}} + F + (F^C + \delta_0^C) + (A + A^C + \delta_d^C) \times d \tag{1.1b}$$

$$D^C \equiv \Delta E(U)^{C-N} = (F^C + \delta_0^C) + (A^C + \delta_d^C) \times d \tag{1.2}$$

Equation 1.2 displays the baseline linear model of discrimination with intercept $(F^C + \delta_0^C)$ and slope $(A^C + \delta_d^C)$, which provide two estimated coefficients. To identify belief-based discrimination, it is necessary to rule out both taste-based parameters, δ_0^C and δ_d^C. The identification challenge is that neither parameter can be separated from their belief-based counterparts, F^C and A^C. Fortunately, it is unlikely for consumers to have believed that Chinese food became safer relative to other restaurant food, meaning $F^C \leq 0$. This allows for pinning down δ_0^C to 0 by finding no discriminatory response toward Chinese take-outs. Specifically, we can rule out a taste-based reduced willingness to do business by finding a null value for the combined parameter $F^C + \delta_0^C$ because F^C is non-positive.

The second parameter, δ_d^C, remains conjoined with beliefs over air-borne risks and cannot be ruled out in this baseline framework. I address this in the following section, which extends this model by allowing for greater heterogeneity in belief-based and taste-based responses. However, this baseline model, while limited, serves as a simple framework that can provide suggestive evidence of which motive is likely driving the overall consumer response.

1.3.2 Empirical Strategy

The triple-differences estimation strategy will compare, in the first two differences, visits to Chinese restaurants after the covid announcement relative to before and in the current year relative to the prior year. In the third difference, it compares the consumer response across cuisines. The DDD framework is preferred over the standard difference-in-differences (DID) framework in order to account for seasonal trends in consumer preferences, such as increased demand for Chinese food on Christmas.[16]

We can interpret the DDD strategy as follows: given that a conventional DID model using other cuisine types as the control group may not account for seasonal variation in consumer preferences, such differences are netted out by first controlling for each cuisine's demand in the prior year, when there was no covid. This is equivalent to first adopting a conventional DID framework, separately for each cuisine, that interacts indicators for post-treatment and year, and includes control variables. Then, in the third difference, cuisine type is interacted with all right-hand-side variables, which will result in a DDD model that has removed cuisine-specific seasonal variation. As such, the identifying assumption in the DDD is that seasonal variation may differ across cuisines but is similar

[16]Christmas and New Year each occurred within the same observational week in 2020 relative to 2019. The Chinese New Year occurred on different weeks in 2020 (February 5) relative to 2019 (January 25), but occurred within the post-treatment period in both years.

within cuisines between 2020 and 2019 (Gruber 1994; Olden and Møen 2022).

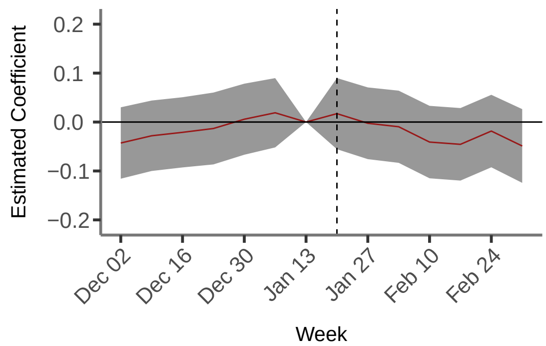
$$\begin{aligned} log(visits)_{r,i,j,k,c} = \quad & \tilde{\beta}_1 Post_i \times 2020_j \\ & + \tilde{\beta}_2 Post_i \times 2020_j \times Chinese_k \\ & + \mathbf{W}_{i,j,k}\mu + \mathbf{X}_{i,j,c}\gamma + \eta_{k,c} + \epsilon_{r,i,j,k,c} \end{aligned} \tag{1.3}$$

I estimate equation 1.3 with county-cuisine fixed effects, separately for six partitions of restaurant dine-share. The subscripts index restaurant r, week i, year j, cuisine k, and county c, and visits are measured in logs. Each estimated coefficient can be approximated as a percentage change in visits due to a one unit increase in the explanatory variable. The model includes the remaining full set of DDD interactions in $\mathbf{W}$ and control variables in $\mathbf{X}$.[17] Finally, the model includes county-cuisine fixed effects to ensure it has the interpretation of a difference of DID models. These fixed effects capture time-invariant characteristics common to each cuisine within a county. Under the DDD identifying assumption of a constant difference in parallel trends, $\tilde{\beta}_1$ represents the impact of the covid announcement on demand for the control group and $\tilde{\beta}_2$ represents the additional impact of the covid announcement on demand for Chinese restaurants.

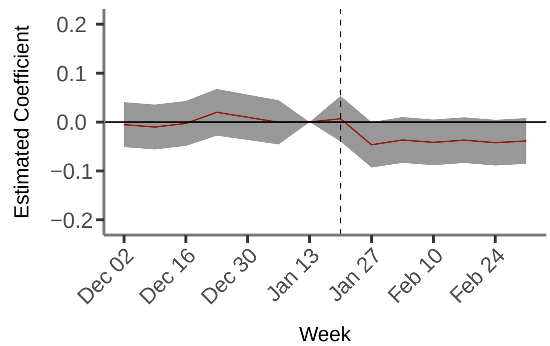
[17] The remaining set of DDD variables include $Post$, 2020, $Post \times Chinese$, and $2020 \times Chinese$. The uninteracted indicator, $Chinese$, is absorbed by the fixed effects. Other than covid cases, which vary by county-week and are imputed 0 in all pre-treatment observations, the remaining controls vary by county or by county-year. These controls enter the model as interactions with $Post \times 2020$ and additionally as interactions with $Post \times 2020 \times Chinese$ to control for differential covid-induced responses to local characteristics for Chinese and non-Chinese restaurants.

Figure 1.2: Chinese Restaurant Visits, by Service Type (Reference = other cuisines)

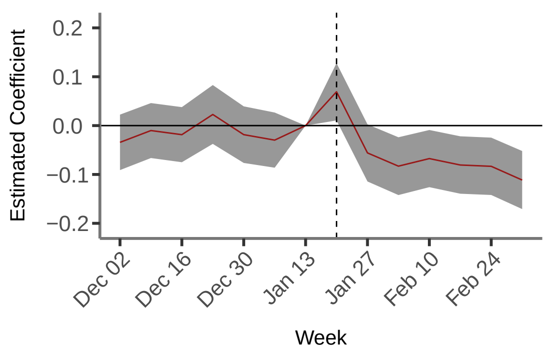
(a) Dine Share: 0% - 20%



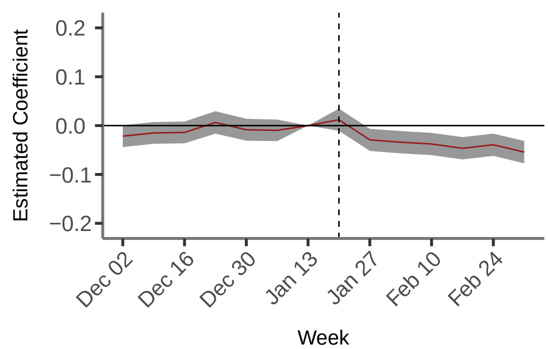
(b) Dine Share: 40% - 50%



(c) Dine Share: 60% - 100%



(d) Full Sample



Note: The figures provide event studies associated with equation 1.3 that represents estimation of the DDD model, separately for three partitions of dine-share: 0-20% (a), 40-50% (b), and 60-100% (c), for the full sample (d).

Figure 1.2 displays event studies associated with equation 1.3 for restaurants falling in three dine-share groups: mostly take-out restaurants that have dine share less than 20% (a), those with a mix of service types with dine share 40%-50% (b), and mostly dine-in restaurants with dine share over 60% (c). The last subfigure (d) displays the event study for the full sample, representing the approximately average dine-share restaurant. The reference group includes all remaining cuisines, with the 95% confidence interval in gray. The figures illustrate that the pre-treatment trends do not violate the DDD identifying assumption. The figures also do not indicate a relative decline in demand for Chinese take-outs, but noticeable decline in demand for Chinese dine-ins.

The estimated coefficients of equation 1.3 are provided in Appendix table A.1.3 and indicate that customers avoided Chinese dine-ins the most, with the coefficient on the triple interaction increasing in dine-share. Additionally, no consumer discrimination is detected against Chinese take-outs, as shown in the first two columns.

1.3.3 Parameter Estimates

To estimate the parameters of interest in the baseline framework, I fully interact dine-share with the three DDD indicator variables.[18] I also interact dine-share with recent covid cases to detect the responsiveness of restaurant-goers to prolonged in-person interactions after a new case is reported locally. The resulting model, shown in equation 1.4, has the benefit of capturing all discernible parameters of interest in the baseline framework.

$$\begin{aligned}
log(visits)_{r,i,j,k,c} = \quad & \beta_1 Post_i \times 2020_j \\
& + \beta_2 Post_i \times 2020_j \times DineShare_r \\
& + \beta_3 Post_i \times 2020_j \times Chinese_k \\
& + \beta_4 Post_i \times 2020_j \times DineShare_r \times Chinese_k \\
& + \mathbf{W}_{r,i,j,k}\mu + \mathbf{X}_{i,j,c}\gamma + \eta_{k,c} + \epsilon_{r,i,j,k,c} \qquad (1.4)
\end{aligned}$$

As before, each coefficient of an explanatory variable that includes $Post \times 2020$ represents an announcement-induced response. The parameters of interest from the

[18] This produces 15 explanatory variables that are displayed, stored in $\mathbf{W}$, or absorbed by fixed effects.

baseline framework correspond to the coefficients of interests as $F^C + \delta_0^C = \beta_3$ and $A^C + \delta_d^C = \beta_4$. Since few restaurants are fully take-out or dine-in, the preferred specification used throughout this paper demeans the dine-share (mean of 0.40) to estimate effects for the average-service restaurant. However, only the specification where dine-share is not demeaned can provide an estimate of $(F^C + \delta_0^C)$.

Table 1.2 summarizes the main results. The estimated model uses all four non-Chinese cuisine types as the reference group, and indicates an 11.9 percentage point decline in visits to take-outs (F), with just a 1.9 percentage point *additional* decline for dine-ins (A).[19] There is no detected discrimination against Chinese take-outs, captured by $F^C + \delta_0^C = \beta_3$, whereas a significant discrimination effect of 8.1 percentage points is detected for Chinese dine-ins, captured by $A^C + \delta_d^C = \beta_4$. Under the baseline framework, the results indicate that consumers did not develop a reduced willingness to do business with Chinese restaurants.

In Appendix table A.1.4, I provide the full set of results corresponding to equation 1.4. Columns 1-4 provide estimated coefficients that include all remaining cuisines as the reference group to Chinese restaurants. In columns 5-8, American restaurants are the reference group to each foreign cuisine. The coefficient on $Post \times 2020$ is the benchmark response parameter associated with take-outs of the reference group and is equivalent to $F = \beta_1$. When this is further interacted with the dine-share, the resulting coefficient captures the consumer response to dine-in restaurants relative to take-outs among the reference group, and estimates $A = \beta_2$. Each estimated model shows that consumers broadly avoided dine-ins only slightly more than take-outs.

The interaction of post-treatment with cuisine, $Post \times 2020 \times Cuisine$, provides an

[19] Estimates in text are adjusted from the log approximation, $exp(-0.127) - 1 = -0.119$.

Table 1.2: Summary of Baseline Framework Parameters

Restaurant Group	**Theoretical Parameter**	**Motive Type**	**Estimated Parameter**	**Coefficient Estimate**
Other Take-Outs	F	Beliefs	β_1	-0.127*** (0.003)
Other Dine-Ins	A	Beliefs	β_2	-0.019*** (0.007)
Chinese Take-Outs	$F^C + \delta_0^C$	Beliefs & Tastes	β_3	0.010 (0.013)
Chinese Dine-Ins	$A^C + \delta_d^C$	Beliefs & Tastes	β_4	-0.085*** (0.029)

* p < 0.1, ** p < 0.05, *** p < 0.01
Note: The table provides estimates of the baseline framework of the consumer response to the covid announcement. F is aversion to restaurant-cooked food. A is additional aversion to sitting in a restaurant for an extended period of time. F^C is additional taste-based aversion to Chinese food. δ_0^C is a taste-based decline in willingness to do business with Chinese restaurants. A^C is additional belief-based aversion to sitting in a Chinese restaurant for an extended period of time. δ_d^C is additional taste-based aversion to sharing space with Chinese people.

estimate of the response to the cuisine's take-outs *relative* to the reference group's take-outs. This estimate, in the model where dine-share is not demeaned, is most likely to detect taste-based discrimination because these Chinese restaurants are perceived to be the least risky. Column 3 indicates no additional aversion to Chinese take-outs, whereas the demeaned model in column 4 displays an approximately 2.3 percentage point discrimination effect on the *average* Chinese restaurant in terms of service type.

Finally, the coefficient on the full interaction, $Post \times 2020 \times Cuisine \times DineShare$, is most likely to capture belief-based discrimination because these restaurants are perceived to be riskiest. The estimate indicates a strong aversion to Chinese dine-ins. A discrimination effect is also detected for Asian restaurants, indicating belief spillover into neighboring cuisines to Chinese restaurants.

As column 2 controls for covid cases without interacting with dine-share, a comparison with the preferred estimated model shows that local cases explain only 15-20% of health-related concerns at Chinese dine-ins. The interaction $Cases \times Cuisine \times DineShare$ supports this: relative to American dine-ins, consumers avoided Chinese dine-ins much more following recently reported covid cases, and this is not limited to Chinese restaurants. In addition to China, other Asian and European countries had rising cases by March 8 and consumers consequently avoided these cuisine's restaurants the most after a new case is reported. Latin American countries had few cases during this time and the associated restaurants were not avoided. These results suggest a strong local salience effect: when more cases are reported locally, consumers avoided foreign restaurants that were associated with rising cases abroad more than they avoided domestic restaurants.

1.4 Extended Framework

While the baseline model tests for taste-based discrimination in the form of a reduced willingness to do business with Chinese restaurants, it cannot test for taste-based discrimination in the form of aversion to prolonged interactions with Chinese people. To address this, it would be useful to know how consumer demand responded to restaurants that typically have a large Chinese clientele. It would be particularly useful to know how Chinese and non-Chinese customers responded differently to such restaurants. If non-Chinese customers, who have potential for anti-Chinese sentiment, responded negatively or worse than Chinese customers, it may suggest some degree of taste-based aversion. On the other hand, if *only* Chinese customers, who can only be belief-driven, responded negatively, it would indicate their heightened concerns over interacting with other Chinese people. This latter finding would also indicate an absence of a taste-based response among non-Chinese customers. This exercise on its own would not rule out aversion to prolonged interactions, δ_d^C in the baseline framework, because it does not isolate the response specifically to Chinese dine-ins, but it would provide us with strong evidence on the likelier motive.

$$\begin{aligned} log(1+visits)_{r,i,j,c} = \quad & \tilde{\delta}_1 Post_i \times 2020_j \\ & + \tilde{\delta}_2 Post_i \times 2020_j \times ChineseShare_r \\ & + \mathbf{W}_{r,i,j}\mu + \mathbf{X}_{i,j,c}\gamma + \eta_r + \epsilon_{r,i,j,c} \end{aligned} \tag{1.5}$$

I adapt the DDD model from equation 1.3 by replacing the indicator for cuisine with the continuous measure of Chinese share.[20] The resulting equation 1.5 is estimated separately for each cuisine. As Chinese restaurants and customers are concentrated in relatively few counties, I use restaurant fixed effects instead of county fixed effects. $\tilde{\delta}_2$ represents the announcement-induced response to restaurants that, just before the covid announcement, had a one standard deviation higher Chinese share of customers. This can be interpreted as a covid-induced aversion to interactions with Chinese people and tests directly for taste-based aversion under the identification strategy.

[20] I also transform the outcome variable by adding 1 before taking the log due to approximately 10% of observations having no Chinese customers, as all visitors originated in fully non-Chinese census tracts. Additionally, approximately 3% of observations do not provide a home census tract and these are omitted (4% among Chinese restaurants). Chen and Roth (2023) emphasize that the transformation of the outcome variable Y into log(1 + Y) cannot provide a percentage interpretation of the average treatment effect. In this setting, there is less concern with the magnitude of the estimated coefficient of interest and more concern with the sign. However, I find nearly identical estimates when observations with 0 visits by Chinese customers are omitted and I simply use log(Y). Following the authors' suggestion, I also estimate the model using the transformation $Y/X - 1$, where X is the restaurant's average weekly visits in the pre-treatment period (see Appendix subsection 1.4.2). This alternate transformation of the outcome can be interpreted as a precise, rather than approximate, percentage change measure. Although point estimates differ under this alternate specification, the signs and the overall interpretation are unchanged. However, for consistency between the two customer types and with the rest of the analysis, and to retain as much of the original sample, I use log(1 + Y) as the preferred transformation.

Table 1.3: Response to Chinese Share of Customers

	Dependent Variable: Log(1 + Visits), by Customer Type							
	Chinese Restaurants				Non-Chinese Restaurants			
Customer Type:	All	Chinese	Non-Chinese	Difference	All	Chinese	Non-Chinese	Difference
Post x 2020 x Chinese Share	0.029***	-0.020***	0.033***	-0.054***	0.060***	-0.014***	0.061***	-0.075***
	(0.003)	(0.006)	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)
Post x 2020	-0.158***	-0.087***	-0.152***	0.065***	-0.129***	-0.080***	-0.125***	0.045***
	(0.002)	(0.004)	(0.002)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Chinese Share	-0.019***	-0.021***	-0.019***	-0.002	-0.026***	-0.026***	-0.026***	0.000
	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Cases	-0.111***	-0.099***	-0.106***	0.007	-0.066***	-0.059***	-0.066***	0.006***
	(0.006)	(0.009)	(0.006)	(0.008)	(0.001)	(0.002)	(0.001)	(0.002)
Total Obs.	674,253	646,079	646,079	646,079	9,769,128	9,442,887	9,442,887	9,442,887
R Squared	0.830	0.841	0.834	0.821	0.863	0.813	0.868	0.771
FE: Restaurant	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.5, representing the response of Chinese and non-Chinese customers to the pre-covid restaurant Chinese share. The dependent variables in columns 1-3 and 5-7 are measured as log(1 plus customer visits). In columns 4 and 8, the dependent variable is the difference, log(1 plus Chinese visits) - log(1 plus non-Chinese visits). Visit counts for customer types are calculated by allocating total visits according to the 2019 ACS 5-year sample of the Chinese and total population and the home census tract of visitors. Chinese share is standardized. The full set of interactions of Post, Year, and Chinese share are included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and Post x Year x Chinese share to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the restaurant level.

Table 1.3 displays the results and Appendix figure A.2.1 displays the associated event studies. The table indicates that Chinese customers avoided restaurants that typically have more Chinese customers, whereas non-Chinese customers responded positively, a finding that holds at both Chinese and non-Chinese restaurants. On the other hand, when a new covid case is reported nearby, both customer types avoided high Chinese-share restaurants, as indicated by the interaction $Cases \times ChineseShare$.

These results are important for three reasons. First, they offer strong evidence that consumers did not develop taste-based aversion to interacting with Chinese people because non-Chinese customers did not reduce demand for restaurants that will have likely led to interactions with Chinese customers. They also indicate that Chinese customers developed concerns that other Chinese people were more likely to have the virus. Second, the similar response to recent cases indicates that both customer types avoided Chinese customers *when* there was a case reported, but non-Chinese customers did not avoid Chinese customers when there wasn't a recent case reported. Under the identification strategy, this behavior is consistent with the *belief* that Chinese people are more likely to have the virus because aversion is detected only when there was an elevated perceived risk of infection. Finally, these results address concerns that non-Chinese customers may be underestimating the presence of Chinese customers at non-Chinese restaurants, or that the Chinese share is simply too low to matter; they would have accurately or overestimated the presence of Chinese customers at Chinese restaurants, yet they were not deterred from visiting such restaurants following the covid announcement.

For greater assurance, I probe for distributional differences in the Chinese share across cuisine and service types. Appendix tables A.1.5a and A.1.5b display each cuisine's restaurant counts and relative representation in different quantiles of the Chinese share distribution, separately for take-outs and dine-ins. They broadly show that the distribution is highly skewed to the right for both service types and all cuisines, with only slightly greater skew among Chinese restaurants.

1.4.1 Model

To reconcile these findings with the baseline framework, I introduce two new sets of parameters, which differ by cuisine. A_c^N and A_c^C represent beliefs that Chinese people are more likely to carry the virus. $\delta_{d,c}^N$ and $\delta_{d,c}^C$ represent changes in the willingness to share space with Chinese people, independently of beliefs. The new taste-based parameter at Chinese restaurants, $\delta_{d,c}^C$, absorbs the corresponding parameter from the baseline framework, δ_d^C, and depends on both the restaurant's dine-share, d, and its Chinese-share, $c \in [0, 1]$. The parameters enter additively into the baseline framework, resulting in the following extended theoretical framework:

$$E(U)_{\text{post}}^N = U_{\text{pre}}^N + F + A \times d + (A_c^N + \delta_{d,c}^N) \times d \times c \tag{1.6a}$$

$$E(U)_{\text{post}}^C = U_{\text{pre}}^C + F + F^C + \delta_0^C + (A + A^C) \times d + (A_c^C + \delta_{d,c}^C) \times d \times c \tag{1.6b}$$

Equations 1.6a and 1.6b fully allocate anti-Chinese sentiment to a decline in the willingness to do business with a Chinese restaurant or to a decline in the willingness to share space with Chinese people. Although the new parameters, $\delta_{d,c}^N$ and $\delta_{d,c}^C$, remain conjoined with beliefs, there is a special case in which $\delta_{d,c}^C$ is confined to 0 such that the combined coefficient $(A_c^C + \delta_{d,c}^C)$ identifies solely a belief-based response. Specifically, it is likely that the typical customer at Chinese restaurants exhibited *no greater* anti-Chinese sentiment relative to the typical customer at non-Chinese restaurants because, prior to the covid announcement, they self-select into their preferred cuisine type. This is simply expressed by $\delta^N_{d,c} \leq \delta^C_{d,c} \leq 0$. As such, finding a null value for the combined parameter $(A^N_c + \delta^N_{d,c})$ at non-Chinese restaurants will pin both $\delta^N_{d,c}$ and $\delta^C_{d,c}$ to 0, so that a negative response among Chinese dine-ins $(A^C_c + \delta^C_{d,c})$ is reduced to beliefs.[21]

Estimation of the extended model, shown in equation 1.7, retains the structure of the baseline model in equation 1.4 and replaces the indicator for cuisine with the continuous measure of Chinese share. The model is estimated separately for each cuisine and includes restaurant fixed effects.

$$\begin{aligned} log(visits)_{r,i,j,c} = \quad & \delta_1 Post_i \times 2020_j \\ & + \delta_2 Post_i \times 2020_j \times DineShare_r \\ & + \delta_3 Post_i \times 2020_j \times ChineseShare_r \\ & + \delta_4 Post_i \times 2020_j \times DineShare_r \times ChineseShare_r \\ & + \mathbf{W}_{r,i,j}\mu + \mathbf{X}_{i,j,c}\gamma + \eta_r + \epsilon_{r,i,j,c} \end{aligned} \tag{1.7}$$

The coefficient of interest, δ_4, corresponds to the theoretical framework as $A^N_c + \delta^N_{d,c}$ for non-Chinese restaurants and $A^C_c + \delta^C_{d,c}$ for Chinese restaurants. The results are provided in the first row of Appendix table A.1.6 and in the accompanying event studies in Appendix figure A.2.2. The results show that consumers reduced demand for Chinese dine-ins with more Chinese customers, but did not avoid other dine-ins with more Chinese customers.

[21] A large body of evidence from a wide range of literatures justifies $\delta^N_{d,c} \leq \delta^C_{d,c}$, which is fundamentally rooted in Allport (1954)'s Contact Theory. Pettigrew and Tropp (2006)'s meta analysis of 515 studies from psychology, sociology, political science, and education finds that intergroup contact broadly reduces prejudice. Studies from the economics literature, such Carrell, Hoekstra and West (2019), Marmaros and Sacerdote (2006), and Corno, La Ferrara and Burns (2022), further provide evidence that greater contact increases willingness for more contact. It is also noted that the triple difference framework first benchmarks the taste-based willingness to interact with Chinese customers within cuisine in the prior year before comparing across cuisines, thus differencing out baseline levels.

The four-way interaction is important for two reasons. First, the coefficients on the response to dine-share and Chinese-share provide insight into the likely motives at each cuisine type. The coefficient on the response to dine-share, $Post \times 2020 \times DineShare$ is large and negative, whereas the coefficient on the response to Chinese-share, $Post \times 2020 \times ChineseShare$, is positive. Second, customers of all backgrounds likely have own-ethnicity preferences in their restaurant choice, which may bias measurement of Chinese share down for Chinese restaurants and up for the remaining cuisines. However, the DDD specification safeguards from this bias by controlling for within-cuisine demand in the prior year.

I next ensure that the overall consumer response is not driven predominantly by Chinese customers. I replicate the results of the baseline framework, separately for Chinese and non-Chinese customers, and show them in Appendix table A.1.7 with county-cuisine fixed effects and in Appendix table A.1.8 with restaurant fixed effects. Together, the two tables indicate that both customer types avoided Chinese dine-ins, with Chinese customers doing so at least as much as non-Chinese customers. The specification using restaurant fixed effects provides estimates that are similar to the baseline specification that uses county x cuisine fixed effects, but with significantly greater explanatory power. This is noteworthy because the granularity of restaurant fixed effects control for many characteristics likely to influence a potential customer's choice that county x cuisine fixed effects do not, such as the restaurant's reputation, menu prices, location within the county, and other demographic characteristics of its customer base.

Combined with the results from the baseline framework, these findings are consistent with a belief-based motive under the identification strategy. They show that in response to the covid announcement, consumers did not avoid Chinese restaurants with the least perceived risk (take-outs), but did avoid those with the most perceived risk (dine-ins), relative to other restaurants. They also show that consumers did not avoid Chinese people where there is less perceived risk (at non-Chinese dine-ins), but did avoid them where there is more perceived risk (at Chinese dine-ins). In summary, restaurant-goers discriminated where there was elevated perceived risk of infection, but otherwise did not. If taste-based motives did contribute to the abrupt decline in demand by the representative consumer, the findings indicate that it must have come in a highly particular and unexpected form: it cannot have been due to a reduced willingness to do business with Chinese restaurants, nor a reduced willingness to engage in brief interactions with Chinese people, nor a reduced willingness to engage in prolonged interactions with Chinese people. While I acknowledge that such sentiment may have risen in the initial weeks following the covid announcement, the overall evidence indicates that among consumers, heightened concerns over health risks provide a better explanation.

1.4.2 Robustness Tests

In this section, I provide results from several robustness tests to validate the results. First, I check whether the results are sensitive to sample selection. In Appendix tables A.1.9 and A.1.10, I replicate the results from the baseline framework of dine-share and the results from table 1.3 on Chinese-share, using the unrestricted sample of limited and full service restaurants that were open on January 20. This sample lifts the restriction of conventional and cuisine-unambiguous restaurants, thus including bars, cafes, other shops, and restaurants that offer multiple cuisines. This sample also lifts the restriction of a minimum 40 raw visits in each of the three pre-covid periods. The estimated parameters of interest for the non-demeaned and preferred specifications continue to provide similar results, as do the corresponding coefficients for covid cases. The main difference with the preferred sample lies in the response to European dine-ins, for which the coefficient representing the response to dine-ins remains negative but is noticeably larger in magnitude. Further inspection indicates this difference is driven by restaurants with typically few visitors in pre-covid periods.

Second, I inspect whether using raw GPS-detected visits instead of SafeGraph's suggested state-weighted variable provides a different interpretation of the results. In table A.1.11, I replicate the baseline results using the logged raw visits as the outcome variable. The overall interpretation is unchanged, with the coefficient on dine-share being slightly higher in magnitude across specifications. The coefficients on covid cases is also slightly higher for non-Latin cuisines but remains insignificant for Latin cuisines.

Next, I examine the sensitivity of the estimated response to a higher Chinese share in equation 1.5 to the transformation $log(1+Y)$. Specifically, I follow the suggestion by Chen and Roth (2023) and re-estimate the results using an alternate transformation of visits: the percent decline in visits relative to the pre-announcement average weekly visits. The results, shown in table A.1.12, are consistent with the preferred specification. Chinese customers were less likely to visit restaurants with typically large Chinese clientele, whereas non-Chinese customers did not reduce demand for such restaurants.

I also examine results using different specifications of control variables. Specifically, instead of covid cases in the *previous* two weeks, I examine results when controlling for covid cases in the combined current and last week. As in the baseline specification, I take the log of 1 plus this count. The results are shown in table A.1.13. This specification, labeled "Cases", indicates slightly smaller effects for Chinese restaurants. However, this specification partly associates restaurant visits with cases that had not yet been reported (for example, visits on Monday when a new case is reported later in the same week), which would attenuate the estimate. In a separate specification, I show that replacing the control for population with the logged population also makes little difference.

Finally, table A.1.14 provides coefficient estimates of the baseline model for control variables. For covid cases, the specification indicates a one log-point increase in recent covid cases is followed by approximately an 8 and 13 percentage point decline in visits to American and Chinese restaurants, respectively. Among dine-ins, Chinese, Asian, and European restaurants saw a greater reduction in traffic following new cases whereas Latin American and American restaurants saw weaker but similar declines, shown in the top row. The table shows that in addition to recent covid cases, the 2016 county share for democrat predicts reduced visits overall and to Chinese restaurants following the covid announcement. However, the size indicates a small effect: a county with a 10 percentage point greater share for Hilary Clinton reduced visits to American and Chinese restaurants by 2 and 3 percentage points, respectively. The remaining controls do not predict a greater decline in demand for Chinese restaurants relative to other restaurants.

1.5 Belief Accuracy

Were consumers justified in avoiding Chinese dine-ins more than other dine-ins? The extent to which they were justified—equivalently the extent to which beliefs were accurate—is proportional to the likelihood that the virus had a greater presence at Chinese restaurants than at others.

Health studies indicate that covid was widely present in the US by December 2019 (Basavaraju et al. 2021). Combined with event studies in figure 1.2, this indicates that uninformed consumers responded swiftly to the CDC's announcement but otherwise lacked information on the health risks of in-person contact. As such, a large decline in demand for restaurants can be justified by some level of risk aversion. However, a discrepancy between the observed discriminatory demand changes and what would be explained by the true (unobserved) differences in risk can be attributed to exaggerated beliefs on the additional health risks of Chinese restaurants.

Small differences in covid cases among the Chinese and non-Chinese population were likely during the initial weeks of the virus's arrival to the US—well before the CDC's announcement—due to international travel from China. However, these differences diminished quickly due to the rapid diffusion of the virus. For example, CDC data shows that Asians overall had similar or lower rates of cases, hospitalizations, and deaths compared to other major race groups by March, 2020 (Latoja and Artiga 2022; Romano 2021).

As data on transmissions at restaurants is not available, I provide suggestive evidence that any differences in virus risks between cuisine types is likely too small to justify the large response. If the virus held a greater presence at Chinese restaurants during the observational period, then it is likely that counties with more Chinese restaurants, or counties where Chinese restaurants were more popular, will have eventually outpaced other counties in reported covid cases, particularly because several weeks passed before social distancing or lockdown measures were enacted to restrict the spread. As such, I use county level counts of and visits to Chinese restaurants in the 2020 pre-treatment period as predictors of future covid cases in equation 1.8. Here, $CumulCases$ is cumulative cases reported in county c at some future week. Using the unrestricted sample of restaurants, $Share$ is the Chinese share of restaurants or the Chinese share of restaurant visits, and $\mathbf{X}$ is the matrix of control variables.[22] I estimate this model for each of 80 weeks of cumulative cases starting on the week of the covid announcement, and plot the corresponding estimates of β in Appendix figure A.2.3.

$$log(1 + CumulCases)_c = \alpha + \beta Share_c + \mathbf{X}_c\gamma + \varepsilon \tag{1.8}$$

On the vertical axis is the coefficient estimate that represents the log point increase in cumulative covid cases predicted by a one percentage point increase in 1) the share of all restaurants that are Chinese restaurants or 2) the share of all restaurant visits that went to Chinese restaurants in the aggregate pre-treatment period. All models control for log population, density, and Democratic share, and subfigures (c) and (d) further control for the Chinese share of the county's population. The vertical dashed line represents the end of the observational period.

The figures do not indicate that counties with greater presence of or popularity of Chinese restaurants were predisposed to covid cases. In fact, all figures suggest a slightly negative correlation between the pre-announcement popularity of Chinese restaurants and covid cases in the initial months after the lockdowns. This may be reflective of several factors that are outside the scope of this paper, such as varying implementation of and adherence to lockdown and social distancing measures.

[22]The unrestricted sample includes non-conventional and cuisine-ambiguous restaurants, thus including bars, cafes, other shops, and restaurants that offer multiple cuisines. It also includes as restaurants that did not meet the minimum 40 raw visits restriction.

1.6 Discussion

Understanding the motives underlying discriminatory behavior has significant policy implications. Behavior driven by false beliefs can be mitigated — or exacerbated — with new information or by making the subject matter more salient to the discriminator, whereas behavior driven by preferences is less responsive. However, distinguishing between the two encounters an identification challenge in observational settings where it is often difficult to tease out the motives. In this paper, I explore a unique observational setting and exploit a wide array of variation to test for the relative roles of belief-based and taste-based discrimination among consumers in a natural experiment affecting the entire restaurant industry.

My findings indicate that the covid announcement generated concerns among restaurant-goers that sitting in Chinese restaurants posed a greater health risk relative to other restaurants. This aversion was also found among other foreign restaurants associated with greater covid cases abroad when covid cases are reported locally. However, I show that non-Chinese customers did not become averse to doing business with Chinese restaurants, nor averse to interacting with Chinese people. Instead, they avoided Chinese restaurants with the greatest perceived risk based on the expected duration of their visit. Although I acknowledge that taste-based discrimination may have influenced consumers, such influence would have come in a particular and unexpected form. For example, the quality of visits—the demeanor of customers and their interactions with other customers and staff—may have changed following the covid announcement, and this would not be captured in the data but may constitute taste-based discrimination.

My findings are timely due to rising concerns over the spread of misinformation on social media, with potentially significant consequences to elections, health outcomes, and the general social cohesion of a rapidly diversifying population. The recent development of artificial intelligence and its utilization on social media platforms is of particular concern, as it can facilitate the speedier creation of increasingly convincing but false content that may impact off-platform market choices of consumers. Furthermore, the accuracy of the informational content need not matter for it to have an adverse impact, as this paper shows; consumers may form false beliefs due to cognitive biases, such as representative heuristics or representative signal distortion, when exposed to information or following events that elevate health concerns. It is therefore incumbent on future research to investigate feasible solutions to these adverse responses.

Chapter 2

Consumer Discrimination and Social Media In Uncertain Times

2.1 Introduction

The growing popularity of social media has led to a recent literature studying the effects of platform usage on a range of important outcomes, including crime (Bursztyn et al. 2019; Cao, Lindo and Zhong 2022; Müller and Schwarz 2021), health (Braghieri, Levy and Makarin 2022), and political attitudes (Allcott and Gentzkow 2017; Levy 2021). However, there is scarce evidence of an impact of social media on the simplest and quintessential economic choice: how we spend our money. In this paper, I build on the social media literature by examining the role of Twitter in restaurant demand changes during the early weeks of the COVID-19 (Covid) pandemic. I show that greater salience of Covid as a topic of conversation on the social media platform altered restaurant goers' choices, with a greater impact on Chinese restaurants relative to other cuisine types.

Beginning in January 2020, consumers responded to the news of the novel respiratory virus by quickly reducing demand for restaurants, with Chinese restaurants seeing the largest overall decline in foot-traffic (Huang et al. 2023; Yi 2023). However, policies that would restrict restaurant choice were not announced until March 2020, allowing for a seven week window during which consumers had unrestricted market choices and rapidly evolving information on the virus. Uruci (2024a) shows that the greater aversion to Chinese restaurants in these early weeks is consistent with belief-based discrimination over health concerns rather than taste-based discrimination over Chinese restaurants and individuals, indicating that informational channels played a significant role in differentially shifting demand for cuisine types.

During highly dynamic and uncertain times such as this, social media platforms have the potential to spread information quickly and broadly, without regard for accuracy (Allcott and Gentzkow 2017). In fact, Cinelli et al. (2020) show that leading platforms, including Twitter, contributed to spreading Covid-related news from unreliable sources at similar rates as those from reliable sources during this time Related, Bartoš et al. (2021) finds experimental evidence of widespread hostility toward foreigners when Covid is exogenously made more salient, indicative of scapegoating during times of crisis (Bursztyn et al. 2022*b*).[1] Thus, greater exposure to informational channels that elevate the salience of Covid may contribute to consumer discrimination against Chinese restaurants. Similar to their strategy, which experimentally manipulates the provision of information, I rely on local Twitter usage before the first case of the virus as plausibly exogenous variation in informational exposure that elevated the salience of Covid-related content in some

[1] Czech subjects in the treatment group are primed with questions focusing on the coronavirus crisis before starting a controlled money-burning task. The study finds increased hostility toward Asian, other European, and American foreigners, indicative of out-group discrimination. These results are consistent with heightened health-related fears of foreigners in response to elevated salience of the virus.

counties more than others. I then study the resulting impact on relative demand for Chinese restaurants.[2]

The main outcome data for consumer demand comes from Safegraph, which provides foot-traffic to the universe of US restaurants (SafeGraph 2022). I obtain data for the time periods December 2018 - March 2019 and December 2019 - March 2020 and use the first year as the counter-factual restaurant demand for each cuisine type. I adopt the triple-differences (DDD) specification from Uruci (2024a), where the consumer response to the covid announcement is estimated as the change in log visits in the seven weeks after January 20 relative to the seven weeks prior. I compare the response within cuisine type to the same period in the prior year, which removes cuisine-specific seasonal variation in restaurant demand. I then take the third difference across cuisines to estimate consumer discrimination.

The Twitter data is obtained by scraping a large dataset of geo-tagged and time-stamped tweets, from which I construct a measure of county-week covid salience, which I call *covid tweeting intensity.* I define *covid tweets* as those that contain the strings "covid", "coronavirus", or "corona virus" in the main text or hashtags, and define the covid tweeting intensity as the county-week number of covid tweets divided by the county population. Conditioning on reported cases and other county characteristics, I examine whether greater covid tweeting intensity reduced visits to Chinese restaurants more than to other restaurants. To identify a causal effect, I use a Bartik-style instrumental variable (IV) that is formed by interacting the national time series of covid tweet counts with the prior year's county-level tweeting intensity, which did not include any covid-related

[2] I abstract from the discussion of beliefs vs. tastes arising from social media and from accurate vs. inaccurate informational sources, instead studying the overall impact of elevated salience of covid-related social media content on consumer behavior in the marketplace.

content. Therefore, the instrument predicts the weekly local covid tweeting intensity using the national covid tweet count that is adjusted according to each county's typical Twitter usage in the pre-covid world.

In the preferred IV specification, I find that a one standard deviation increase in Covid salience reduced visits to non-Chinese restaurants by 3.7 percentage points and to Chinese restaurants by an *additional* 1.2 percentage points.[3] A similar effect is found for Asian and European restaurants but no effect is found for Latin American restaurants.

The central result is consistent with the finding in Uruci (2024a) that, following a new case of covid in a county, consumers discriminated against Chinese, other Asian, and European restaurants, but not Latin American restaurants. Since covid cases were rising in China, other Asian countries, and Europe, but not in Latin America during this time, both results are consistent with a salience effect: whether new information about covid is accompanied by real risk (new local covid cases) or not accompanied by real risk (more covid-related Twitter content) consumers avoided the foreign restaurants associated with the greatest risk abroad.[4]

The central contribution of this paper is to provide evidence of a social media effect on off-platform consumption choices. Past studies have examined social media effects on hate crimes and hate incidents toward minorities (Bursztyn et al. 2019; Müller and Schwarz 2021, 2023). More related studies find increased hate incidents following President Trump's "Chinese virus" tweet in the time period immediately following the one in this study (Cao, Lindo and Zhong 2022; Lu and Sheng 2022). Hate incidents and crimes are high-stakes and extreme actions that are performed by a small segment of

[3]This can be translated as follows: if the county-week increase in the number of tweets about covid equals 1 percent of the population, the effect is a 1.5 percentage point decline in visits to non-Chinese restaurants and an additional 0.5 percentage point decline to Chinese restaurants.

[4]These results are also consistent with experimental evidence by Bartoš et al. 2021.

the population. This paper instead examines changes in demand for restaurants that are differentiated by ethnicity, which is a relatively low-stakes and covert response, but one with much greater prevalence in the population and one that avoids the social costs of discrimination.[5]

The remainder of this paper is organized as follows. Section 1 describes the observational setting and data sources. Sections 2 describes the identification strategy and section 3 provides the results. Finally, Section 4 concludes with a brief discussion of the overall results.

2.2 Observational Setting

On January 20, 2020, the Center for Disease Control (CDC) confirmed the first case of covid in the US (Center for Disease Control 2020).[6] Borrowing from Uruci (2024a), several important features in the following seven week period allow for uncovering an off-platform market impact of social media.

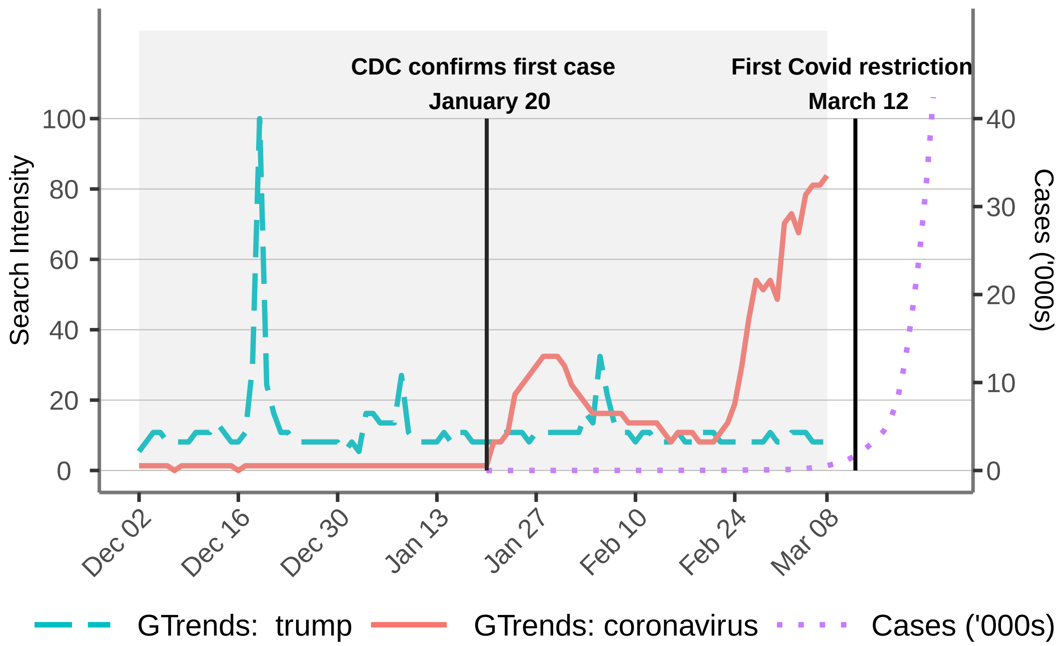
During this time, consumers were aware that the virus was in the US but no policies had yet been announced that would restrict restaurant choice. Consumers were therefore free to make unrestricted market choices but with new information on the presence of the virus. Second, the virus was a major topic of public concern immediately following the announcement. As shown in Figure I, "coronavirus" was the top searched word on Google during this time period (Google 2022).

[5]Other studies on media examine positive coverage of an ethnic minority group (Zussman (2023)), the spread of misinformation (Allcott and Gentzkow 2017), health (Braghieri, Levy and Makarin 2022; Bursztyn et al. 2022*a*), and public opinion (Levy 2021; Simonov et al. 2020; Wang 2021). Evidence of mass media effects on ethnic prejudice include Adena et al. (2015); DellaVigna et al. (2014), and Yanagizawa-Drott (2014).

[6]The first positive test was confirmed on Monday, January 20 and the announcement was made the following day. I consider the week starting on January 20th as the first treatment week.

Third, this period lacks the influence of high-profile political figures using terms that pointedly associate the virus to Chinese people. For example, Donald Trump's use of "Chinese virus," "Wuhan virus," or "Kung Flu" did not first occur until March 16, and the first instance of any US politician to use such terms was on March 6 (Cao, Lindo and Zhong 2022; Huang et al. 2023). This suggests that the growing prominence of covid in news or on social media was a result of updates reported by health organizations that generated organic public dialogue rather than instigated by political figures.

Figure I: Google Trends Search Intensity, December 2019 - March 2020



Note: Observational period in shaded region. The figure displays the Google Trends search intensity of "trump" and "coronavirus" between December 2019 - March 2020 (left axis) and the number of reported covid cases in the US, in thousands (right axis). The search intensity is normalized to the time period and scaled to a range of 0 to 100. The peak search for "trump" occurred on December 19, 2019, the day after Trump's impeachment by the House of Representatives. The maximum search intensity for "coronavirus", occurring on March 8, is approximately 85% of that of "trump" on December 19. Coronaviruses are a group of viruses that cause respiratory illnesses and are generally transmitted through close personal contact. Despite the official name of "COVID-19" announced on February 11, "coronavirus" was the top searched topic in the seven week period after the CDC announcement.

Finally, covid cases remained low in the US during this time whereas cases in China had already spiked, as shown in Appendix table B.1.1, which fortifies the association of the virus with China. Cases were also rising in other parts of Asia and Europe, but not in Latin America, which provides cross-cuisine variation in perceived health risks of foreign restaurants.

2.2.1 SafeGraph Data

The SafeGraph data comes in two sets called Places and Patterns for the universe of full-service and limited-service establishments. Places contains descriptive variables on each restaurant, including its name, location, open and close dates, and a description of the restaurant, with labels such as “Chinese”, “Dinner”, “Food Truck”, and “Late Night”. Using these labels, I classify restaurants into five broad cuisines: Chinese, Asian (non-Chinese), American, Latin, and European. I keep establishments that can be considered conventional restaurants and remove those with an ambiguous cuisine group, but otherwise classify restaurants as American if they are not suggestive of another major group. This restriction, which removes 22% of restaurants, is adopted because a strong and clear association between cuisine type and ethnicity is necessary.

The Patterns data provides a weekly panel of visit counts to each restaurant. I obtain data for the time periods December 3, 2018 - March 10, 2019 and December 2, 2019 - March 8, 2020.[7] For the remainder of the paper, “2019” will refer to the former period and “2020” will refer to the latter period. This way, there are three pre-treatment periods and one post-treatment period. I also keep restaurants with at least 40 raw visits in each of the three pre-treatment periods.[8] Finally, I reduce the sample to counties with a

[7]Thanksgiving occurred one week prior in 2019 but two weeks prior in 2018, making the last two weeks of November less comparable between the two years.

[8]Restaurants with few detected visits may be less precisely comparable across the three pre-treatment

Table 1.1: Summary Statistics of Restaurant Visits

Cuisine	Mean Visits 2019	Mean Visits 2020	Close Rate (%) Post-2020	Restaurants	Obs.
American	83.3	102.3	0.51	207,435	5,801,884
Asian	52.5	59.4	0.70	31,493	880,450
Chinese	52.8	59.7	0.63	23,411	654,597
European	62.2	72.8	0.71	26,938	753,128
Latin	68.7	80.0	0.60	68,521	1,916,283
Sample	74.2	89.3	0.57	357,798	10,006,342

Note: The table provides summary statistics of the SafeGraph data, displaying for each cuisine the average weekly recorded raw visits, the percent of restaurants that closed between January 20 and March 8, 2020, counts of restaurants, and total restaurant-week observations. The close rate is calculated using the "closed on" date variable in the data. The higher visit count in 2020 is due to an increase in the pool of smartphone devices that SafeGraph tracked over the year, which is accounted for in the state-weighted visits and in all triple difference models.

minimum population of 25,000 due to the low level of Twitter usage in small counties that contribute to measurement of the central explanatory variable, which I discuss in the following subsection.

The resulting sample includes 357,798 limited or full service establishments that were open on January 20 and 1) are conventional restaurants, 2) are unambiguously classified into one of the five large cuisine groups, and 3) met the minimum 40 raw visits threshold in all three pre-treatment periods. This represents approximately 70% of all limited and full service restaurants reported by the US Census in 2019 (US Census 2019*b*).[9]

Table 1.1 provides summary statistics of the final restaurants sample, including mean visits, the percent of restaurants that closed post-announcement, counts of restaurants, and total observations. The low close rate indicates that restaurants did not preemptively close in anticipation of reduced demand or due to the virus, and shows that Chinese restaurants had similar close rates to other cuisines.

periods. This restriction removes 14% of restaurants before cuisine assignment.

[9]This also represents 64% of the unrestricted sample in the SafeGraph data, which includes limited or full service establishments that were open on January 20 but without the remaining three restrictions.

Control variables include cases in the last two weeks, the 2016 share of votes for the Democratic party, population, population density, and the share of the population of Chinese ancestry.[10]

2.2.2 Twitter Data

I use data from Twitter, a mainstream social media platform for quickly spreading online social and political commentary, to construct a measure of local salience of covid (Cinelli et al. 2020). One might wonder how tweets made in a local area correspond to consumers in the same area being exposed to greater covid-related content: after all, once a tweet is made publicly, *anyone* can see it. An important phenomenon that facilitates heterogeneity in exposure to online content is that social media networks are highly local. The typical user is much more likely to follow and be followed by other users who live close by than those who live far away. For example, Takhteyev, Gruzd and Wellman (2012) show that 39% of all Twitter followings are less than 100 kilometers, based on the geographic coordinates of the two. This is much greater than the approximate 2% of followings that would exist within this distance if Twitter followings were random across users worldwide; a similar pattern is found among the top 15 metropolitan clusters of Twitter users in the US. The phenomenon is also found in Bailey et al. (2018)'s study of social connectedness, which shows that for the median Facebook user in the US, over

[10] Data on reported covid cases is provided by New York Times (2022) and includes the cumulative reported covid cases for each county-day. I take one plus the count of new cases in the prior two weeks, in logs, as the primary control variable. This data set does not distinguish the five counties of New York. I group these counties into one county representing New York City throughout the analysis. I obtain the county share of 2016 Democratic party votes from MIT Election Data and Science via Harvard Dataverse (2023), and this variable is centered at 0.5. I obtain data on the 2018 and 2019 county populations and population density from the US Census and ACS 2019 5-Year Sample (US Census 2019*a*; Ruggles et al. 2023). I use the 2015-2019 American Community Survey to construct the county level Chinese ancestry share of the population, which is limited to the population aged 16 and over.

55% of friends live within 50 miles, in contrast to the less than 1% of the population that lives within the same distance.

There is also significant inequity in who tweets: the top 10% of users by tweet volume post 92% of all tweets. At the same time, about 66% of users report visiting the site at least once a week, indicating that just few users are predominantly the posters while most users are just observers (Pew Research Center 2020, 2021).[11] This is supported by the Twitter data used in this study, which yield a Gini-coefficient of 0.69 in tweeting inequality of users. Collectively, these statistics suggest that a small subset of highly active users produce the online content that, due to the local concentration of follower networks, is predominantly consumed by their neighbors. Accordingly, at the time of the covid announcement, these typically active users provided the lion share of the covid content that was predominantly observed by nearby followers.

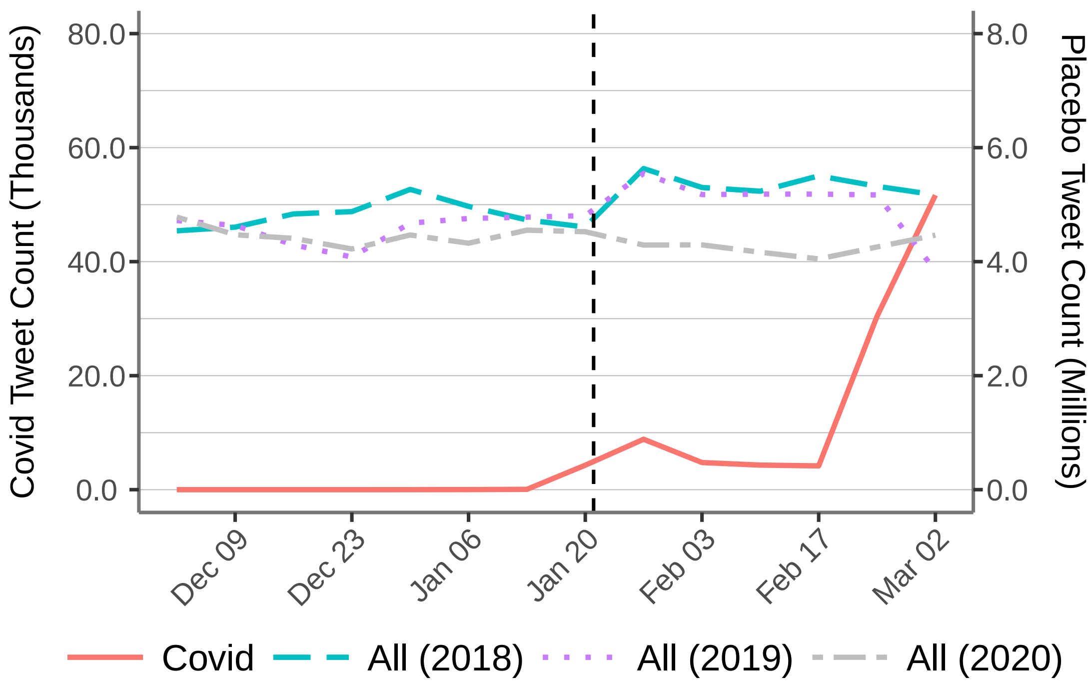
I scrape tweets using the python library *snscrape*. I clean the data according to the procedure outlined in Appendix 2.A and construct two data sets. The first provides county-week-year level counts of tweets that contain the strings "covid", "coronavirus", or "corona virus". I refer to these as covid tweets.

A second data set is obtained from tweets made during the fourteen week observational period between December - March for three years: 2018 - 2020. This dataset of tweets does not restrict to any keyword. Instead, it samples any geo-tagged tweets made in the US, comprising a total count of 16.1 million. I aggregate tweet counts to the county

[11] Twitter is used by approximately 27% of US adults, compared to 70% for Facebook and 82% for YouTube, thus its content reaches a smaller audience compared to the larger platforms. However, this concern is allayed when considering that Twitter users skew towards residing in urban areas where Chinese restaurants are also more likely to be located, and that Twitter users report a higher rate of regularly getting news from the platform (53%) compared to users on other platforms (Facebook 44%, and Youtube 30%) (Pew Research Center 2013, 2022). The observed effects from Twitter in this paper do not preclude nor prove discrimination effects driven by other platforms, and it is possible that platforms with wider reach may have contributed a larger effect.

level for 2018 and 2019 to obtain a measure of typical local tweeting levels one year before the observational period. I separately aggregate counts to the county-week level for 2019 and 2020 to obtain a measure of typical Twitter content concurrent to the observational period. I refer to the former as unrelated tweets and the latter as placebo tweets, however they are drawn from the same data source and they are identical in the overlapping year, 2019. Appendix table B.1.3 provides descriptive statistics on these counts and Appendix figure B.2.1 provides an illustrative timeline of how each constructed dataset is used in the analysis.[12]

Figure II: Tweet Counts, December 2019 - March 2020



Note: The figure displays the national time series of geo-tagged covid tweets (left axis), which include any tweets with the string "covid", "coronavirus", or "corona virus", and all tweets (right axis), which are unrestricted in string text. The sample of unrestricted tweets, described in the Appendix, are drawn from the entire population of Twitter content and include covid tweets. The right axis is scaled according to the sample rate of 8.33%.

[12]Placebo tweets include tweets that reference the virus, although estimates are nearly identical when these are removed as they represent a small share. Based on the sampling rate of 8.33% discussed in Appendix section B.1, covid tweets represent approximately 0.36% of placebo tweets.

Figure II plots the count of covid and placebo tweets. The figure displays a similar pattern to the Google Trends search intensity in Figure I, with an initial bump in the first two weeks, a decline in mid-February, then a sharp rise in the last few weeks of the observational period as cases escalated. Overall tweeting during this time frame remained constant across all three years.

2.3 Empirical Strategy

2.3.1 Consumer Response

Following Uruci (2024a), the triple-differences estimation strategy will compare, in the first two differences, visits to Chinese restaurants after the covid announcement relative to before and in the current year relative to the prior year. In the third difference, it compares the consumer response across cuisines. The DDD framework is preferred over the standard difference-in-differences (DID) framework in order to account for seasonal trends in consumer preferences, such as increased demand for Chinese food on Christmas.[13]

The DDD strategy can be interpreted as follows: given that a conventional DID model using other cuisine types as the control group may not account for seasonal variation in consumer preferences, such differences are netted out by first controlling for each cuisine's demand in the prior year, when there was no covid. This is equivalent to first adopting a conventional DID framework, separately for each cuisine, that interacts indicators for post-treatment and year, and includes control variables. Then, in the third difference,

[13] Christmas and New Year each occurred within the same observational week in 2020 relative to 2019. The Chinese New Year occurred on different weeks in 2020 (February 5) relative to 2019 (January 25), but occurred within the post-treatment period in both years.

cuisine type is interacted with all right-hand-side variables, which will result in a DDD model that has removed cuisine-specific seasonal variation. As such, the identifying assumption in the DDD is that seasonal variation may differ across cuisines but is similar within cuisines between 2020 and 2019 (Gruber 1994; Olden and Møen 2022).

$$\begin{aligned} log(visits)_{r,i,j,k,c} = \quad & \tilde{\alpha}_1 Post_i \times 2020_j \\ & + \tilde{\alpha}_2 Post_i \times 2020_j \times Chinese_k \\ & + \mathbf{W}_{i,j,k}\mu + \mathbf{X}_{i,j,c}\gamma + \eta_{k,c} + \epsilon_{r,i,j,k,c} \end{aligned} \tag{2.1}$$

Equation 2.1 represents the baseline specification, capturing the consumer response to the covid announcement. The subscripts index restaurant r, week i, year j, cuisine k, and county c, and visits are measured in logs. Each estimated coefficient can be approximated as a percentage change in visits due to a one unit increase in the explanatory variable. The model includes the remaining full set of DDD interactions in $\mathbf{W}$ and control variables in $\mathbf{X}$.[14] Finally, the model includes county-cuisine fixed effects to ensure it has the interpretation of a difference of DID models. These fixed effects capture time-invariant characteristics common to each cuisine within a county. Under the DDD identifying assumption of a constant difference in parallel trends, $\tilde{\beta}_1$ represents the impact of the covid announcement on demand for the control group and $\tilde{\beta}_2$ represents the additional impact of the covid announcement on demand for Chinese restaurants.

Results from 2.1 are provided in table B.1.4 and indicate that following the covid

[14] The remaining set of DDD variables include $Post$, 2020, $Post \times Chinese$, and $2020 \times Chinese$. The uninteracted indicator, $Chinese$, is absorbed by the fixed effects. Other than covid cases, which vary by county-week and are imputed 0 in all pre-treatment observations, the remaining controls vary by county or by county-year. These controls enter the model as interactions with $Post \times 2020$ and additionally as interactions with $Post \times 2020 \times Chinese$ to control for differential covid-induced responses to local characteristics for Chinese and non-Chinese restaurants.

announcement, restaurant demand declined by approximately 13 percentage point, and an additional 2 percentage points for Chinese restaurants. In columns 5-9, which uses American restaurants as the reference group, Chinese restaurants saw the largest declines following a newly reported case of covid, and effect that is also found among Asian and European restaurants, but not Latin American restaurants.

2.3.2 Twitter Intensity

I define the following measure of local tweeting intensity:

$$TwitterIntensity_{i,j,c} = \frac{Tweets_{i,j,c}}{Users_c} \times \frac{Users_c}{Pop_c} = \frac{Tweets_{i,j,c}}{Pop_c}$$

The first term in the interaction captures the tweeting intensity of the typical Twitter user during week i, in year j, in county c, while the second term weighs the measure by the share of the local population that uses Twitter. Their interaction simplifies to the number of tweets per the local population, which is then standardized. When the measure is restricted to tweets that include "covid", "coronavirus", or "corona virus", the estimated coefficient of $CovidT.I.$ will approximately represent the marginal effect of a one standard deviation increase on the percentage point change in the outcome variable. Since this results in imprecise measures of covid tweeting intensity for small counties, many of which have few restaurants, even fewer Chinese restaurants, or no tweets made in the observational period, I limit the sample in this section to counties with a minimum population of 25,000.[15]

Equation (2.2) displays the baseline model for the consumer's response to covid tweet-

[15] This restriction drops 47% of counties, but only 3% of Chinese restaurants and 4% of all restaurants.

ing intensity. $\widetilde{\beta}_1$ approximates the marginal effect of a one standard deviation increase in $CovidT.I.$ on the percent change in visits to the reference group and $\widetilde{\beta}_2$ is the additional marginal effect on Chinese restaurant visits. To capture an exacerbation effect, I control for the full set of RHS variables from the baseline DDD model in equation (2.1), which are stored in $\mathbf{X^{DDD}}$. In other words, $\widetilde{\beta}_2$ answers the question: conditional on the overall discrimination by consumers and controlling for covid cases and other country characteristics, did consumers avoid Chinese restaurants *more* when Covid was a more salient topic on Twitter?

$$
\begin{aligned}
log(visits)_{r,i,j,k,c} = \quad & \widetilde{\beta}_1 CovidT.I._{\cdot i,j,c} \\
& + \widetilde{\beta}_2 CovidT.I._{\cdot i,j,c} \times Chinese_k \\
& + \mathbf{X^{DDD}}\gamma + \eta_{k,c} + \epsilon_{r,i,j,k,c}
\end{aligned} \tag{2.2}
$$

Since nearly all observations for $CovidT.I.$ in the pre-treatment period are 0, the standalone variable is interchangeable with the interaction $CovidT.I. \times Post \times 2020$. The placebo specification, which provides variation before and after treatment, requires the inclusion of the interactions with the post-treatment variable in order to detect whether regular tweeting activity changed *in response* to the covid announcement. Placebo specifications therefore include the full set of interactions of $PlaceboT.I.$, $Post$, 2020. Their triple interaction becomes comparable to $CovidT.I.$ in equation (2.2).

2.3.3 Identification

Equation (2.2) raises concerns over whether $\widetilde{\beta}_2$ identifies a covid salience effect driven purely through Twitter. As recent cases are controlled for, $\widetilde{\beta}_2$ would not be identified if another source of information brings attention to covid on Twitter and also affects restaurant visits. One example is a non-covid illness, which may lead consumers to suspect that they've contracted covid, avoid Chinese restaurants, and promptly express their alarm on Twitter. Spatial and temporal variation in immune system strength would then bias the estimated coefficient of interest (negatively) by increasing covid content on Twitter and reducing visits to Chinese restaurants.

Another concern is mismeasurement in covid tweets, which is likely given the low tweet count for many counties. For example, in the post-treatment period, 34 counties account for half of covid tweets while approximately one-fourth of counties, mostly the least populated ones that also have the fewest restaurants, do not have any covid tweets in the data. This would result in attenuated estimated coefficients using ordinary least squares (OLS) estimation.

An instrumental variable would be necessary to more accurately estimate the impact that would have been generated solely by the typical informational exposure by the platform, and in the absence of external channels and mismeasurement. I use a Bartik-style instrumental variable that interacts the tweeting intensity one year prior to the covid announcement with the national time series of covid tweet counts. Specifically, I construct the following interaction:

$$Z_{i,j,c} = \frac{UnrelatedTweets_{(j-1),c}}{Pop_{(j-1),c}} \times CovidTweets_{i,j}$$

The first term of the interaction is a collapsed measure of $TwitterIntensity$ for the entire fourteen week period in the prior year. It mimics the "share" component of the conventional Bartik instrument, and provides a county level cross section of regular tweeting intensity that is free of any covid-related content. The large sample of unrelated tweets from the prior year is useful because it provides a more precise measure of typical local Twitter use. The second term is the weekly time series of covid tweet counts made in the US and mimics the "shift" component of the Bartik instrument. It provides a measure of the national salience of covid on Twitter in a given week, and is shown in figure II. When interacted, the prior tweeting intensity re-weighs the national time series of covid tweets according to each county's typical Twitter usage just before the covid announcement. The resulting first and second stage estimated models are displayed in equations (2.3) and (2.4), respectively.

First Stage:

$$\begin{aligned} CovidT.I._{i,j,c} = \quad & \gamma_2 Z_{i,j,c} \times Chinese_k \\ & + \gamma_1 Z_{i,j,c} \\ & + \mathbf{X^{DDD}}\theta + \mu_r + e_{r,i,j,c} \end{aligned} \tag{2.3}$$

Second Stage:

$$\begin{aligned} log(visits)_{r,i,j,k,c} = \quad & \beta_2 \widehat{CovidT.I}._{i,j,c} \times Chinese_k \\ & + \beta_1 \widehat{CovidT.I}._{i,j,c} \\ & + \mathbf{X^{DDD}}\gamma + \eta_r + \epsilon_{r,i,j,k,c} \end{aligned} \tag{2.4}$$

Table 1.2: Consumer Discrimination, Covid Salience on Twitter

	Dependent Variable: Log Visits							
Reference Group:	Covid				Placebo			
All Other Cuisines	OLS		IV		OLS		IV	
Term T. I. x Post x 2020 x	-0.004	-0.004*	-0.011**	-0.012***	0.011	0.010	0.029	0.026*
Cuisine	(0.004)	(0.002)	(0.005)	(0.003)	(0.027)	(0.014)	(0.029)	(0.015)
Term T. I. x Post x 2020	-0.026***	-0.026***	-0.033***	-0.033***	0.049***	0.049***	0.045***	0.046***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.007)	(0.003)	(0.008)	(0.003)
Post x 2020 x Cuisine	-0.025***	-0.024***	-0.024***	-0.023***	-0.028***	-0.028***	-0.026***	-0.025***
	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)
Post x 2020	-0.127***	-0.129***	-0.126***	-0.128***	-0.126***	-0.128***	-0.127***	-0.129***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Cuisine	-0.040***	-0.042***	-0.041***	-0.043***	-0.062***	-0.064***	-0.062***	-0.064***
	(0.012)	(0.006)	(0.012)	(0.006)	(0.011)	(0.005)	(0.011)	(0.005)
Cases	-0.063***	-0.063***	-0.063***	-0.063***	-0.082***	-0.082***	-0.082***	-0.082***
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)
Total Obs.	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342
R Squared	0.075	0.863	0.075	0.863	0.074	0.862	0.074	0.862
Tweet Count	108,526	108,526	108,526	108,526	11,454,880	11,454,880	11,454,880	11,454,880
FE: County x Cuisine	X		X		X		X	
FE: Restaurants		X		X		X		X
Full DDD Controls	X	X	X	X	X	X	X	X
Control Covid Cases	X	X	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X	X	X
Controls[4] x Nat. Tweets	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
The table provides estimates obtained from equations 2.1 and 2.4, representing the main OLS and IV specifications for covid and placebo tweets. The reference group in all models is non-Chinese restaurants. The full set of interactions of indicators for Placebo, Post, Year, and cuisine are included in placebo models. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. Control variables are also included as interactions with the national time series of tweets to capture the differential impact of the control in the term's salience. All controls are also included as interactions with cuisine. Standard errors clustered at the county x cuisine level for specifications including county x cuisine fixed effects, and clustered at the restaurant level for specifications including restaurants fixed effects.

2.4 Main Findings

Table 1.2 provides OLS and IV estimates of the covid tweets specifications, separately with country-cuisine and restaurant fixed effects, in columns 1-4. I contrast these results with the corresponding placebo tweets specifications in columns 5-8. The reference group in all columns includes the remaining four cuisine types. The columns estimated with OLS do not indicate a significant correlation between restaurant visits with covid tweets nor with placebo tweets. However, IV estimates for Covid T.I. are stronger and negative in magnitude for both Chinese and non-Chinese restaurants, as indicated in columns 3-4. On the other hand, IV estimates for Placebo T.I. (columns 7-8) remain insignificant, but slightly larger in magnitude compared to their OLS counterparts using the same fixed effects. This contrast suggests measurement bias is likely present in OLS specifications.

Including restaurant fixed effects instead of county-cuisine fixed effects improves precision but makes little difference to point estimates: a one standard deviation increase in covid tweets per county population results in approximately a 3-4 percentage point decline in visits to non-Chinese restaurants and an additional 1 percentage point decline in visits to Chinese restaurants. Back of the envelope calculations that use a standard deviation of 0.025 in Covid T.I., implies a decline of 1.5 percentage points in customer visits when the increase in covid tweets is approximately 1 percent of the county population, and an additional 0.5 percentage points in customer visits for Chinese restaurants.[16]

I also provide cuisine-specific results in table 1.3, with American restaurants as the reference group, and include county-specific week trends in addition to county-cuisine fixed effects. Consistent with the earlier results on covid cases, consumers exposed to

[16]Calculation approximates 1% of all tweets are geo-tagged.

Table 1.3: Consumer Discrimination by Cuisine, Covid Salience on Twitter

	Dependent Variable: Log Visits							
Reference Group:	Covid Tweets				Placebo Tweets			
American Restaurants	Chinese	Asian	Latin	European	Chinese	Asian	Latin	European
Term T. I. x Post x 2020 x	-0.011**	-0.017***	0.000	-0.010*	0.030	0.027	-0.001	0.025
Cuisine	(0.005)	(0.005)	(0.004)	(0.005)	(0.029)	(0.029)	(0.019)	(0.032)
Term T. I. x Post x 2020	-0.038***	-0.038***	-0.037***	-0.038***	0.027**	0.027***	0.026**	0.027**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.011)	(0.010)	(0.010)	(0.011)
Post x 2020 x Cuisine	-0.007	0.075***	0.040***	0.022***	-0.012**	0.067***	0.035***	0.015***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Post x 2020	-0.141***	-0.142***	-0.142***	-0.141***	-0.149***	-0.148***	-0.149***	-0.148***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine	-0.043***	-0.028***	0.007	-0.031**	-0.060***	-0.030***	0.012*	-0.030**
	(0.012)	(0.009)	(0.007)	(0.012)	(0.012)	(0.008)	(0.007)	(0.012)
Cases	-0.065***	-0.065***	-0.065***	-0.064***	-0.085***	-0.085***	-0.084***	-0.084***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Total Obs.	6,456,481	6,682,334	7,718,167	6,555,012	6,456,481	6,682,334	7,718,167	6,555,012
R Squared	0.066	0.064	0.054	0.059	0.065	0.064	0.054	0.058
Counties	1,591	1,591	1,591	1,591	1,591	1,591	1,591	1,591
Total Restaurants	230,846	238,928	275,956	234,373	230,846	238,928	275,956	234,373
Tweet Count	108,526	108,526	108,526	108,526	10,677,411	10,677,411	10,677,411	10,677,411
FE: County x Cuisine	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Control Covid Cases	X	X	X	X	X	X	X	X
Controls[4] x Nat. Tweets	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
The table provides estimates obtained from equation 2.4, representing the main IV specification for covid and placebo tweets. The reference group in all models is American restaurants. The full set of interactions of indicators for Placebo, Post, Year, and cuisine are included in placebo models. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. Control variables are also included as interactions with the national time series of tweets to capture the differential impact of the control in the term's salience. All controls are also included as interactions with cuisine. Standard errors clustered at the county x cuisine level.

more covid-related Twitter content reduced visits to all restaurants, but to a greater extent against Chinese, other Asian, and European restaurants, whereas Latin American restaurants saw a similar decline to American restaurants. The table further indicates a similar pattern in response to local covid cases, as found in Uruci (2024a). When a new case of covid is reported locally, consumers avoided Chinese, Asian, and European restaurants more than American restaurants, whereas Latin American restaurants did not see the same relative decline. On the other hand, placebo specifications indicate the same pattern of demand changes in response to new covid cases, but do not indicate a response to placebo tweets. Combined with the earlier results, this evidence is consistent with a salience effect of covid that resulted in reduced visits to foreign restaurants associated with rising cases abroad.

2.5 Robustness Tests

Bartik instruments are commonly used in the trade and migration literatures to generate plausibly exogenous variation that predicts the endogenous explanatory variable. The conventional instrument differs from the one in this paper in that it uses the pre-existing mix of some local exposure characteristic in the form of shares (industry shares in the case of trade, or national origin shares of foreign born in the case of migration) rather than in the form of a stock measure (here, the Twitter intensity). Each, however, is the interaction of an aggregate treatment measure with a local treatment exposure measure, and can essentially be interpreted as a difference-in-differences research design: the instrument is a projection of pre-treatment local trends into the treatment period.

The wide use of Bartik-style instruments has prompted increased attention to the underlying assumptions necessary for identification. Goldsmith-Pinkham, Sorkin and Swift (2020) provide an in-depth discussion of these assumptions, including for Bartik instruments of the kind used in this paper. One that directly applies to the current context is that the exposure variable need not be uncorrelated with *levels* of the outcome variable when the concern is with regard to *changes* in the outcome variable. This is analogous to the parallel trends assumption in difference-in-differences: initial levels of treatment and control group outcomes need not be the same, but their evolution over time would be parallel in the absence of treatment. This concern is raised in Jaeger, Joyce and Kaestner (2020), which comments on the study by Kearney and Levine (2015) on the impact of the reality show *16 and Pregnant* on teenage pregnancy. In the current setting, this translates to the correlation between restaurant demand and local twitter use exhibiting parallel trends across cuisines. An advantage of the current setting is that the placebo set of tweets is obtained from the same universe of tweets as the instrument, with the only difference being a one year lag. This allows for an obvious test of the parallel trends assumption: although there may typically exist correlation between the Twitter use and restaurant demand, did this change in response to the covid announcement?

Columns 5-9 from table 1.3 indicates that this baseline correlation increased slightly for American restaurants, but the relative correlation for foreign restaurants is not statistically significant. Curiously, the magnitudes are slightly larger in magnitude for Chinese, Asian, and European restaurants, bringing into question whether the placebo specification is a good test of the parallel trends assumption. However, the estimated triple difference framework requires a different interpretation of the parallel trends assumption. Specifically, the assumption is that the correlation between restaurant demand and local Twitter use may be different in levels between the two years and may be different for foreign relative to American restaurants, but the evolution of this correlation remained parallel over time. Since variation in the treatment exposure measure (the prior year's local Twitter use) is at the county-year level, the different levels between the two years can be accounted for by including county-year fixed effects, just as the different levels between the cuisines are accounted for with county-cuisine fixed effects. Table B.1.5 provides IV estimates of baseline models for Covid and Placebo T.I. specifications with the inclusion of county-year fixed effects. The first row shows that while covid tweets continue to predict reduced relative demand for Chinese restaurants after the covid announcement, the magnitude for placebo tweets disappears when both sets of fixed effects are included.

For greater granularity accounting for restaurant characteristics, such as location and reputation, I replace county-cuisine fixed effects with restaurant fixed effects and estimate cuisine-specific specifications in Table B.1.6 for covid tweets and in Table Table C.1.3 for placebo tweets. The two tables provide a sharp contrast. Covid-related tweeting continues to predict a reduced relative demand for Chinese, Asian, and European, but not Latin American restaurants, however, the placebo specification indicates no relative decline in demand for foreign restaurants, consistent with the main findings.

2.6 Discussion

My findings are timely due to rising concerns over the spread of misinformation on social media. Since this study focuses on a highly uncertain period during which there was little information confirmed by health officials about the novel virus, the results show that consumer behavior can be quickly influenced by elevating the salience of pertinent topics. Here, I show that demand for restaurants was influenced by elevated social media salience of a novel respiratory virus, without regard for accuracy of information. The recent propagation of artificially created content on social media platforms should raise obvious concerns. The speedy creation of increasingly convincing but potentially incorrect information may therefore impact the off-platform market choices of consumers with greater efficacy and with potentially deleterious consequences. It is incumbent on future researchers and policymakers to investigate feasible solutions to adverse consequences such as this.

Chapter 3

Isolating the Effect of Grant Rates on Asylum Applications: A Relative Discrete Choice Approach

3.1 Introduction

Individuals applying for asylum must weigh a number of factors in deciding in which country to seek refuge. For example, will the country provide refuge from the factors causing the emigration? What are the economic conditions in the destination country? Will there be compatriots who can provide community upon arrival? Another important factor is the grant rate (recognition rate), the rate at which the destination country grants asylum to applicants. Quantifying the effect of grant rates on the number of future applications is important to destination countries, especially following foreign events that spark mass emigration. For example, European countries received over 1.2 million applications for asylum in 2015, prompted by the Syrian civil war and the unrest in the region. That year, over 300,000 migrants were admitted under refugee or humanitarian status, whereas just over 900,000 were admitted for occupational, family, or educational reasons, per European Commission (2021). Thus, asylum can represent a significant share of demand for entry into Europe.

Although there is a modest literature on the effects of asylum policies on the destination choices of refugees (Hatton (2005, 2009); Thielemann (2006); Brekke, Roed and Schøne (2016); Bertoli, Brücker and Fernández-Huertas Moraga (2022); Diop-Christensen and Diop (2021)), robust estimates have remained difficult to attain, in part because it is difficult to control for unobserved determinants of migrants' destination choices. In this paper, I address this issue using a simple technique that allows for a much richer set of unobserved determinants of location choices than existing approaches. Specifically, I develop a *relative discrete choice* model that leverages the similarities in (unobserved) characteristics between asylees and non-asylum foreigners within origin country x destination country x year cells. Essentially, if unobserved, time-varying factors that are specific to an origin-destination country pair, such as evolving ethnic networks and the political and economic conditions of the country pair, affect asylees and other migrants equally, my approach controls for all these unobserved differences by modeling the *difference* between asylees' and other migrants' location choices at the cell level. Identification then results from the fact that only the refugees' location choices should be affected by asylum grant rates.

My model combines the estimation approach of Cadena (2013), who develops a discrete choice model of immigrant inflows in response to native born labor supply increases (using pre-reform welfare participation rates as an instrument), with insights from Borjas

(2001)'s seminal paper, which uses the relative geographic distribution of new immigrant inflows versus the incumbent population to show that new immigrants arbitrage interstate wage differentials.[1] In the present context, I use the assumption that refugees but not other immigrants (i.e. those with long term residency permits for family unification, occupation, or education) respond to asylum grant rates to estimate that an increase in the origin-destination pair grant rate by 10 percentage points increases the share of applications lodged in that destination country by 6.5 percent. In contrast to existing literature, this effect is both precisely measured and robust to a wide variety of specification changes.

In addition to leveraging the similarities between asylum applicants and non-asylum foreigners, my model directly accounts for variation in the origin-destination gains in utility from migration, which varies across both origin and destination countries, without affixing these into origin and destination fixed effects, as past work has done. Finally, my model has the added benefit that while structurally derived from a utility-maximizing choice framework, it also has a reduced form interpretation that does not require the structural assumptions to hold. While a number of approaches have been used to address endogeneity concerns in discrete choice models, including the BLP method (Berry, Levinsohn and Pakes (1995)), control functions (Petrin and Train (2010)), instrumental variables (Cadena (2013)), and the multiple indicator solution (Guevara et al. (2020)), the relative choice approach used in this paper provides researchers in the migration and other literatures with a simple, additional approach to addressing such concerns.

[1]In contrast to my approach, which estimates the absolute effect of grant rates on application flows, Borjas (2001) estimates the effect of higher wages on the inflow of new immigrants *relative* to the location choices of prior immigrants and natives. A similar specification is used by Albert and Monras (2018) who examine immigrant location choices in response to local prices *relative* to that of natives, with total consumption modeled as a function of consumption locally and abroad (via remittances).

Section 2 of this paper provides background information and a brief discussion of the relevant literature. I derive the model in Section 3, describe the data in Section 4, and provide estimation results in Section 5. I conclude in Section 6.

3.2 Background

Asylum applicants are individuals who migrate to a foreign country and, upon arrival, request protection in the receiving country under the Refugee (Geneva) Convention of 1951. A core principle of the agreement is non-refoulement: individuals who would face threats to their life or freedoms if forced to return to their origin country should not be turned away and instead should be provided refuge. Prospective applicants must be present in the destination to apply for asylum and are then promptly entered into the application process. Until a decision is made, applicants reside in government facilities where they are provided food and medical care. Upon a grant of asylum, the individual attains refugee status in the destination country. Those denied asylum on the grounds of not meeting the Geneva standard for refugee status may be considered for humanitarian (also called subsidiary) protection, which is a country-specific status with varying definitions and awards. Not all European countries offer humanitarian status, but if either status is granted, the individual may generally reside in the country for five years, after which may adjust their status to long-term resident (European Commission (2020)). They are also afforded the same broad rights as others, such as freedom of mobility to other European Union countries (Brekke and Brochmann (2015); Zorlu and Mulder (2008)). This makes the grant rate a particularly important factor in the decision make process, although migration costs and long processing times would attenuate this.

Despite agreeing to the principles of the Geneva Convention, many countries attempt to deter future applications as part of larger immigration policies to manipulate which types of foreigners arrive and are admitted and which are denied entry. Toshkov and de Haan (2013) show, in fact, that European countries engage in a so-called "race to the bottom" in terms of applications, seeking to receive fewer applications than their neighbors. This is also found by Holzer, Schneider and Widmer (2000), who show that Switzerland enacted policies that deterred future applications between 1986-1995 in response to an uptick in filings. Vink and Meijerink (2003) furthermore find that the variance of the distribution of asylum applications across countries decreased between 1982-2001, a result of the main EU destinations becoming more proactive in deterring future applications with restrictive policies. This suggests a dynamic relationship between the supply of applications and policy responses.

Due to the ease of mobility within Europe, a concern among countries is the extent to which refugees can exploit the application process. To address concerns of asylum "shopping"–the submission of applications to multiple countries–or asylum "hopping"–using entry into one country as a means of entry into a more desirable country for application purposes (such as within the Schengen Area)–most European countries agreed to the Dublin Regulation in 1990, which has seen major updates twice, most recently in 2013. With some exceptions, the core principle of all versions of the agreement stipulate that the country in which a foreigner first arrives is responsible for their asylum application.[2] Thus, any applications lodged outside this initial country are transferred

[2]Exceptions to this include 1) minors applying for asylum, in which case responsibility of the application falls on the country in which the closest family member resides, and 2) applicants who have immediate relatives who have either been previously granted asylum or have an asylum case pending, in which case the country in which they reside is responsible. As discussed in Section 5 and shown in the Appendix, controling for the refugee stock has no impact on baseline model estimates. Therefore, I consider these exceptions to be minimal in impact to the model and I abstract from them in this

to the first country. Furthermore, since all European countries are recognized as safe countries, the decision made in the initial country is final and recognized as such by all other countries in the agreement. Therefore any repeat submission is summarily denied, barring any changes in circumstance to the applicant (European Commission (2020)).

These principles are not always uniformly upheld, however. EU countries near the periphery are more likely to receive applications due to their proximity to sending countries, and especially so during large waves of asylum applications. In 2015, Hungary suspended agreement with the core principle of the Geneva Convention - taking responsibility of applications for first-time entrants - due to the spike in applications from Syria (United Nations High Commissioner for Refugees (2017)). In response, Germany adopted a policy of not turning away Syrian applicants, regardless of whether they've entered via other EU countries. However, these large waves may result in longer processing times and worse conditions for applicants, which would prompt other actions by member countries. For example, in 2011, two EU court rulings barred member countries from transferring applicants back to Greece on human rights grounds, as Greece was in the midst of its own refugee crisis (European Commission (2016)). Although these withdrawals from the principle are not common, they do suggest that core Dublin principles are not ironclad, and individual countries will take matters into their own hands when they are overburdened. The inclusion of dyadic fixed effects in my model addresses such destination-year level changes by member countries. Furthermore, in robustness checks I account for these suspensions of the Geneva principle by excluding the relevant countries altogether and my results remain unchanged.

The most straightforward policy tool a country can use to affect future applications is

discussion.

stricter interpretations of the definition of refugee or humanitarian status, which can be implemented by lowering grant rates. A number of studies have examined this response, but have found wide-ranging estimates. For example, Toshkov (2014), using 29 European countries from 1987-2010, takes a dynamic approach by including lags of the dependent variable in the specification, and finds a small estimate of a 1 percentage point increase in the share of applications received by a country in response to a 15 percentage point increase in the grant rate. He further finds a negative effect of applications on grant rates, although effects are largely limited to within-country variation. However, that analysis is limited to a destination-year analysis due to sample size, and does not account for origin country heterogeneity.[3]

More recently, Bertoli, Brücker and Fernández-Huertas Moraga (2022) estimate a discrete choice model, although not a relative one, that considers processing time of applications as a factor in the destination choice. They find that a 10 percentage point increase in the grant rate is associated with an increase in application shares by 1.7 percent in their unweighted model that is most similar to mine and find no effect in their weighted model. Their inclusion of the interaction of grant rates with processing times increases the estimated effect to 2.8 percent in their weighted model. As I will show Section 3, the interaction of controls with grant rates is important because the *utility gains* (or *losses*, in the case of processing time) at the time of the decision will depend on the grant rate. Hatton (2009) examines the relationship between total applications and grant rates from 1997-2006 and suggests that earlier estimates of grant rate effects may be biased towards zero due to drops in the grant rate deterring those most likely to be rejected. He constructs an index of policy stance based on legislative enactments -

[3] Neumayer (2004) constructs a measure of asylum burden taken on by a destination country and also finds a small but positive effect. However the outcome variable in that analysis is not comparable here.

specifically policies related to applications with manifestly unfounded claims and those with subsidiary/humanitarian status - and uses the composite index as an instrumental variable for grant rates, finding that a 10 percentage point increase in the grant rate increases total applications by 16 percent. Although the use of other policy stances may provide a good predictor of grant rates, they may not satisfy the exclusion restriction required for valid instrumental variables: those legislative enactments should not correlate with applications through channels other than the grant rate. Furthermore, legislative changes generally apply equally to all origin countries, thus potentially omitting the differential discretion that destination countries often exhibit towards foreigners from different origin countries. At the other extreme, Diop-Christensen and Diop (2021) use a Poisson pseudo-maximum likelihood model of total applications with data on on EU15 countries from 2008-2015, and finds a large effect of a lagged 1 percentage point increase from the mean grant rate resulting in a 17.2 percent increase in total applications at the origin-destination level.

To facilitate comparisons among the estimated effects summarized above, suppose that Europe receives a constant flow of 50,000 applications from some particular origin country, and a particular destination country usually receives 20% of these applications. The estimates of the four aforementioned studies suggest that an increase of the grant rate by 10 percentage points for that origin-destination pair would result in 48, 170, 1600, and 17200 more applications, respectively.[4] Under the same circumstances, I find an estimate of 650 more applications in my preferred specification. The range of these estimates suggest that methodology matters. My approach improves on previous approaches in a few

[4]Toshkov (2014)'s specification is at the destination-year level rather than the origin-destination-year level, so the estimated effect is in response to an increase of the grant rate by 10 percentage points from the mean for that destination country (rather than origin-destination pair).

ways. First, as previously noted, I leverage the similarities between asylum applicants and non-asylum foreigners at the origin x destination x year cells so that unobserved characteristics are controlled for, under the identification assumption that the *differences* in these two groups' characteristics (rather than just the asylee population's characteristics) are unrelated to the grant rate. Second, my model directly accounts for variation in the origin-destination *gains* in utility from migration, which varies across both origin and destination countries, without affixing these into separate origin and destination fixed effects, whereas past work have generally just included destination country controls. Finally, my model has the added benefit that although structurally derived from a utility-maximizing choice framework, it also has a reduced form interpretation that does not require the structural assumptions to hold.

Finally, some past work has also examined how the asylee/refugee population compares to immigrants. In the US context, Jaeger (2007) arguably provides the most thorough examination of the location choices of different types of immigrants in the US using Immigration and Naturalization Services (INS) data. The paper provides an abundance of evidence that asylees and refugees are similar to economic immigrants regarding other aspects of their location preferences: both are highly attracted to locations with high wages and locations that are closer to the origin country, and refugees/asylees are slightly more attracted to locations with a large prior immigrant presence than other immigrants, possibly due to greater network attachment upon arrival. Connor (2010) also finds evidence that refugees/asylees have similar levels of employment to immigrants, but lower wages, which is attributed to lower levels of education upon arrival. That study also finds refugees/asylees live in areas with higher percentage foreign born and lower median income, but no statistical difference in terms of unemployment or percent living in poverty relative to other immigrants.

3.3 A Model of Relative Choices

The model is motivated by the possibility that, in the absence of grant rate effects, the asylee population would still choose to live in countries in such a way that correlates with grant rates for unobserved reasons, such as pro- or anti- immigrant sentiment, government enacted policies of immigrant deterrence, or spurious correlation. In such a scenario, a simple OLS regression of applications on grant rates does not constitute a causal relationship, and a valid instrument would be necessary. Finding and justifying such an instrument is often difficult within the immigration literature.

In the absence of such exogenous variation, I provide a method that, although not causal in interpretation, removes much of the unobserved selection and endogeneity concerns in studies of migrant destination choices. I rely on the fact that asylees and other immigrants are likely to share in many of the factors that determine their location choices, such as likelihood of employment, a reliance on home country networks in the destination country, and similar treatment by the native population. The location choices of recently arrived immigrants who are not asylum seekers would provide a close counterfactual because they have no incentive to select on the harshness of the country's grant rates, but still share many of the common observed and unobserved characteristics with asylum seekers at the origin-destination-year level. To illustrate this, consider two Syrian nationals migrating in 2015, both caught under similar circumstances of civil war, and further suppose that the first reaches Europe with the intent to file an asylum application, due to having no other means of adjusting status, while the second applies for a non-asylum permit such as family reunification or employment, due to having family connections or higher skills. If the latter has family in some European country and applies for a family-based permit, such a country is also desirable to the asylum applicant because that country is more likely to have a Syrian community that can provide a safety net and ease the adjustment of life in a new country. If the latter instead applies for an occupational permit, similarly, the country is also desirable to the asylum applicant as it is more likely to offer work opportunity to Syrian nationals in that year.

The assumption that grant rates are uncorrelated with origin-destination-year level characteristics of asylum applicants is too strong to allow for a causal interpretation. For example, after accounting for origin-destination-level variation, grant rates may be correlated with the education level of asylum applicants: destination countries may be broadly more receptive of foreigners with higher education, and this would be reflected in a higher grant rate to those individuals. If this is the case, then the destination country is also likely to favor more educated non-asylum immigrants when granting occupational permits. Differencing this favoritism for the two groups would therefore result in a weaker identification assumption: grant rates are no longer required to be uncorrelated with the education level of asylum applicants, but instead uncorrelated with the *difference* in education levels between asylum applicants and non-asylum immigrants. In the model that follows, I allow for differences to exist between asylum applicants and non-asylum immigrants, and allow for grant rates to be correlated with the characteristics of each type, but rely on the identifying assumption that grant rates are uncorrelated with the differences in asylum-immigrant characteristics, whether they are observed or unobserved.

3.3.1 Specifications

The primary outcome variable will be the relative share of asylum applications to immigrant inflows, measured as awarded permits for residency. Asylees will refer to first-time applicants for asylum (who have already arrived at the destination country, a requirement of requesting asylum) and immigrants will refer to the recently arrived non-asylum foreign born. Adapting from Borjas (2001), let $A_{s,j,t}$ denote the share of applications from origin country s filed in year t that are lodged in destination country j, and let $I_{s,j,t}$ denote the share of the non-asylum inflow of immigrants from origin country s in year t who migrate to destination country j. Therefore,

$$Z_{s,j,t} = \frac{A_{s,j,t}}{I_{s,j,t}} \tag{3.1}$$

is the relative concentration of new asylees to new immigrants in destination j. If $Z_{s,j,t}$ is larger than one, then there are relatively more new asylees concentrated in that country than new immigrants. Let $g_{s,j,t}$ denote the grant rate of applications of origin country s applying in destination j that are decided in year t.

Following the literature, I model the location decisions of both types as depending on four variables: grant rates, labor market expected outcomes, a measure of civil/political rights, and the size of refugee population already in the destination country to capture network effects. Labor market expected outcomes are included based on the abundance of evidence in the literature (Borjas (1987, 2001), Card (2001); Card and Lewis (2007); Cadena (2013); Cadena and Kovak (2016), Basso, D'Amuri and Peri (2019)), that immigrants are attracted to high-employment locations. Civil/political rights are included because, although it is relevant to both asylum applicants and immigrants, it is a distinguishing feature of the asylum application process. The size of the existing refugee population in the destination country is included due to Geneva Convention policy allowing applicants to have their case heard in a country where they have family members who have previously been approved for refugee status, which also accounts for network effects.

3.3.2 Expected Outcomes

Expected outcomes k periods after arriving to the destination country will depend on the probability of being granted asylum, $\mathbb{P}(G_{s,j,t} = 1) = g_{s,j,t}$, which is less than 1 for asylum applicants and mechanically equal to 1 for non-asylum immigrants who have already been given residency permits, and on $V^l_{s,j,t}$ and $V^r_{s,j,t}$, which are the present discounted value of expected labor and rights outcomes, respectively. Let $\delta < 1$ be the discount factor such that expected outcomes can be expressed by:

$$V^l_{s,j,t} = \sum_{k=0}^{T} \delta^k \mathbb{E}(m^{l,k}_{s,j,t} | G_{s,j,t}) \tag{3.2a}$$

$$V^r_{s,j,t} = \sum_{k=0}^{T} \delta^k \mathbb{E}(m^{r,k}_{s,j,t} | G_{s,j,t}) \tag{3.2b}$$

where $m^{l,k}_{s,j,t}$ is the labor market gain from migrating and $m^{r,k}_{s,j,t}$ is the rights gain from migrating. The outcome of application decisions is binary, with “granted” including anyone who may stay (granted refugee or humanitarian status) and “denied” including anyone who is denied both statuses. Expected outcomes can be rewritten as:

$$V^l_{s,j,t} = \sum_{k=0}^{T} \delta^k [g_{s,j,t}\mathbb{E}(l^k_{j,t}) + (1 - g_{s,j,t})\mathbb{E}(l^k_{s,t})] \tag{3.3a}$$

$$V^r_{s,j,t} = \sum_{k=0}^{T} \delta^k [g_{s,j,t}\mathbb{E}(r^k_{j,t}) + (1 - g_{s,j,t})\mathbb{E}(r^k_{s,t})] \tag{3.3b}$$

where $l_{j,t}$ and $l_{s,t}$ are measures of employment outcomes in the destination country and origin country, respectively, and similarly for rights outcomes. Reorganizing results in:

$$V^l_{s,j,t} = \sum_{k=0}^{T} \delta^k \mathbb{E}(l^k_{s,t}) + g_{s,j,t} \sum_{k=0}^{T} \delta^k [\mathbb{E}(l^k_{j,t}) - \mathbb{E}(l^k_{s,t})] = V^{l,d}_{s,t} + g_{s,j,t} V^{l,g}_{s,j,t} \tag{3.4a}$$

$$V^r_{s,j,t} = \sum_{k=0}^{T} \delta^k \mathbb{E}(r^k_{s,t}) + g_{s,j,t} \sum_{k=0}^{T} \delta^k [\mathbb{E}(r^k_{j,t}) - \mathbb{E}(r^k_{s,t})] = V^{r,d}_{s,t} + g_{s,j,t} V^{r,g}_{s,j,t} \tag{3.4b}$$

where superscripts d and g denote denied and granted, respectively. Note that expected outcomes if granted is the value of the destination-origin country difference, or the *gains* from migrating.

3.3.3 Random Utility

Let $U_{i,s,j,t}$ be the utility of individual i from origin s who chooses destination j in year t. The decision rule must be such that the individual chooses destination j if and only if $U_{i,s,j,t} \geq U_{i,s,h,t}\ \forall h \neq j$. I model individual i's utility as being linear in expected outcomes and other observable characteristics at the $s \times j \times t$ level.

$$U_{i,s,j,t} = \gamma V^l_{s,j,t} + \delta V^r_{s,j,t} + \mathbf{X}_{s,j,t}\beta + u_{i,s,j,t}$$

Since the presence of same-country earlier refugees is a feature of just the destination country, I include this as a control in $\mathbf{X}_{s,j,t}$ rather than modeling it as I do labor and rights outcomes. From McFadden (1973), under the assumption that all $u_{i,s,j,t}$ are i.i.d with Type I extreme value distribution, the coefficients can be estimated by the probability choice taking conditional logit form. This assumption is certain to fail because 1) unobserved characteristics shared by asylees from the same origin country will result in correlated errors, and 2) the errors are likely to be correlated with expected outcomes due to, for example, the correlation of employment or rights with unobserved characteristics, inducing omitted variable bias. To overcome this, I follow Cadena (2013) and Scanlon et al. (2002).[5] I model the error term to be separable in the unobserved group term and the unobserved individual term: $u_{i,s,j,t} \equiv \eta_{s,j,t} + e_{i,s,j,t}$, resulting in:

$$U_{i,s,j,t} = \gamma V^{l}_{s,j,t} + \delta V^{r}_{s,j,t} + \mathbf{X}_{s,j,t}\beta + \eta_{s,j,t} + e_{i,s,j,t} \tag{3.5}$$

where the $e_{i,s,j,t}$ are now i.i.d Type I extreme value. Consequently, $\pi_{s,j,t}$ is the probability that the individual chooses destination j, which takes the conditional logit form:

$$\pi_{s,j,t} = \frac{e^{\left(\gamma g_{s,j,t} V^{l,g}_{s,j,t} + \delta g_{s,j,t} V^{r,g}_{s,j,t} + \mathbf{X}_{s,j,t}\beta + \eta_{s,j,t}\right)}}{D_{s,t}} \quad , \quad \text{where:} \tag{3.6}$$
$$D_{s,t} = \sum_{h=1}^{J} e^{\left(\gamma g_{s,h,t} V^{l,g}_{s,h,t} + \delta g_{s,h,t} V^{r,g}_{s,h,t} + \mathbf{X}_{s,h,t}\beta + \eta_{s,h,t}\right)}$$

The terms $V^{l}_{s,j,t}$ and $V^{r}_{s,j,t}$ from equation (3.5) are replaced by $g_{s,j,t}V^{l,g}_{s,j,t}$ and $g_{s,j,t}V^{r,g}_{s,j,t}$ because the first terms in equations (3.4a) and (3.4b) are present in the numerator and every term in the denominator of $\pi_{s,j,t}$ and are therefore canceled, and more intuitively

[5] Scanlon et al. (2002) models the optimal choice of health care plans and Cadena (2013) models immigrant geographic responses to native labor supply shocks.

because these terms are invariant to the destination and do not factor into the choice. The derivation so far also suggests that the control for the present refugee population should be interacted with grant rate. The intuition behind this is that the utility gains from choosing the destination are not realized unless the application is approved, which occurs with probability $g_{s,j,t}$.

$\pi_{s,j,t}$ equals the expected share of applications and inflows into j: $\mathbb{E}(A_{s,j,t}) = \pi^A_{s,j,t}$ and $\mathbb{E}(I_{s,j,t}) = \pi^I_{s,j,t}$. Due to sampling error, the measured shares can be expressed as the choice probability plus an uncorrelated, mean-zero error term:

$$A_{s,j,t} = \pi^A_{s,j,t} + v^A_{s,j,t} \tag{3.7a}$$

$$I_{s,j,t} = \pi^I_{s,j,t} + v^I_{s,j,t} \tag{3.7b}$$

Taking the natural log of both sides and then the Taylor approximation around $(v^A_{s,j,t}, v^I_{s,j,t}) = (0, 0)$ gives:

$$\ln(A_{s,j,t}) \cong \ln(\pi^A_{s,j,t}) + \frac{v^A_{s,j,t}}{\pi^A_{s,j,t}} \tag{3.8a}$$

$$\ln(I_{s,j,t}) \cong \ln(\pi^I_{s,j,t}) + \frac{v^I_{s,j,t}}{\pi^I_{s,j,t}} \tag{3.8b}$$

which can be subtracted so that the difference equals the log of the relative concentration measure:

$$\ln(A_{s,j,t}) - \ln(I_{s,j,t}) = \ln(Z_{s,j,t}) \tag{3.9}$$

Finally, plugging in (3.6) and pulling out the control for the refugee network, $S_{s,j,t}$, interacted with grant rate gives the following 3 equations:

$$\begin{aligned}\ln(A_{s,j,t}) \cong\ & \gamma^A g_{s,j,t} V^{l,g}_{s,j,t} + \delta^A g_{s,j,t} V^{r,g}_{s,j,t} + \psi^A g_{s,j,t} S_{s,j,t} \\ & + \mathbf{X}_{s,j,t}\beta^A + \eta^A_{s,j,t} - \ln(D^A_{s,j,t}) + \frac{\upsilon^A_{s,j,t}}{\pi^A_{s,j,t}} \end{aligned} \tag{3.10a}$$

$$\begin{aligned}\ln(I_{s,j,t}) \cong\ & \gamma^I g_{s,j,t} V^{l,g}_{s,j,t} + \delta^I g_{s,j,t} V^{r,g}_{s,j,t} + \psi^I g_{s,j,t} S_{s,j,t} \\ & + \mathbf{X}_{s,j,t}\beta^I + \eta^I_{s,j,t} - \ln(D^I_{s,j,t}) + \frac{\upsilon^I_{s,j,t}}{\pi^I_{s,j,t}} \end{aligned} \tag{3.10b}$$

$$\begin{aligned}\ln(Z_{s,j,t}) \cong\ & \gamma g_{s,j,t} V^{l,g}_{s,j,t} + \delta g_{s,j,t} V^{r,g}_{s,j,t} + \psi g_{s,j,t} S_{s,j,t} \\ & + \mathbf{X}_{s,j,t}\beta + \eta^A_{s,j,t} - \eta^I_{s,j,t} - \ln\left(\frac{D^A_{s,j,t}}{D^I_{s,j,t}}\right) + \frac{\upsilon^A_{s,j,t}}{\pi^A_{s,j,t}} - \frac{\upsilon^I_{s,j,t}}{\pi^I_{s,j,t}} \end{aligned} \tag{3.11}$$

Equations (3.10a) and (3.10b) are the individual discrete choice models. Equation (3.11) is the relative model, which has a reduced form interpretation: each coefficient represents the effect of a change in one unit of the explanatory variable on the percentage change of the relative concentration of asylum applications to immigrant inflows. With the primary explanatory variable of interest being the grant rate rather than the grant rate interaction, the flexibility of random utility models allows for the inclusion of $g_{s,j,t}$ within $\mathbf{X}_{s,j,t}$, so that β^A becomes the coefficient of interest. While the $\upsilon_{s,j,t}$ terms arise due to sampling error and are exogenous, the main obstacle to identification in all three equations are the unobserved common group level variables, $\eta^A_{s,j,t}$ and $\eta^I_{s,j,t}$.

There are a number of advantages to using the relative model rather than the individual model of equation (3.10a). First, the coefficients are additive due to the logged dependent variable, so that $\beta = \beta^A - \beta^I$. Second, since asylees have an incentive to respond positively to grant rates but immigrants do not, we expect $\beta^A > 0$ but $\beta^I = 0$, so that $\beta = \beta^A$. This holds because any correlation between grant rates and unobserved immigrant characteristics is captured in $\eta^I_{s,j,t}$ of (3.11), rendering β^I as the true response of immigrants to grant rates. Finally, and critically, the identifying assumption of the relative model is that grant rates are uncorrelated with the *difference* in asylee-immigrant characteristics rather than uncorrelated with unobserved asylee characteristics. This is a weaker assumption because it allows for the case where there exists correlation that cannot be controlled for. In the previous example at the beginning of the section, I used education to highlight this: grant rates are not allowed to be correlated with the education levels of asylees or immigrants in the individual models (3.10a) and (3.10b) to consistently estimate β^A and β^I, however, grant rates are allowed to be correlated with education in the relative model, so long as the correlation is the same for asylees and immigrants. This can be seen from the fact that $cov(\eta^A_{s,j,t}, g_{s,j,t}) \neq 0$ will bias the estimation of β^A in (3.10a), however, it will not bias the estimate of $\beta = \beta^A$ in (3.11) under the weaker assumption that $cov(\eta^A_{s,j,t}, g_{s,j,t}) = cov(\eta^I_{s,j,t}, g_{s,j,t})$.

3.4 Data and Descriptives

European Commission (2021) and the United Nations High Commissioner for Refugee, United Nations High Commissioner for Refugees (2021), are the two main data sources for applications and decisions used in past work. The main source for non-asylum immigrant inflows is Eurostat. Both acquire data from each receiving nation's foreign ministry. Results remain unchanged from using either data sources, however, to maintain consistency in data sources, I rely on Eurostat for all three variables, which are used to construct the primary response variable and calculate grant rates. In robustness checks, where I include specifications that include the US, I add US data from UNHCR. The World Bank (2021) provides annual data on a number of economic conditions on all countries, and these are used in models that include controls in the appendix. The World Bank also provides data on civil and human rights indices under the Worldwide Governance Project. Finally, I rely on an additional measure of civil/human rights using Economist Intelligence Unit (2021)'s Democracy Index, obtained from each year's report.

Counts of applications are restricted to first time applicants for asylum lodged to a destination country plus all subsidiary beneficiaries under the same application. There are 3 decision categories: those granted asylum (Geneva) status, those denied asylum but granted humanitarian status, and those denied either status. Unless noted otherwise, I consider a positive or accepted grant as one where the decision is either a grant of asylum or humanitarian status.

Data for foreign inflows is measured by the number of permanent (long-term, at least 12 months) permits to reside in the destination country. There are four categories for such permits: educational, family, occupational, or other. Due to variation in destination country definitions, the 'other' category may or may not include positive decisions made for asylum applicants. I use the former 3 categories to measure the count of long-term permits awarded to non-asylum immigrants into Europe.[6] One concern with this measure may be that it is not comparable to applications of asylum, as it is a measure of *granted* permits rather than applications for permits. However, I argue that this specification is preferred to applications for permits for three reasons. First, foreigners may apply to a number of destination countries for non-asylum reasons - asylum applicants cannot -

[6]In the appendix, I provide estimates with individual specifications of permits, with results remaining unchanged for occupational and family permits, but increasing for educational permits

making a count of applications less representative of their first choice preference. Second, asylees must be present in the destination country in order to apply for asylum, and since they are not allowed to apply for asylum in another Geneva-agreement country after initial arrival, a more comparable measure for immigrants is the actual award of a permit. Although no official reports on the processing delay between applying and being awarded for long term residence were found, some government websites suggest that most applicants generally take roughly 3 months[7][8], and any mismeasurement due to this delay is unlikely to be correlated with the response variable.

Table 3.1a provides descriptive statistics on applications, grant rates, and inflows for the top 10 destination countries in the sample, accounting for 85.3% of all applications lodged and 81.0% of immigrant (occupational, family, educational) inflows during the 2008-2019 period. Almost a third of applications we lodged in Germany, with the next highest receiving country, France, receiving 12.3% of applications. Thus, central European countries that are not geographically closest make up a significant share of received applications. Table 3.1b provides the same statistics for origin countries, and shows that the top 10 asylum sending countries account for nearly 60% of all asylum applications lodged but just under 17% of immigrant permits. The grant rate for these countries varies considerably and overall stands at 61.1%, however, when excluding Syria the remaining nine countries have an overall grant rate of 47.9%.

The baseline measure capturing labor market conditions is the employment to population ratio because other reasonable measures, such as wages and/or unemployment rate, are more sensitive to equilibrium conditions and because data for those variables

[7]For Germany, 1-3 months: https://www.germany.info/us-en/service/visa/residence-visa/922288 (website accessed March 3, 2022)

[8]For Finland, 3 months for occupational permit: https://migri.fi/en/processing-times (website accessed March 3, 2022)

Table 3.1a: Descriptive Statistics, Top 10 Destinations, 2008-2019, (Thousands)

Country	Applications	App Share (%)	Immigrant Inflows	Inflow Share (%)	Geneva Grants	Total Grants	Total Decisions	Grant Rate (%)
Germany	2,074.5	31.5	1,095.0	8.5	1,048.4	1,139.5	2,121.4	53.7
France	812.0	12.3	1,763.2	13.7	195.6	195.6	835.2	23.4
Italy	616.1	9.4	1,605.4	12.5	230.5	342.6	698.7	49.0
Sweden	480.1	7.3	559.5	4.3	218.9	230.3	445.3	51.7
United Kingdom	382.3	5.8	2,845.7	22.1	107.9	124.8	334.6	37.3
Greece	290.0	4.4	230.7	1.8	52.1	52.3	202.1	25.9
Hungary	263.8	4.0	151.2	1.2	4.6	5.1	29.8	17.2
Spain	246.8	3.7	1,806.5	14.0	57.9	93.2	153.8	60.6
Switzerland	230.8	3.5	121.3	0.9	116.2	164.9	236.6	69.7
Belgium	220.7	3.4	264.5	2.1	85.1	85.1	227.3	37.4
Total (Top 10)	5,617.0	85.3	10,443.0	81.0	2,117.1	2,433.4	5,284.6	46.0
Total	6,585.3	100.0	12,890.7	100.0	2,521.8	2,877.6	6,168.6	46.6

Counts, in thousands, are displayed for total applications lodged (2), number of immigrant (occupational, family, educational) permits issued (4), number of asylum (Geneva) grants (6), total grants that include humanitarian status (7), and total decisions made on applications (8). Columns 3 and 5 represents the share of applications and inflows out of all destination countries, and column 9 is the overall grant rate, calculated as total grants out of total decisions, over the 2008-2019 period.

Table 3.1b: Descriptive Statistics, Top 10 Origins, 2008-2019, (Thousands)

Country	Applications	App Share (%)	Immigrant Outflows	Outflow Share (%)	Geneva Grants	Total Grants	Total Decisions	Grant Rate (%)
Syria	1,195.5	18.2	239.9	1.9	973.0	991.7	1,053.7	94.1
Afghanistan	681.7	10.4	68.0	0.5	285.3	373.0	644.9	57.8
Iraq	482.1	7.3	124.3	1.0	242.1	259.4	452.3	57.4
Pakistan	266.6	4.0	448.4	3.5	35.6	46.2	246.5	18.7
Eritrea	261.2	4.0	46.7	0.4	200.9	209.8	243.4	86.2
Nigeria	248.6	3.8	249.3	1.9	44.4	71.0	254.7	27.9
Russia	201.2	3.1	427.0	3.3	43.3	49.0	199.3	24.6
Albania	201.2	3.1	357.7	2.8	9.9	12.4	188.4	6.6
Iran	199.9	3.0	114.7	0.9	85.3	89.8	181.0	49.6
Somalia	197.2	3.0	70.9	0.5	123.4	143.8	209.9	68.5
Total (Top 10)	3,935.3	59.8	2,146.8	16.7	2,043.2	2,246.0	3,674.2	61.1
Total	6,585.3	100.0	12,890.7	100.0	2,521.8	2,877.6	6,168.6	46.6

Counts, in thousands, are displayed for total applications lodged (2), number of immigrant (occupational, family, educational) permits issued (4), number of asylum (Geneva) grants (6), total grants that include humanitarian status (7), and total decisions made on applications (8). Columns 3 and 5 represents the share of applications and outflows out of all origin countries, and column 9 is the overall grant rate, calculated as total grants out of total decisions, over the 2008-2019 period.

are either not available (wages) or are less complete for origin countries (unemployment rate). However, in specifications with controls, I include the unemployment rate for the

Table 1c: Descriptive Statistics, Controls, 2008-2019

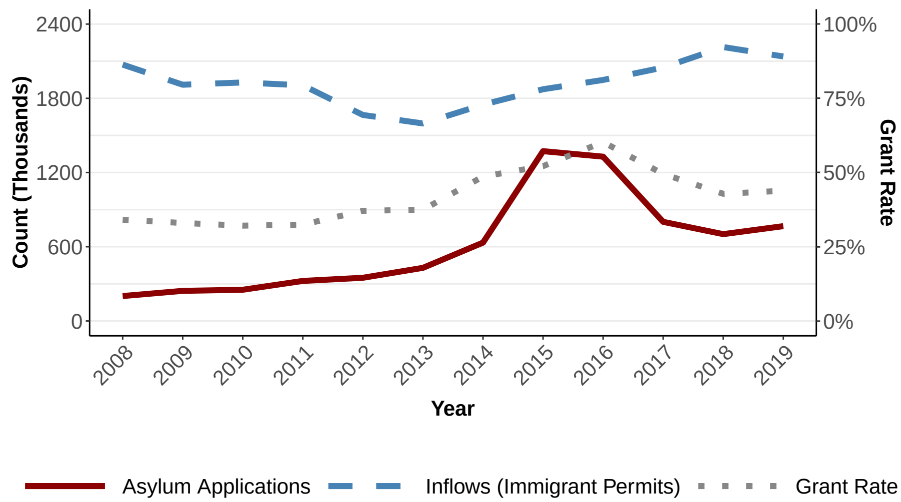
Variable	Origin	Destination
Employment to population ratio, 15+, total (%)	58.07	54.54
Employment to population ratio, 15+, male (%)	68.96	60.68
GDP growth (annual %)	2.73	1.04
GDP per capita (current US$)	12,277	41,424
GNI growth (annual %)	3.19	1.23
Household consumption growth (annual %)	3.38	0.74
Household consumption pc (constant 2015 US$)	5,060	18,150
Labor force participation rate, total, (15-64), (%)	66.47	72.99
Unemployment, total (%)	7.74	8.23
Unemployment, male (%)	7.01	8.34
Control of Corruption (CC)	2.85	6.87
Government Effectiveness (GE)	3.00	6.73
Political Stability/Nonviolence (PV)	4.97	8.01
Rule of Law (RL)	4.24	8.03
Regulatory Quality (RQ)	4.76	7.82
Voice and Accountability (VA)	4.10	7.68
Democracy Index	4.88	8.11

The top section provides the mean of economic measures across countries over the entire 2008-2019 period. The bottom section provides the mean index of different measures of civil/political rights, across countries over the entire 2008-2019 period.

15-64 population and other economic controls, such as GDP growth, labor force participation rate, and household consumption. Furthermore, past work such as Cadena (2013) show that foreigners are more sensitive to employment probabilities than wages.

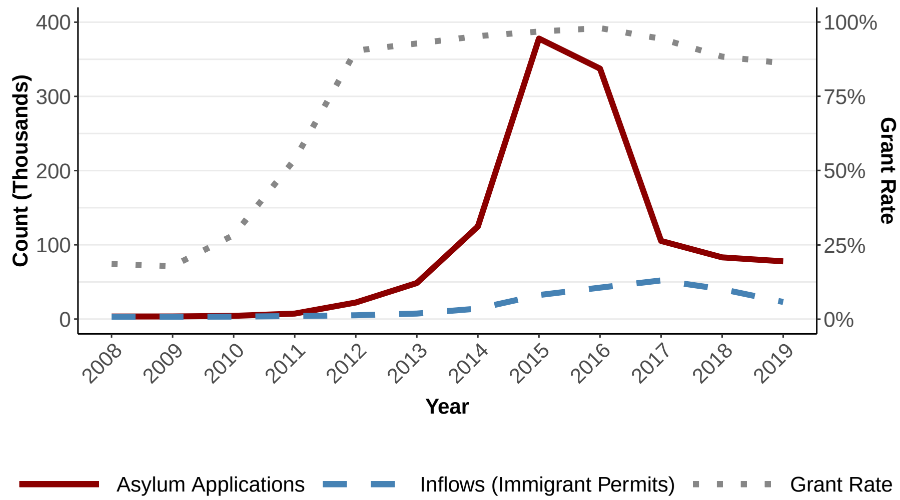
Measures of civil/political rights are obtained from the World Bank's Worldwide Governance Project, which provides six annual indicators: Control of Corruption (CC), Government Effectiveness (GE), Political Stability/Nonviolence (PV), Rule of Law (RL), Regulatory Quality (RQ), and Voice and Accountability (VA). The indicators are composite indices based on a number of data sources that include surveys of households and firms, commercial business information providers, non-governmental organizations, and public sector organizations. Among these sources is the Economist's Intelligence Unit's measure of democratic freedoms, called Democracy Index, which I use as the primary measure of rights. Specifically, it is a weighted average based on answers to a number of public opinion surveys from the respective countries. Although the indicators are not perfect measures, they are plausibly valid annual measures of the state of the social, civil, and political conditions and freedoms in each country. All six indicators are rescaled to range from 0 to 10.[9] Descriptive statistics on the explanatory variables for destination and origin countries are provided in Table 1c.

Figure I: Applications, Immigrant Inflows, and Grant Rates, 2008-2019



[9]Democracy Index already ranges from 0 to 10; the other six indices range from 0 to 1.

Figure II: Syria Trends, 2008-2019



The sample to be used in the baseline specification includes 93 origin countries and 30 destination countries during 2008-2019, which comes after limiting observations to origin countries that are responsible for a cumulative total of 2000 applications across all destinations and years. I also restrict to origin-destination-year cells with at least 10 applications and 10 non-asylum immigrant permits, and where the lagged grant rate is calculated based on at least 10 decisions, as is described in detail in Section 5.1. Due to missing origin-destination-year cells, total observations are reduced to 7,132, with slightly more or fewer in alternate specifications.

Figure I displays trends in annual asylum applications, foreign inflows, and grant rates across all destination countries, and highlights the spike in applications between 2015-2016 due to the Syrian civil war. Although applications have dropped afterwards, they did not returned to pre-2015 levels, and the figure generally suggests a rising trends despite the crisis. This is also illustrated in Figure II, which displays the same trends for Syrian nationals to Europe. Syria led all countries in asylum applications lodged in Europe in 2015, accounting for a third of all applications. To ensure that results are not driven by the leading countries of origin or destination, I provide results where I remove Germany and Syria for the 2015-2016 period from the data, finding similar results.

3.5 Model Estimation

3.5.1 Grant Rates

The grant rate will be the share of decisions that were approved in some period prior to year t. It is not immediately clear if the most immediate prior period should be used nor what the length of this period should be. For example, if applicants update their information from networks on changes in grant rates of the destination country relatively quickly and make their decisions based on only recent information, using a simple lagged grant rate would best capture grant rate effects. Table 3.2, provides coefficient estimates of grant rates in identical models of the main outcome variable, but with specifications of grant rate that vary by period duration (1-5 years) and lags (1-5 years). All 3 dyadic fixed effects (origin x year, destination x year, and origin x destination) are included and no controls are included. As an example, the 3-year grant rate with a 1-year lag specification would associate outcome observations in 2018 with the grant rate calculated over decisions made over the 2015-2017 period. Additionally, observations where the grant rate is calculated based on fewer than 10 decisions of the period are removed, as are observations with fewer than 10 applications or permits. In the last column, I provide a 2019 cross-sectional model based on decisions over the entire 2008-2018 period, with origin and destination fixed effects.

Table 3.2: Models of Different Specifications of Grant Rates

	Duration of Calculated Grant Rate					
	1-Year	2-Year	3-Year	4-Year	5-Year	2019 CS
1 Lag	0.567***	0.715***	0.743***	0.969***	1.268***	0.673
	(0.140)	(0.162)	(0.187)	(0.211)	(0.259)	(0.603)
2 Lags	0.481***	0.511***	0.633***	0.751***	0.821***	
	(0.143)	(0.162)	(0.176)	(0.206)	(0.250)	
3 Lags	0.150	0.229	0.278	−0.029	−0.085	
	(0.154)	(0.167)	(0.194)	(0.229)	(0.283)	
4 Lags	0.069	0.144	−0.143	−0.115		
	(0.150)	(0.169)	(0.197)	(0.239)		
5 Lags	0.139	−0.281	−0.108			
	(0.144)	(0.178)	(0.221)			

Note: Table provides the coefficient estimates of different specifications of grant rate in identical models. Each model is a regression of the logged relative concentration measure on the grant rate, with origin x time, destination x time, and origin x destination fixed effects. Specifications differ by the duration for which the grant rate is calculated (columns) and by number of lags (rows). The final column provides the estimate using just a 2019 cross-section, where the grant rate is calculated as total granted divided by total decisions over the entire 2008-2018 period, with origin and destination fixed effects.

The table displays two findings. First, specifications with greater than 2 lags remove grant rate effects, so that the most recent few periods are essential to any specification, as is expected. This suggests that information on the grant rate leniency travels quickly to origin country prospective applicants, which is reinforced by the strong positive coefficient for the 1-year, 1-lag specification. The window length is less clear: a longer horizon increases the effect, but also results in the specification losing precision as time variation is removed when widening the window. In the extreme, the 2019 cross section provides a positive coefficient, but is much less precise. In order to balance the possibility of applicants relying on a longer horizon of historical rates while also retaining time variation, I use the 3-year, 1-lag specification in baseline estimates. I provide estimates of baseline models with different specifications of grant rate in appendix Table C.1.2, with results remaining consistent with those in Table 3.2.

3.5.2 Model Estimation

I estimate the following models:

$$\begin{aligned} \ln A_{s,j,t} &\cong \alpha^A + \beta^A g_{s,j,t} + \gamma^A g_{s,j,t} E:P_{s,j,t} + \delta^A g_{s,j,t}\text{Rights}_{s,j,t} \\ &\quad + \psi^A g_{s,j,t}\text{Network}_{s,j,t} + \phi_{s,t} + \phi_{s,j} + \phi_{j,t} + e^A_{s,j,t} \end{aligned} \quad (3.12)$$

$$\begin{aligned} \ln I_{s,j,t} &\cong \alpha^I + \beta^I g_{s,j,t} + \gamma^I g_{s,j,t} E:P_{s,j,t} + \delta^I g_{s,j,t}\text{Rights}_{s,j,t} \\ &\quad + \psi^I g_{s,j,t}\text{Network}_{s,j,t} + \phi_{s,t} + \phi_{s,j} + \phi_{j,t} + e^I_{s,j,t} \end{aligned} \quad (3.13)$$

where the outcome variable and grant rates are calculated directly from the data; E:P is the difference in destination-origin employment to population ratio and is the proxy for labor outcomes; Rights is the difference in destination-origin Democracy Index; Network captures the size of the existing refugee population and is calculated as the log of 1 plus the accumulated granted applications since 2008 for each origin-destination pair; and the ϕ terms are origin $\times$ year (capturing the $ln(D^A/D^I)$ term), origin $\times$ destination, and destination $\times$ year fixed effects. Other than the grant rate, all other covariates in all specifications are lagged by 1 year and demeaned at the origin-year level to capture the response from the mean value of the covariate at the place and time of the decision. The model for applications is weighed by the stock of applications at the origin-year level and the model for immigrant inflows is weighed by the stock of inflows at the origin-year level. Weights are necessary in order to ensure that individual applications (and permits) have the same weight across origin-year groups, and to account for potential heteroskedasticity.[10]

[10] Specifically, if total applications lodged for a particular origin-year group are larger, then each application would count for less if origin-year cells are weighed equally (i.e. an unweighted model)

Table 3.3: Baseline Results

	Dependent Var: log(Asylum Applications Share/Foreign Inflow Share)							
	GR-Only	Baseline	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Grant Rate	0.743***	0.655***	0.174*	0.312*	0.605***	0.780***	0.800***	1.468***
	(0.187)	(0.168)	(0.102)	(0.163)	(0.165)	(0.185)	(0.175)	(0.320)
GR x E:P (15-64)		-0.111***	-0.058**	-0.043	-0.066**			-0.047
		(0.035)	(0.024)	(0.039)	(0.026)			(0.087)
GR x Rights		0.584*	0.474**	0.389		0.187		-0.983
		(0.312)	(0.198)	(0.288)		(0.231)		(0.790)
GR x Network		0.311***	0.172***	0.217***			0.345***	0.423***
		(0.072)	(0.048)	(0.072)			(0.071)	(0.124)
E:P (15-64)								-0.040
								(0.027)
Rights								1.294***
								(0.217)
Network								0.265***
								(0.073)
N	7,132	7,132	7,132	7,132	7,540	7,132	7,654	7,132
R^2 Apps	0.954	0.956	0.928		0.953	0.954	0.953	0.895
R^2 Inflows	0.98	0.98	0.975		0.979	0.98	0.979	0.963
R^2				0.907				
FE: Org x Year	X	X	X	X	X	X	X	X
FE: Org x Dest	X	X	X	X	X	X	X	X
FE: Dest x Year	X	X	X	X	X	X		
Weighted	X	X		X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Note: Models are weighted by the origin-year level total applications lodged and immigrant inflows across destinations (see Equations 3.12 and 3.13). Standard errors are heteroskedasticity-robust. Explanatory variables are demeaned at the origin-year level. Grant Rate uses the 3 year, 1 lag specification, and E:P (15-64) and Rights refer to the difference in the employment to population ratio (destination - origin)for those 15 to 64 and the difference in rights index (destination - origin), respectively, which are both lagged 1 year. Network is measured as log(1 + accumulated stock of granted applications since 2008). Destination x Year fixed effects are omitted in the last column to avoid collinearity.

The relative model is not calculated directly due to the use of different weights for the different populations. Instead, the relative model estimates are calculated from estimates of these two models. The estimated coefficients in the relative models represent the difference in estimated coefficients from the individual models. The standard errors represent the standard error of the difference of the coefficients, which is found by first obtaining the variance covariance matrix of coefficients using the residuals of the two individual models, then directly calculating the standard error using also the individual models' standard errors. I also report the R^2 of each model.

Table 3.3 displays model estimates using just the grant rate, the baseline specification, and alternate specifications. The coefficient estimate of grant rate is positive and significant across all models. Model 1 provides the unweighted model, with the estimate remaining significant only at the 10% level and smaller than the baseline, suggesting that high-application years may be driving some of the results. To ensure that each application counts the same in the estimation, however, the weighted model is preferred. The exclusion of different covariates from the baseline model does little to alter the existing coefficients estimates (Models GR-Only, 2, 3, and 4), but confirms that non-asylum immigrants have stronger preferences for employment whereas asylum applicants have stronger preferences for civil and political rights and are more likely to apply to countries with a large existing refugee population from the same origin. In Model 5, I include the individual (non-interacted) controls while omitting destination-year fixed effects, which provide the least explanatory power of the three dyadic fixed effects, in order to avoid collinearity. Their omission, however, would fail to capture variation due to sweeping changes in asylum policy by destination countries, such as those previously highlighted or any others that differentially impact asylum applicants but not non-asylum immigrants. In Appendix Table C.1.1, I provide estimated specifications with the addition of different controls, such as GDP, household consumption, and alternative measures of labor market outcomes and civil/political rights in the appendix.

As supporting evidence for my approach, I also provide estimates of the individual models separately, without invoking the assumption that the coefficient on grant rates for immigrant inflows is zero. In Table 3.4, I show that immigrant permits are not responsive to grant rates; thus my estimated relative effect is nearly identical to the absolute effect derived from a more simpler estimation approach, but much more robust.

Table 3.4: Individual vs. Relative Models

	Dependent Var: log Share					
	Apps	Inflows	Ratio	Apps	Inflows	Ratio
Grant Rate	0.737***	-0.006	0.743***	0.656***	0.001	0.655***
	(0.174)	(0.075)	(0.187)	(0.152)	(0.074)	(0.168)
GR x E:P (15-64)				-0.138***	-0.026	-0.111***
				(0.031)	(0.016)	(0.035)
GR x Rights				0.911***	0.327**	0.584*
				(0.270)	(0.141)	(0.312)
GR x Network				0.215***	-0.096***	0.311***
				(0.062)	(0.035)	(0.072)
N	7,132	7,132	7,132	7,132	7,132	7,132
R^2 Apps	0.954		0.954	0.956		0.956
R^2 Inflows		0.980	0.980		0.980	0.980
FE: Org x Year	X	X	X	X	X	X
FE: Org x Dest	X	X	X	X	X	X
FE: Dest x Year	X	X	X	X	X	X
Weighted	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Note: Apps and Inflows models are weighted by the origin-year level total applications lodged and immigrant inflows, respectively. Standard errors are heteroskedasticity-robust. Explanatory variables are demeaned at the origin-year level. Grant Rate uses the 3 year, 1 lag specification. E:P (15-64) and Rights refer to the difference (destination - origin) in the employment to population ratio for those 15 to 64 and rights index, respectively, which are both lagged 1 year. Network is measured as log(1 + accumulated stock of granted applications since 2008).

3.6 Discussion

Asylum policy is a major component of most western countries' immigration policies. In this paper, I examine the destination choices of individuals applying for asylum in Europe, and how this is impacted by the grant rate. Although the effect is unambiguously positive, it is of policy interest to destination counties to know the response of applications to grant rates. Using a *relative* discrete choice model, I show that an increase of 10 pp in the origin-destination-year level grant rate increases the share of applications to Europe that a particular country receives by 6.5 percent. These results are robust to a number of specifications, and robustness checks generally result in point estimates varying from a 3 to 8 percent estimate. To the best of my knowledge, the method used is novel and has potentially wide appeal: when utilizing discrete choice models, the researcher can rely on the choices of a counterfactual group that is similar to the population of interest other than in the incentives to respond to the explanatory variable of interest. If one group is expected to have a response while the other is expected to have no response, this relative discrete choice model may provide a simple and convenient method to attain nearly-causal estimates on the responding group.

This approach can potentially be applied to other settings where traditional tools for identification, such as instrumental variables, may not be feasible. Within migration, for example, the relative choices of immigrants from similar backgrounds (such as those originally from the same region of the world) but who receive differential treatment by the native population may shed light into the effects of existing perceptions and attitudes. In another context, the choice of health insurance plans of workers can be examined by comparing the choices of union and non-union workers, in settings where union workers may not be responsive to particular aspects of the plan due to union membership.[11] In educational settings, the choice of schools and majors between international students and domestic students can be compared in settings where state/national policies induce a non-responsive for one group. Other classic settings include choice of occupation and industry, mode of transport, product brand, etc. In any case, the researcher must make a reasonable judgement that the counterfactual group is not responsive to the explanatory variable of interest but otherwise similar such that the explanatory variable is uncorrelated with the differences in characteristics between the two groups.

Examining grant rate effects in the American context would provide additional information to governments on how responsive prospective applicants are when making the decision to emigrant and apply for asylum. The US context notably differs from the European context in that prospective applicants are allowed entry into the US, and then the applicant is eligible to apply for asylum for a period of 1 year upon arrival. During this year, prospective applicants are free to migrate anywhere within the US, which has been well-documented to have significant variation in grant rates across asylum offices and immigration courts. Comparing results with those in this paper may lead to novel findings on the effects of policy on mobility and migration.

[11] This is similar to the setting in Scanlon et al. (2002), although union workers in that setting were expected to have a positive response.

Chapter 1 Appendix

A.1 Tables

Table A.1.1: Timeline of Key Events

Jan 30	WHO declares international emergency	**Mar 11**	WHO declares COVID-19 a pandemic
Jan 31	US places travel restrictions on travel from China	**Mar 12**	First covid restriction in US takes effect; NY closes public schools. Announcement made March 10
Feb 3	US declares a public health emergency	**Mar 13**	Trump declares national emergency, announces travel restrictions on 26 countries
Feb 11	WHO announces official name "COVID-19"	**Mar 15**	First statewide bans on public events (NY, OH)
Mar 6	US Secretary of State uses "Wuhan virus" on Twitter	**Mar 16**	Trump's first "Chinese virus" tweet

Note: Timeline events obtained from Cao, Lindo and Zhong (2022) and National Public Radio (2020).

Table A.1.2: Cumulative Covid Cases (Deaths) Worldwide

Date	US	China	Asia	Latin America	Europe	All Others
Jan 20	1 (0)	202 (4)	8 (0)	0 (0)	7 (0)	0 (0)
Jan 31	8 (0)	9,720 (213)	86 (0)	0 (0)	24 (0)	20 (0)
Feb 15	23 (0)	66,576 (1,663)	225 (2)	0 (0)	86 (4)	328 (0)
Feb 29	69 (1)	79,389 (2,838)	3,363 (22)	6 (0)	1,269 (28)	1,306 (49)
Mar 8	464 (19)	80,859 (3,101)	7,967 (58)	130 (2)	10,448 (278)	7,880 (215)
Mar 15	3,929 (69)	81,048 (3,204)	9,977 (115)	742 (15)	53,885 (1,978)	16,717 (760)
Mar 31	173,143 (3,327)	82,545 (3,314)	21,893 (533)	18,402 (572)	450,099 (32,538)	81,561 (3,587)

Note: Counts obtained from World Health Organization (2023). Asia includes eastern and southern Asian countries except China. Latin America includes central and South America and Caribbean countries. Europe includes western and eastern European countries and Russia, and excludes Turkey and Middle Eastern countries. Other includes all other excluded countries and regions (Africa, Turkey, Middle East, the Caucasus, Pacific Islands, Australia, New Zealand, and Canada).

Table A.1.3: Consumer Discrimination, by Dine Share Group

Reference Group:	Dependent Variable: Log Visits, by Dine Share (%)					
All Other Cuisines	0-20	20-30	30-40	40-50	50-60	60-100
Post x 2020 x Chinese	-0.011	-0.008	-0.029***	-0.036***	-0.037***	-0.046***
	(0.016)	(0.010)	(0.009)	(0.010)	(0.012)	(0.013)
Post x 2020	-0.149***	-0.129***	-0.120***	-0.123***	-0.130***	-0.163***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Cases x Chinese	-0.065	-0.063**	-0.035*	-0.062***	-0.069**	-0.103***
	(0.047)	(0.025)	(0.021)	(0.024)	(0.027)	(0.030)
Cases	-0.076***	-0.079***	-0.080***	-0.085***	-0.082***	-0.082***
	(0.010)	(0.006)	(0.006)	(0.007)	(0.008)	(0.008)
Total Obs.	1,290,730	2,228,781	2,213,345	1,773,775	1,361,854	1,575,642
R Squared	0.166	0.143	0.120	0.126	0.163	0.235
Counties	2,478	2,617	2,649	2,597	2,614	2,628
Total Restaurants	46,155	79,695	79,151	63,435	48,698	56,324
Chinese Restaurants	1,962	4,878	6,399	4,912	3,276	2,688
FE: County x Cuisine	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.3 representing estimation of the partitioned DDD model, separately for six partitions of dine-share. The reference group is the pooled renaming cuisine groups for all models. The full set of interactions of indicators for Post, Year, and Chinese are included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

Table A.1.4: Consumer Discrimination, Baseline Framework

	Dependent Variable: Log Visits							
	Chinese Restaurants (Ref = All Other Cuisines)				Cuisine (Ref = American Restaurants)			
	No Controls	Controls	Not Demeaned	Preferred	Chinese	Asian	Latin	European
Post x 2020 x Cuisine x Dine Share	-0.103***	-0.106***	-0.085***	-0.085***	-0.079***	-0.057**	-0.026	-0.018
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.026)	(0.018)	(0.029)
Post x 2020 x Cuisine	-0.020***	-0.024***	0.010	-0.024***	-0.008	0.077***	0.038***	0.022***
	(0.004)	(0.005)	(0.013)	(0.005)	(0.005)	(0.005)	(0.003)	(0.005)
Post x 2020 x Dine Share	-0.004	-0.017**	-0.019***	-0.019***	-0.025***	-0.025***	-0.025***	-0.025***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Post x 2020	-0.134***	-0.134***	-0.127***	-0.134***	-0.150***	-0.150***	-0.150***	-0.150***
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine x Dine Share			-0.414***	-0.414***	-0.464***	-0.349***	-0.008	-0.288***
			(0.069)	(0.069)	(0.071)	(0.051)	(0.043)	(0.076)
Cases x Cuisine	-0.054***	-0.062***	0.108***	-0.056***	-0.061***	-0.027***	0.014**	-0.032***
	(0.011)	(0.011)	(0.031)	(0.011)	(0.012)	(0.009)	(0.007)	(0.012)
Cases x Dine Share			0.058***	0.058***	0.108***	0.108***	0.108***	0.108***
			(0.018)	(0.018)	(0.024)	(0.024)	(0.024)	(0.024)
Cases	-0.105***	-0.081***	-0.103***	-0.080***	-0.075***	-0.075***	-0.075***	-0.075***
	(0.003)	(0.003)	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Total Obs.	10,444,127	10,444,127	10,444,127	10,444,127	6,796,847	7,010,617	8,113,155	6,891,206
R Squared	0.096	0.096	0.096	0.096	0.086	0.085	0.073	0.086
Counties	3,008	3,008	3,008	3,008	2,996	2,995	3,005	2,997
Total Restaurants	373,458	373,458	373,458	373,458	243,021	250,671	290,085	246,399
Cuisine Restaurants	24,115	24,115	24,115	24,115	24,115	31,765	71,179	27,493
FE: County x Cuisine	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year		X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework. The reference group in columns 1-4 includes all non-Chinese cuisines. The reference group in columns 5-8 is American restaurants. The full set of interactions of indicators for Post, Year, Chinese, and Dine Share are included. The full set of interactions of Cases, Cuisine, and Dine Share are also included. Dine share is demeaned in all models except in column 3. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Cuisine to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

Table A.1.5a: Chinese Share of Customers, Take-outs

Share	0.2%	0.5%	1.3%	3.3%	5.7%	15.8%	51.8%	Top 5 Percent Share
Percentile	25%	50%	75%	90%	95%	99%	100%	Top Cities
Chinese	2,537	2,690	2,827	1,891	695	685	434	New York, NY
	(0.216)	(0.229)	(0.240)	(0.161)	(0.059)	(0.058)	(0.037)	Brooklyn, NY
American	37,221	32,877	28,006	14,187	4,135	3,018	473	Chicago, IL
	(0.310)	(0.274)	(0.234)	(0.118)	(0.034)	(0.025)	(0.004)	Los Angeles, CA
Asian	862	1,321	2,251	2,456	1,303	1,218	285	Houston, TX
	(0.089)	(0.136)	(0.232)	(0.253)	(0.134)	(0.126)	(0.029)	Philadelphia, PA
Latin	8,306	9,116	9,489	5,799	1,906	1,531	280	San Francisco, CA
	(0.228)	(0.250)	(0.260)	(0.159)	(0.052)	(0.042)	(0.008)	Bronx, NY
European	1,291	1,724	2,373	2,031	821	579	98	Las Vegas, NV
	(0.145)	(0.193)	(0.266)	(0.228)	(0.092)	(0.065)	(0.011)	Miami, FL

Note: The table provides the distribution of the Chinese share of customers for take-outs (below median dine-share). The top two rows display the Chinese share of customers at different percentiles of the distribution over the entire sample. For each cuisine, the first row provides the number of restaurants in the indicated range and the second row provides the relative representation of the cuisine in the industry, for the quantile of the indicated length. For example, a value greater than 0.01 in the top percentile indicates the cuisine is overrepresented within cuisine among the top percentile of Chinese shares of customers; a value greater than 0.04 indicates the cuisine is overrepresented within cuisine among restaurants in the 95th - 99th percentile of Chinese share of customers, etc. The last column lists, in order, the most represented cities among restaurants within the top 5 percent of Chinese share of customers.

Table A.1.5b: Chinese Share of Customers, Dine-ins

Share	0.2%	0.5%	1.3%	3.3%	5.7%	15.8%	44.9%	Top 5 Percent Share
Percentile	25%	50%	75%	90%	95%	99%	100%	Top Cities
Chinese	2,728	2,499	2,638	1,872	796	951	871	Houston, TX
	(0.221)	(0.202)	(0.214)	(0.152)	(0.064)	(0.077)	(0.070)	New York, NY
American	26,635	26,795	25,665	13,225	3,710	2,557	391	Los Angeles, CA
	(0.269)	(0.271)	(0.259)	(0.134)	(0.037)	(0.026)	(0.004)	San Francisco, CA
Asian	1,859	3,546	5,554	5,504	2,568	2,388	649	Chicago, IL
	(0.084)	(0.161)	(0.252)	(0.249)	(0.116)	(0.108)	(0.029)	San Diego, CA
Latin	9,069	8,616	8,966	5,181	1,543	1,208	162	Brooklyn, NY
	(0.261)	(0.248)	(0.258)	(0.149)	(0.044)	(0.035)	(0.005)	Orlando, FL
European	2,850	4,173	5,589	3,868	1,195	802	92	Seattle, WA
	(0.153)	(0.225)	(0.301)	(0.208)	(0.064)	(0.043)	(0.005)	Austin, TX

Note: The table provides the distribution of the Chinese share of customers for dine-ins (above median dine-share). The top two rows display the Chinese share of customers at different percentiles of the distribution over the entire sample. For each cuisine, the first row provides the number of restaurants in the indicated range and the second row provides the relative representation of the cuisine in the industry, for the quantile of the indicated length. For example, a value greater than 0.01 in the top percentile indicates the cuisine is overrepresented within cuisine among the top percentile of Chinese shares of customers; a value greater than 0.04 indicates the cuisine is overrepresented within cuisine among restaurants in the 95th - 99th percentile of Chinese share of customers, etc. The last column lists, in order, the most represented cities among restaurants within the top 5 percent of Chinese share of customers.

Table A.1.6: Consumer Discrimination, Extended Framework

	Dependent Variable: Log Visits				
	Chinese	American	Asian	Latin	European
Post x 2020 x Chinese Share x Dine Share	-0.027***	0.001	0.000	0.031***	0.039***
	(0.007)	(0.006)	(0.008)	(0.010)	(0.015)
Post x 2020 x Chinese Share	0.031***	0.064***	0.028***	0.077***	0.039***
	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)
Post x 2020 x Dine Share	-0.125***	-0.025***	-0.103***	-0.031***	-0.031***
	(0.013)	(0.004)	(0.012)	(0.006)	(0.012)
Post x 2020	-0.158***	-0.141***	-0.081***	-0.102***	-0.133***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Cases x Chinese Share x Dine Share	0.008	0.000	0.002	0.022	0.031
	(0.013)	(0.010)	(0.013)	(0.014)	(0.022)
Cases x Dine Share	-0.052	0.010	0.024	0.012	0.067*
	(0.042)	(0.010)	(0.028)	(0.015)	(0.036)
Cases x Chinese Share	-0.019***	-0.026***	-0.021***	-0.025***	-0.026***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
Cases	-0.111***	-0.065***	-0.086***	-0.054***	-0.095***
	(0.006)	(0.002)	(0.005)	(0.002)	(0.006)
Total Obs.	674,253	6,122,262	888,024	1,990,393	768,449
R Squared	0.830	0.871	0.818	0.850	0.845
Counties	2,056	2,995	1,379	2,541	1,762
Total Restaurants	24,115	218,906	31,765	71,179	27,493
FE: Restaurant	X	X	X	X	X
Full DDD Controls	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.7, representing estimation of the extended theoretical framework. The Chinese share of customers is calculated using the home census tract of visitors in the entire 2020 pre-treatment period, based on the Chinese share of the population provided by the 2019 ACS 5-year sample. Dine-share is demeaned and Chinese share is standardized. The full set of interactions of indicators for Post, Year, Dine share, and Chinese share are included. The full set of interactions of Cases, Dine share, and Chinese share are also included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese share to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the restaurant level.

Table A.1.7: Baseline Framework, by Customer Type

	Dependent Variable: Log(1 + Visits)					
Reference Group:	Not Demeaned			Demeaned		
All Other Cuisines	All Customers	Chinese	Non-Chinese	All Customers	Chinese	Non-Chinese
Post x 2020 x Chinese x Dine Share	-0.085***	-0.137***	-0.091***	-0.085***	-0.137***	-0.091***
	(0.029)	(0.035)	(0.028)	(0.029)	(0.035)	(0.028)
Post x 2020 x Chinese	0.010	0.040***	0.018	-0.024***	-0.014**	-0.018***
	(0.012)	(0.015)	(0.012)	(0.005)	(0.006)	(0.005)
Post x 2020 x Dine Share	-0.019***	-0.037***	-0.010	-0.019***	-0.037***	-0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)
Post x 2020	-0.126***	-0.066***	-0.129***	-0.134***	-0.081***	-0.133***
	(0.003)	(0.004)	(0.003)	(0.001)	(0.002)	(0.001)
Cases x Chinese x Dine Share	-0.413***	0.290***	-0.453***	-0.413***	0.290***	-0.453***
	(0.069)	(0.094)	(0.069)	(0.069)	(0.094)	(0.069)
Cases x Chinese	0.108***	-0.188***	0.135***	-0.056***	-0.073***	-0.045***
	(0.031)	(0.042)	(0.030)	(0.011)	(0.015)	(0.011)
Cases x Dine Share	0.058***	0.266***	0.031*	0.058***	0.266***	0.031*
	(0.018)	(0.021)	(0.017)	(0.018)	(0.021)	(0.017)
Cases	-0.103***	-0.181***	-0.095***	-0.080***	-0.075***	-0.083***
	(0.008)	(0.009)	(0.007)	(0.003)	(0.004)	(0.003)
Total Obs.	10,444,127	10,089,316	10,089,316	10,444,127	10,089,316	10,089,316
R Squared	0.096	0.410	0.095	0.096	0.410	0.095
FE: County x Cuisine	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework, separately for customer types using county-cuisine fixed effects. The reference group in all models the pooled non-Chinese restaurants. Columns 1-3 provide estimates where dine-share is not demeaned and Columns 4-6 provide estimates where dine-share is demeaned. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

Table A.1.8: Baseline Framework by Customer Type, Restaurant FE

	Dependent Variable: Log(1 + Visits)					
Reference Group:	Not Demeaned			Demeaned		
All Other Cuisines	All Customers	Chinese	Non-Chinese	All Customers	Chinese	Non-Chinese
Post x 2020 x Chinese x Dine Share	-0.096***	-0.108***	-0.107***	-0.096***	-0.108***	-0.107***
	(0.013)	(0.021)	(0.013)	(0.013)	(0.021)	(0.013)
Post x 2020 x Chinese	0.015***	0.028***	0.023***	-0.023***	-0.014***	-0.020***
	(0.006)	(0.009)	(0.005)	(0.002)	(0.004)	(0.002)
Post x 2020 x Dine Share	-0.015***	-0.032***	-0.013***	-0.015***	-0.032***	-0.013***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Post x 2020	-0.130***	-0.068***	-0.128***	-0.136***	-0.081***	-0.133***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Chinese x Dine Share	-0.147***	-0.173***	-0.162***	-0.147***	-0.173***	-0.162***
	(0.036)	(0.054)	(0.034)	(0.036)	(0.054)	(0.034)
Cases x Chinese	-0.003	0.011	0.008	-0.061***	-0.058***	-0.056***
	(0.016)	(0.024)	(0.015)	(0.005)	(0.009)	(0.005)
Cases x Dine Share	0.004	-0.014	0.001	0.004	-0.014	0.001
	(0.007)	(0.011)	(0.007)	(0.007)	(0.011)	(0.007)
Cases	-0.082***	-0.069***	-0.081***	-0.081***	-0.075***	-0.081***
	(0.003)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)
Total Obs.	10,444,127	10,089,316	10,089,316	10,444,127	10,089,316	10,089,316
R Squared	0.862	0.815	0.867	0.862	0.815	0.867
FE: Restaurants	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X

* p < 0.1, ** p < 0.05, *** p < 0.01

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework, separately for customer types using restaurant fixed effects. The reference group in all models the pooled non-Chinese restaurants. Columns 1-3 provide estimates where dine-share is not demeaned and Columns 4-6 provide estimates where dine-share is demeaned. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the restaurant level.

Table A.1.9: Baseline Framework, Unrestricted Sample

	Dependent Variable: Log Visits							
	Chinese Restaurants (Ref = All Other Cuisines)				Cuisine (Ref = American Restaurants)			
	No Controls	Controls	Not Demeaned	Preferred	Chinese	Asian	Latin	European
Post x 2020 x Cuisine x Dine Share	-0.104***	-0.106***	-0.080**	-0.080**	-0.081**	-0.090***	-0.010	-0.071**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)	(0.031)	(0.021)	(0.031)
Post x 2020 x Cuisine	-0.016***	-0.023***	0.009	-0.022***	0.000	0.110***	0.049***	0.038***
	(0.005)	(0.006)	(0.015)	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)
Post x 2020 x Dine Share	-0.052***	-0.061***	-0.059***	-0.059***	-0.058***	-0.058***	-0.058***	-0.058***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)	(0.010)
Post x 2020	-0.100***	-0.101***	-0.078***	-0.101***	-0.124***	-0.124***	-0.124***	-0.124***
	(0.001)	(0.001)	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine x Dine Share			-0.497***	-0.497***	-0.495***	-0.284***	0.143***	-0.160**
			(0.078)	(0.078)	(0.080)	(0.056)	(0.050)	(0.081)
Cases x Cuisine	-0.041***	-0.053***	0.150***	-0.045***	-0.051***	-0.022**	0.015**	-0.026**
	(0.012)	(0.012)	(0.035)	(0.012)	(0.013)	(0.009)	(0.007)	(0.012)
Cases x Dine Share			-0.034*	-0.034*	-0.037	-0.037	-0.037	-0.037
			(0.018)	(0.018)	(0.026)	(0.026)	(0.026)	(0.026)
Cases	-0.108***	-0.088***	-0.074***	-0.088***	-0.082***	-0.082***	-0.082***	-0.082***
	(0.003)	(0.003)	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Total Obs.	15,119,246	15,119,246	15,119,246	15,119,246	9,263,782	9,578,324	10,884,629	9,490,839
R Squared	0.134	0.134	0.134	0.134	0.118	0.119	0.105	0.120
Counties	3,074	3,074	3,074	3,074	3,060	3,060	3,063	3,062
Total Restaurants	556,723	556,723	556,723	556,723	340,934	352,579	400,076	349,114
Cuisine Restaurants	29,543	29,543	29,543	29,543	29,543	41,188	88,685	37,723
FE: County x Cuisine	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year		X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework, using the unrestricted sample. The reference group in columns 1-4 are the pooled non-Chinese cuisines. The reference group in columns 5-8 is American restaurants. The full set of interactions of indicators for Post, Year, Chinese, and Dine Share are included. The full set of interactions of Cases, Cuisine, and Dine Share are also included. Dine share is demeaned in all models except in column 3. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Cuisine to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

Table A.1.10: Response to Chinese Share of Customers, Unrestricted Sample

	Dependent Variable: Log(1 + Visits), by Customer Type							
	Chinese Restaurants				Non-Chinese Restaurants			
Customer Type:	All	Chinese	Non-Chinese	Difference	All	Chinese	Non-Chinese	Difference
Post x 2020 x Chinese Share	0.040***	-0.022***	0.042***	-0.064***	0.066***	-0.018***	0.068***	-0.086***
	(0.004)	(0.006)	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)
Post x 2020	-0.127***	-0.076***	-0.132***	0.057***	-0.098***	-0.067***	-0.104***	0.037***
	(0.003)	(0.004)	(0.002)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Chinese Share	-0.017***	-0.018***	-0.018***	0.000	-0.026***	-0.027***	-0.027***	0.000
	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Cases	-0.117***	-0.102***	-0.109***	0.006	-0.072***	-0.065***	-0.070***	0.005***
	(0.006)	(0.009)	(0.006)	(0.008)	(0.001)	(0.002)	(0.001)	(0.002)
Total Obs.	787,852	712,800	712,800	712,800	14,179,039	13,060,087	13,060,087	13,060,087
R Squared	0.857	0.840	0.850	0.814	0.875	0.816	0.873	0.766
FE: Restaurant	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.5, representing the response of Chinese and non-Chinese customers to the pre-covid restaurant Chinese share, using the unrestricted sample. The dependent variables in columns 1-3 and 5-7 are measured as log(1 plus customer visits). In columns 4 and 8, the dependent variable is the difference, log(1 plus Chinese visits) - log(1 plus non-Chinese visits). Visit counts for customer types are calculated by allocating total visits according to the 2019 ACS 5-year sample of the Chinese and total population and the home census tract of visitors. The restaurant Chinese share of customers is calculated similarly, using the entire 2020 pre-treatment period. Chinese share is standardized. The full set of interactions of Post, Year, and Chinese share are included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese share to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the restaurant level.

Table A.1.11: Baseline Framework Using Raw Visits

	Dependent Variable: Log Visits							
	Chinese Restaurants (Ref = All Other Cuisines)				Cuisine (Ref = American Restaurants)			
	No Controls	Controls	Not Demeaned	Preferred	Chinese	Asian	Latin	European
Post x 2020 x Cuisine x Dine Share	-0.130***	-0.130***	-0.106***	-0.106***	-0.102***	-0.082***	-0.029	-0.040
	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)	(0.026)	(0.018)	(0.029)
Post x 2020 x Cuisine	-0.023***	-0.018***	0.024*	-0.018***	-0.001	0.083***	0.035***	0.030***
	(0.004)	(0.005)	(0.012)	(0.005)	(0.005)	(0.005)	(0.003)	(0.005)
Post x 2020 x Dine Share	-0.004	-0.036***	-0.036***	-0.036***	-0.041***	-0.041***	-0.041***	-0.041***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Post x 2020	-0.034***	-0.013***	0.001	-0.013***	-0.030***	-0.030***	-0.030***	-0.030***
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine x Dine Share			-0.464***	-0.464***	-0.515***	-0.386***	-0.007	-0.372***
			(0.069)	(0.069)	(0.071)	(0.050)	(0.042)	(0.076)
Cases x Cuisine	-0.038***	-0.041***	0.151***	-0.032***	-0.038***	-0.016*	0.004	-0.005
	(0.011)	(0.011)	(0.031)	(0.011)	(0.012)	(0.009)	(0.007)	(0.012)
Cases x Dine Share			0.013	0.013	0.064***	0.064***	0.064***	0.064***
			(0.017)	(0.017)	(0.024)	(0.024)	(0.024)	(0.024)
Cases	-0.086***	-0.016***	-0.021***	-0.016***	-0.010**	-0.010**	-0.010**	-0.010**
	(0.003)	(0.003)	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Total Obs.	10,444,127	10,444,127	10,444,127	10,444,127	6,796,847	7,010,617	8,113,155	6,891,206
R Squared	0.128	0.129	0.129	0.129	0.118	0.119	0.105	0.119
Counties	3,008	3,008	3,008	3,008	2,996	2,995	3,005	2,997
Total Restaurants	373,458	373,458	373,458	373,458	243,021	250,671	290,085	246,399
Cuisine Restaurants	24,115	24,115	24,115	24,115	24,115	31,765	71,179	27,493
FE: County x Cuisine	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year		X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework, using raw GPS visits detected. The reference group in columns 1-4 are the pooled non-Chinese cuisines. The reference group in columns 5-8 is American restaurants. The full set of interactions of indicators for Post, Year, Chinese, and Dine Share are included. The full set of interactions of Cases, Cuisine, and Dine Share are also included. Dine share is demeaned in all models except in column 3. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Cuisine to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

Table A.1.12: Response to Chinese Share of Customers with Alternate Outcome

	Dependent Variable: Percent Decline, by Customer Type							
	Chinese Restaurants				Non-Chinese Restaurants			
Customer Type:	All	Chinese	Non-Chinese	Difference	All	Chinese	Non-Chinese	Difference
Post x 2020 x Chinese Share	0.050***	-0.134***	0.065***	-0.199***	0.062***	-0.299***	0.071***	-0.370***
	(0.005)	(0.048)	(0.005)	(0.048)	(0.002)	(0.017)	(0.002)	(0.017)
Post x 2020	-0.076***	0.283***	-0.074***	0.358***	-0.060***	0.230***	-0.061***	0.291***
	(0.004)	(0.064)	(0.004)	(0.064)	(0.001)	(0.017)	(0.001)	(0.017)
Cases x Chinese Share	-0.018***	-0.016	-0.019***	0.003	-0.024***	-0.009	-0.028***	0.019***
	(0.003)	(0.032)	(0.003)	(0.031)	(0.001)	(0.007)	(0.001)	(0.007)
Cases	-0.092***	-0.122	-0.097***	-0.025	-0.062***	-0.105***	-0.063***	-0.042**
	(0.011)	(0.162)	(0.012)	(0.161)	(0.002)	(0.020)	(0.002)	(0.019)
Total Obs.	674,253	645,074	646,079	645,074	9,769,128	9,423,003	9,442,887	9,423,003
R Squared	0.305	0.175	0.401	0.170	0.290	0.115	0.436	0.110
FE: Restaurant	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 1.5, representing the response of Chinese and non-Chinese customers to the pre-covid restaurant Chinese share. The dependent variables is the percent decline in visits relative to the restaurant's average weekly visits in the 2020 pre-treatment period. Visit counts for customer types are calculated by allocating total visits according to the 2019 ACS 5-year sample of the Chinese and total population and the home census tract of visitors. The restaurant Chinese share of customers is calculated similarly, using the entire 2020 pre-treatment period. Chinese share is standardized. The full set of interactions of Post, Year, and Chinese share are included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Chinese share to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the restaurant level.

Table A.1.13: Baseline Framework With Alternate Controls

	Dependent Variable: Log Visits								
Reference Group:	Chinese Restaurants			Latin Restaurants			European Restaurants		
American Restaurants	Baseline	Current Weeks	Log Pop.	Baseline	Current Weeks	Log Pop.	Baseline	Current Weeks	Log Pop.
Post x 2020 x Cuisine x Dine Share	-0.079***	-0.058*	-0.080***	-0.026	-0.023	-0.030*	-0.018	-0.002	-0.019
	(0.029)	(0.030)	(0.029)	(0.018)	(0.018)	(0.018)	(0.029)	(0.029)	(0.029)
Post x 2020 x Cuisine	-0.008	-0.007	-0.006	0.038***	0.039***	0.029***	0.022***	0.023***	0.034***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
Post x 2020 x Dine Share	-0.025***	-0.027***	-0.025***	-0.025***	-0.027***	-0.025***	-0.025***	-0.027***	-0.025***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Post x 2020	-0.150***	-0.150***	-0.159***	-0.150***	-0.150***	-0.159***	-0.150***	-0.150***	-0.159***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine x Dine Share	-0.464***	-0.305***	-0.465***	-0.008	-0.038*	-0.002	-0.288***	-0.256***	-0.290***
	(0.071)	(0.032)	(0.071)	(0.043)	(0.022)	(0.043)	(0.076)	(0.035)	(0.076)
Cases x Cuisine	-0.061***	-0.028***	-0.058***	0.014**	0.001	0.013*	-0.032***	-0.007	-0.026**
	(0.012)	(0.005)	(0.012)	(0.007)	(0.004)	(0.007)	(0.012)	(0.005)	(0.012)
Cases x Dine Share	0.108***	0.063***	0.110***	0.108***	0.063***	0.110***	0.108***	0.063***	0.110***
	(0.024)	(0.012)	(0.024)	(0.024)	(0.012)	(0.024)	(0.024)	(0.012)	(0.024)
Cases	-0.075***	-0.075***	-0.076***	-0.075***	-0.075***	-0.076***	-0.075***	-0.075***	-0.076***
	(0.004)	(0.002)	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)	(0.002)	(0.004)
Total Obs.	6,796,847	6,796,847	6,796,847	8,113,155	8,113,155	8,113,155	6,891,206	6,891,206	6,891,206
R Squared	0.086	0.087	0.086	0.073	0.073	0.073	0.086	0.086	0.086
Counties	2,996	2,996	2,996	3,005	3,005	3,005	2,997	2,997	2,997
Total Restaurants	243,021	243,021	243,021	290,085	290,085	290,085	246,399	246,399	246,399
Cuisine Restaurants	24,115	24,115	24,115	71,179	71,179	71,179	27,493	27,493	27,493
FE: County x Cuisine	X	X	X	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X	X
Controls[3] x Post x Year	X	X	X	X	X	X	X	X	X
Cases: Prior 2 Weeks	X		X	X		X	X		X
Cases: Current 2 Weeks		X			X			X	
Population	X	X		X	X		X	X	
Log Population			X			X			X

* p < 0.1, ** p < 0.05, *** p < 0.01

Note: The table provides estimates obtained from equation 1.4, representing estimation of the baseline theoretical framework, using raw GPS visits detected. The reference group in all columns are American restaurants. Models labeled baseline replicate the cuisine-specific results in table A.1.4. Models labeled Cases use the logged 1 plus combined covid cases in the current and previous week. Models labeled Pop. replace the population control for the logged population. The full set of interactions of indicators for Post, Year, Chinese, and Dine Share are included. The full set of interactions of Cases, Cuisine, and Dine Share are also included. Dine share is demeaned in all models. Controls are included as interactions with Post x Year to capture the differential impact of the control in response to the announcement. Controls are also included as interactions with cuisine. Standard errors are clustered at the county x cuisine level.

Table A.1.14: Baseline Framework, Controls Variables

	Dependent Variable: Log Visits					
	Chinese Restaurants (Ref = All Other Cuisines)		Cuisine (Ref = American Restaurants)			
	Not Demeaned	Preferred	Chinese	Asian	Latin	European
Cases x Dine Share x Cuisine	-0.414***	-0.414***	-0.464***	-0.349***	-0.008	-0.288***
	(0.069)	(0.069)	(0.071)	(0.051)	(0.043)	(0.076)
Cases x Cuisine	0.108***	-0.056***	-0.061***	-0.027***	0.014**	-0.032***
	(0.031)	(0.011)	(0.012)	(0.009)	(0.007)	(0.012)
Cases x Dine Share	0.058***	0.058***	0.108***	0.108***	0.108***	0.108***
	(0.018)	(0.018)	(0.024)	(0.024)	(0.024)	(0.024)
Cases	-0.103***	-0.080***	-0.075***	-0.075***	-0.075***	-0.075***
	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Post x 2020 x Cuisine x	-0.001***	-0.001***	-0.001***	0.000	0.000**	0.000*
2016 Dem. Share (pp.)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Post x 2020 x	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
2016 Dem. Share (pp.)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Post x 2020 x Cuisine x	0.001	0.001	0.000	0.000	-0.003***	0.000
Chinese Pop. Share (pp.)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post x 2020 x	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
Chinese Pop. Share (pp.)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Post x 2020 x Cuisine x	0.001	0.001	0.001	0.001	0.001	-0.001
Density (per 10,000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post x 2020 x	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
Density (per 10,000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Post x 2020 x Cuisine x	0.003**	0.003**	0.002	-0.001	-0.004***	0.007***
Total Pop. (Millions)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Post x 2020 x	-0.005***	-0.005***	-0.004***	-0.004***	-0.004***	-0.004***
Total Pop. (Millions)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Total Obs.	10,444,127	10,444,127	6,796,847	7,010,617	8,113,155	6,891,206
R Squared	0.096	0.096	0.086	0.085	0.073	0.086
FE: County x Cuisine	X	X	X	X	X	X
Full DDD Controls	X	X	X	X	X	X

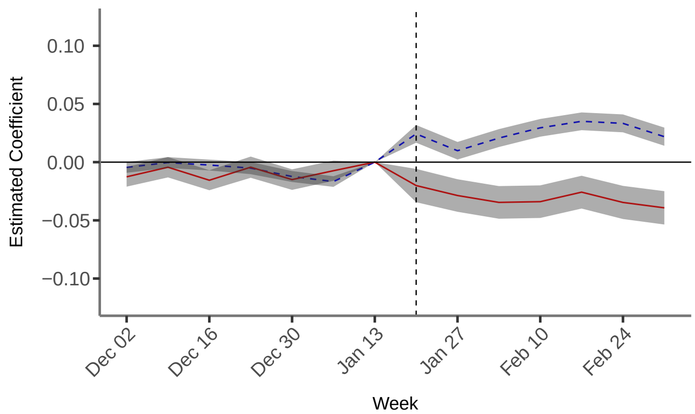
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates coefficients of control variables obtained from equation 1.4, representing estimation of the baseline theoretical framework. The reference group in columns 1-2 are the pooled non-Chinese cuisines. The reference group in columns 3-6 is American restaurants. The full set of interactions of indicators for Post, Year, Chinese, and Dine Share are included. The full set of interactions of Cases, Cuisine, and Dine Share are also included. Dine share is demeaned in all models except in column 1. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year and with Post x Year x Cuisine to capture the differential impact of the control in response to the announcement. Standard errors are clustered at the county x cuisine level.

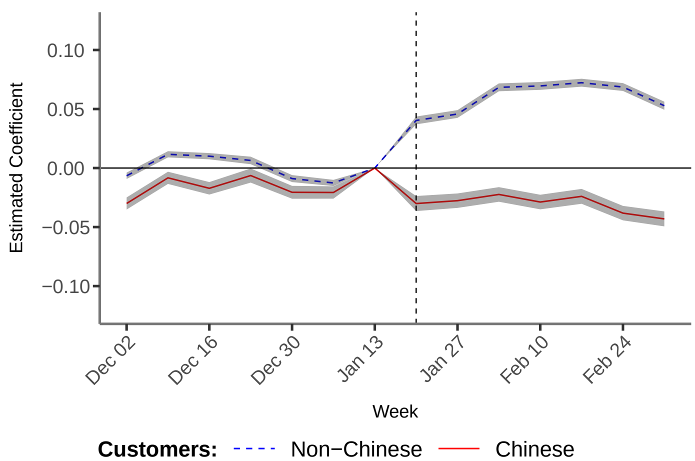
A.2 Figures

Figure A.2.1: Consumer Response to High Chinese-Share Restaurants

(a) Chinese Restaurants



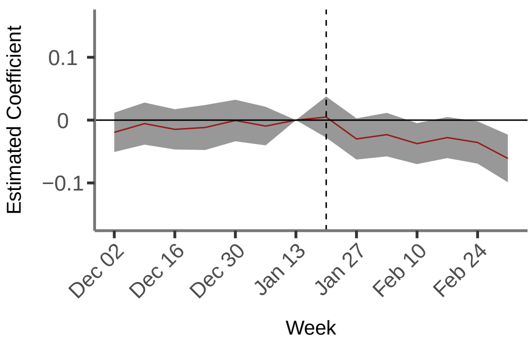
(b) Non-Chinese Restaurants



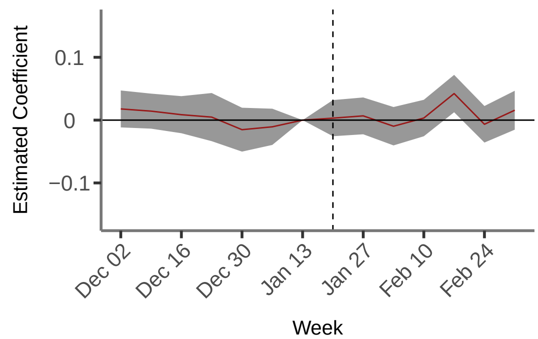
Note: The figure provides event studies representing the response of Chinese and non-Chinese customers to the pre-covid restaurants with a one standard deviation greater Chinese share of customers, and correspond to $\tilde{\delta}_2$ of equation 1.5. Event studies are shown separately for Chinese restaurants (a) and non-Chinese restaurants (b).

Figure A.2.2: Consumer Response to Sharing Space with Chinese Customers

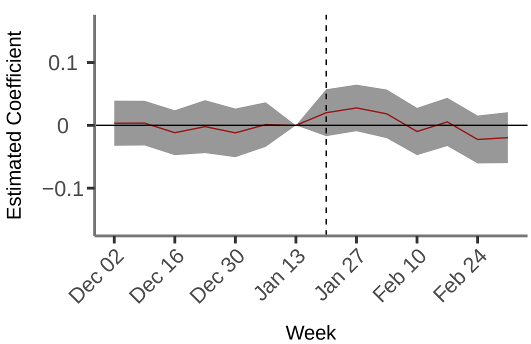
(a) Chinese Restaurants



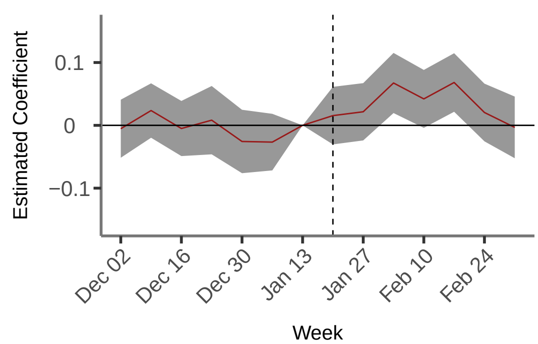
(b) American Restaurants



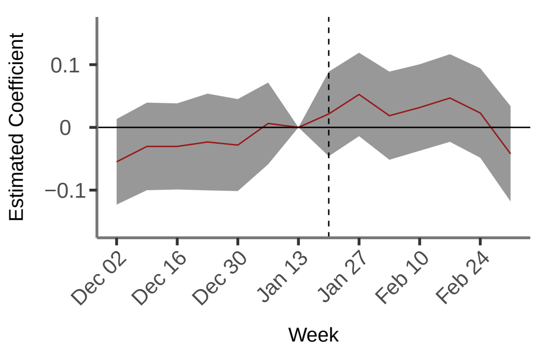
(c) Asian Restaurants



(d) Latin Restaurants



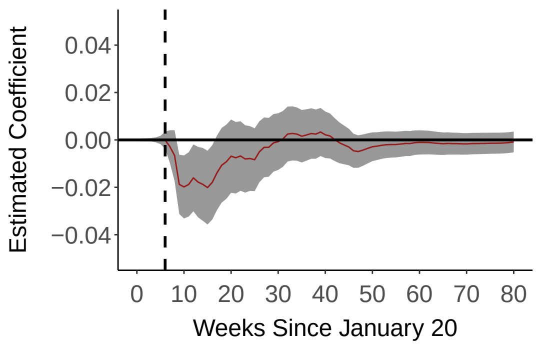
(e) European Restaurants



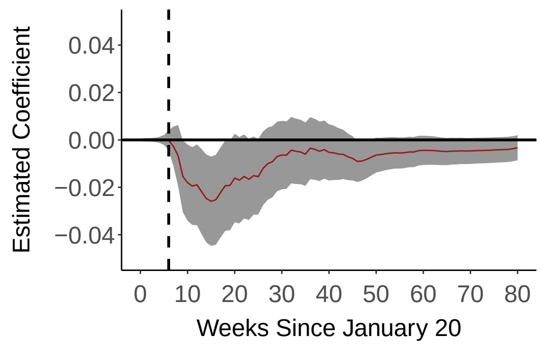
Note: The figures display event studies of the consumer response to dine-in restaurants with a one standard deviation greater Chinese share of customers, and correspond to δ_4 in equation 1.7, separately for each cuisine group.

Figure A.2.3: Chinese Restaurants Predicting Covid Cases

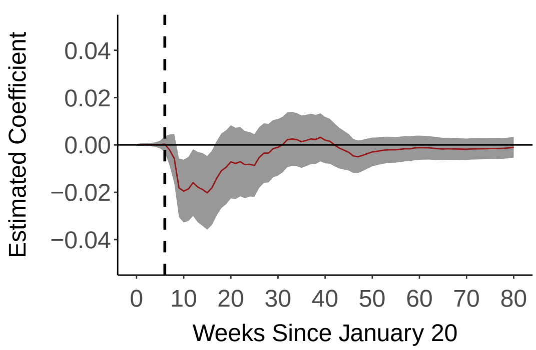
(a) Restaurant Share



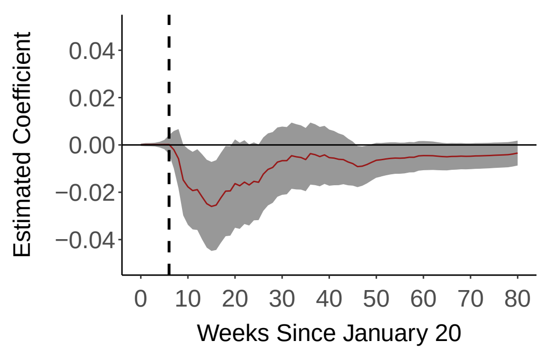
(b) Visit Share



(c) Restaurant Share, Control Chinese Pop. Share



(d) Visit Share, Control Chinese Pop. Share



Note: The figures display coefficient estimates of county-level Chinese share of restaurants in (a) and (c), and the share of all customers that visited a Chinese restaurant in (b) and (d), as predictors for future covid cases. The unrestricted sample of restaurants is included in each estimated model. Each figure correspond to β in equation 1.8, which is restated below. All models include the log population, density, and the 2016 democratic share as controls. Models in subfigures (c) and (d) also include the Chinese ancestry share of the county's population as a control. The primary explanatory variable, Share, is obtained using 2020 pre-announcement consumer outcomes. The outcome variable is cumulative covid cases for each of 80 weeks following the covid announcement. Coefficient estimates for the first seven weeks are very small and alone are not meaningful. The vertical dashed line indicates the end of the observational period in this study, March 8.

Estimated model (equation 1.8 in text):

$$log(1 + CumulCases)_c = \alpha + \beta Share_c + \mathbf{X}_c \gamma + \varepsilon$$

Chapter 2 Appendix

B.1 Data

B.1.1 Control Variables

Data on reported covid cases is provided by the New York Times and includes the cumulative reported covid cases and deaths for each county-day. I take one plus the number of new cases reported in the prior two weeks in logs as the primary control variable that is included in all baseline and preferred specifications.[1]

I obtain the county share of 2016 Democratic party votes from the Harvard Dataverse. I obtain data on 2019 county populations and population density from the US Census. I use to the 2015-2019 American Community Survey to construct the county level Chinese ancestry share of the population, which is limited to the population aged 16 and over.

B.1.2 Twitter

The python library "snscrape" is free software under the General Public License, as published by the Free Software Foundation, and is a scraper for social networking services, including Twitter, Facebook, Instagram, Reddit, and others. When scraping tweets, it has the advantage that it is not constrained by Twitter's API scraping limits, such as a limited number of tweets per 15 minutes, as other scrapers are. Instead, snscrape

[1] As this data set does not distinguish the five boroughs of New York, I group observations in New York County, Kings County, Queens County, Bronx County, and Richmond County into one county representing New York in all datasets.

circumvents Twitter's limits at the cost of several querying features and functionalities that are ultimately unnecessary for this analysis.[2]

There are two noteworthy limitations when scraping historical geo-tagged tweets with snscrape. First, any tweets that are removed either by the user or by Twitter cannot be scrapped. Second, the geographic coordinates of tweets are only available for users who have affirmatively permitted this in their Twitter profile settings. This ultimately represents approximately 1% of all tweets. However, the sheer number of tweets made each day still provide a sizable enough sample to be used in the analysis. While this may typically raise a selection concern, if individuals who turn on the geo-tag setting are different from the population of interest, this is less of a concern in panel studies that examine changes in behavior over time. The self-selection into turning on the geo-tag feature only matters if there is reason to believe that these Twitter users would respond differently to treatment than others, which is not obvious in this observational setting. The Twitter dataset used in this study is collected from the universe of available geo-tagged tweets that satisfy these two restrictions. With these caveats in mind, I describe the data cleaning process next.

One querying limitation of snscrape is that the user cannot filter the scraping query for a particular country or a list of sub-national entities, but can instead specify a geographic point and radius. All available geo-tagged tweets that fall within that geographic circle are then retrieved. I use the geographic center of the contiguous 48 US states, located in

[2] Twitter has recently made available an Academic Access application for researchers to access to historical tweets with higher scraping limits, however, for the purposes of this study, that access is unnecessary. Nevertheless, I adhere to all Twitter policies regarding data scraping and provision of data. Specifically, only aggregate data is presented in this paper, and any sharing of tweet information is strictly limited to the tweet id, which others may use to hydrate tweet attributes. Hydration is the process by which one can acquire all variables associated with a tweet using just the tweet id, and can be done done via the Twitter API.

the small town of Lebanon, Kansas, as my point, and a radius of 3,000 kilometers. This circle fully encompasses the contiguous 48 US states plus significant areas of Canada and Mexico. Once tweets are retrieved, I drop those made in Canada and Mexico and then use the latitude and longitude to assign to counties using census shapefiles obtained from the tidycensus package in R. Some tweets do not have an exact match to a county, due to assignment to locations that practically but not legally fall under the jurisdiction of a county such as on lakes/waters, due to measurement error in the recording of coordinates at the time of the tweet, or due to precision loss in the digits of the coordinates. For these tweets, I obtain the closest county centroid if it falls within 50 miles. These are then verified for accuracy using the associated place name variable of the tweet.

The place name is a user-tagged named location chosen from one of Twitter's suggested options that generally corresponds to their location, but is not necessarily accurate because the user can forgo this option and enter any other place altogether. For example, a user can choose to tag "California, USA", or "Los Angeles, CA", or "Venice Beach, Los Angeles", or a place in another state entirely. In most cases the named place correspond to the nearest county, and I consider these sufficiently close to the true location of the user; these tweets are included in the sample. For unmatched tweets where the place name is another nearby county (for example, for an unmatched tweet where the place name is Los Angeles, CA but the nearest county centroid is San Bernardino County, CA), I assign these based on the nearest centroid rather than the county associated with the place name to avoid over-assignment of large metropolitan counties. I include these in the sample as well and drop all remaining tweets. Of all scraped tweets, approximately 81% match exactly using coordinates, 14% are matched to the nearest county centroid and supported by the place name, and the rest are dropped.

Scraping was carried out three times, on September 1, 2022, December 18, 2022, and February 24, 2023. The three batches of scrapes have over 90% overlap of tweets, with the rest arising due to the de-activation or re-activation of Twitter accounts. I pool the three sets of data files and keep unique tweets. Estimated results using any one of the three individually scraped datasets provide nearly identical results.

Snscrape allows for querying tweets within a specific time window, with specificity down to the second. However, it does not provide a random sample of tweets within that search period. Instead, it begins scraping tweets starting from the end to the beginning of the specified time period. For example, a query for tweets on January 20 will start collecting tweets made at 11:59:59 pm UTC and work back, scraping each available tweet within the geographic radius until 12:00:00 am UTC. This is not a concern for the covid dataset because I collect the universe of available geo-tagged tweets - a much smaller dataset of approximately 108,000 - but it does create a data collection issue for the dataset used to construct the instrument and using in placebo specifications, which do not query for any keywords. Obtaining the universe of available geotagged tweets without keyword restrictions is both computationally demanding and unnecessary for the analysis. Limiting the query to particular days is not a desirable solution because some counties may be more likely to tweets on certain days of the week, which would result in their Twitter activity being oversampled.

To overcome this, I randomly sample time segments within the observational setting and scrape all tweets made within those time frames. I implement the following sampling procedure: for each day in the observational period, I randomly draw 60 disjoint two-minute intervals without replacement, for a total of 120 minutes. As an example, one of these intervals could be 02:06:00 am - 02:07:59 am, when people are less likely to be tweeting, while another could be 10:16:00 pm - 10:17:59 pm, when people are more likely to be tweeting. Each interval has an ex ante equal likelihood of being drawn, and once drawn, I collect all tweets made within that interval. This approximates a random sampling of 120 out of 1,440 minutes each day, or approximately an 8.33% sample. This procedure overcomes snscrape's limitation because even if some counties are more likely to tweet on different days or at different times each day, this procedure obtains the random sample by selecting a random 8.33% of time each day.

I scrape tweets without specifying any keywords for three observational years: December 4, 2017 - March 11, 2018, December 3, 2018 - March 10, 2019, and December 2, 2019 - March 8, 2020, which align the 14 week period according to weekdays each year. This dataset represents approximately 16.1 million tweets. Appendix figure B.2.1 provides an illustrative timeline of how each sample of tweets is used in the analysis.

B.2 Tables

Table B.1.1: Cumulative Covid Cases (Deaths) Worldwide

Date	US	China	Asia	Latin America	Europe	All Others
Jan 20	1 (0)	202 (4)	8 (0)	0 (0)	7 (0)	0 (0)
Jan 31	8 (0)	9,720 (213)	86 (0)	0 (0)	24 (0)	20 (0)
Feb 15	23 (0)	66,576 (1,663)	225 (2)	0 (0)	86 (4)	328 (0)
Feb 29	69 (1)	79,389 (2,838)	3,363 (22)	6 (0)	1,269 (28)	1,306 (49)
Mar 8	464 (19)	80,859 (3,101)	7,967 (58)	130 (2)	10,448 (278)	7,880 (215)
Mar 15	3,929 (69)	81,048 (3,204)	9,977 (115)	742 (15)	53,885 (1,978)	16,717 (760)
Mar 31	173,143 (3,327)	82,545 (3,314)	21,893 (533)	18,402 (572)	450,099 (32,538)	81,561 (3,587)

Counts obtained from WHO (2023). Asia includes eastern and southern Asian countries except China. Latin America includes central and South America and Caribbean countries. Europe includes western and eastern European countries and Russia, and excludes Turkey and Middle Eastern countries. Other includes all other excluded countries and regions (Africa, Turkey, Middle East, the Caucasus, Pacific Islands, Australia, New Zealand, and Canada)

Table B.1.2: Timeline of Key Events

Jan 30	WHO declares international emergency	**Mar 11**	WHO declares COVID-19 a pandemic
Jan 31	US places travel restrictions on China	**Mar 12**	First covid restriction in US; NY closes public schools
Feb 3	US declares a public health emergency	**Mar 13**	Trump declares national emergency, announces travel restrictions on 26 countries
Feb 11	WHO announces official name "COVID-19"	**Mar 15**	First statewide bans on public events (NY, OH)
Mar 6	US Secretary of State uses "Wuhan virus" on Twitter	**Mar 16**	Trump's first "Chinese virus" tweet

Timeline events obtained from Cao, Lindo and Zhong (2022), Google (2022), and National Public Radio (2020)

Table B.1.3: Summary Statistics of Tweets

	2019		2020		Total
	Pre	Post	Pre	Post	
Covid Tweets	7	4	73	108,442	108,526
Mean Tweets/Pop	—	—	—	0.027	
Median Tweets/Pop	—	—	—	0.000	
SD Tweets/Pop	—	—	—	0.238	
Placebo Tweets	2,661,966	2,909,378	2,602,430	2,503,719	10,677,493
Mean Tweets/Pop	10.6	11.5	10.2	9.7	
Median Tweets/Pop	2.0	2.2	2.0	1.9	
SD Tweets/Pop	98.8	104.4	84.4	78.9	
	2018		2019		Total
	Pre	Post	Pre	Post	
Unrelated Tweets	2,819,113	3,064,479	2,661,966	2,909,378	11,454,936
Mean Tweets/Pop	11.6	12.4	10.6	11.5	
Median Tweets/Pop	2.3	2.4	2.0	2.2	
SD Tweets/Pop	107.5	114.8	98.8	104.4	

The table provides counts and county-level statistics over the specified period. Covid tweets are those that include the strings “covid”, “coronavirus”, or “corona virus”. Placebo and Unrelated tweets are collected from three years and do not query for any keywords. The term coronavirus is the name of a group of viruses that includes the common cold. References to covid prior to January 2020 were related to any one of these but not COVID-19. Means, medians, and standard deviations are calculated at the county level over the aggregate seven week period, and are scaled to account for the 1% of tweets that are geo-tagged (covid, placebo, and unrelated tweets) and the sampling rate (8.33%, placebo and unrelated tweets).

Table B.1.4: Consumer Response to Covid Announcement

	Dependent Variable: Log Visits						
	Chinese Restaurants (Ref = All Others)			Cuisine (Ref = American Restaurants)			
	No Controls	Controls	Baseline	Chinese	Asian	Latin	European
Post x 2020 x Cuisine	-0.024***	-0.021***	-0.026***	-0.009*	0.070***	0.039***	0.019***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Post x 2020	-0.137***	-0.132***	-0.131***	-0.147***	-0.147***	-0.147***	-0.147***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Cases x Cuisine		-0.052***	-0.062***	-0.065***	-0.035***	0.015**	-0.034***
		(0.011)	(0.011)	(0.012)	(0.008)	(0.007)	(0.012)
Cases		-0.105***	-0.081***	-0.078***	-0.078***	-0.078***	-0.078***
		(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Total Obs.	10,006,342	10,006,342	10,006,342	6,456,481	6,682,334	7,718,167	6,555,012
R Squared	0.074	0.074	0.074	0.065	0.063	0.053	0.058
Counties	1,591	1,591	1,591	1,591	1,591	1,591	1,591
Total Restaurants	357,798	357,798	357,798	230,846	238,928	275,956	234,373
Cuisine Restaurants	23,411	23,411	23,411	23,411	31,493	68,521	26,938
FE: County x Cuisine	X	X	X	X	X	X	X
Full DDD Controls		X	X	X	X	X	X
Controls[4] x Post x Year			X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 2.1 representing estimation of the partitioned DDD model. The reference group in all models is non-Chinese restaurants. The full set of interactions of indicators for Post, Year, and Chinese are included. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. All controls are also included as interactions with cuisine. Standard errors clustered at the county x cuisine level.

Table B.1.5: Consumer Response to Twitter Content, IV with Alt. Fixed Effects

	Dependent Variable: Log Visits					
	Covid			Placebo		
Term T. I. x Post x 2020 x Cuisine	-0.011**	-0.023***	-0.011**	0.029	-0.011	0.002
	(0.005)	(0.005)	(0.005)	(0.029)	(0.029)	(0.029)
Term T. I. x Post x 2020	-0.033***	-0.033***	-0.034***	0.045***	0.062***	0.061***
	(0.002)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)
Post x 2020 x Cuisine	-0.024***	-0.097***	-0.031***	-0.026***	-0.033***	-0.033***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Post x 2020	-0.126***	-0.130***	-0.134***	-0.127***	-0.131***	-0.131***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Cases x Cuisine	-0.041***	-0.022*	-0.039***	-0.062***	-0.046***	-0.060***
	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Cases	-0.063***	-0.073***	-0.072***	-0.082***	-0.092***	-0.091***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Total Obs.	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342	10,006,342
R Squared	0.075	0.047	0.076	0.074	0.047	0.076
Tweet Count	108,526	108,526	108,526	11,454,880	11,454,880	11,454,880
FE: County x Cuisine	X		X	X		X
FE: County x Year		X	X		X	X
Full DDD Controls	X	X	X	X	X	X
Control Covid Cases	X	X	X	X	X	X
Controls[4] x Post x Year	X	X	X	X	X	X
Controls[4] x Nat. Tweets	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 2.4, representing the main IV specification for covid and placebo tweets. The reference group in all models is non-Chinese restaurants. The full set of interactions of indicators for Placebo, Post, Year, and cuisine are included in placebo models. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. Control variables are also included as interactions with the national time series of tweets to capture the differential impact of the control in the term's salience. All controls are also included as interactions with cuisine. Standard errors clustered at the county x cuisine level for specifications including county x cuisine fixed effects, and clustered at the county x year level for specifications including county x year fixed effects.

Table B.1.6: Consumer Visit Response to Covid Tweets, IV with Alt. Fixed Effects

	Dependent Variable: Log Visits							
	Chinese	Asian	Latin	European	Chinese	Asian	Latin	European
Term T. I. x Post x 2020 x Cuisine	-0.015***	-0.022***	-0.001	-0.012***	-0.015***	-0.020***	-0.002	-0.011***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Term T. I. x Post x 2020	-0.030***	-0.030***	-0.030***	-0.030***	-0.032***	-0.032***	-0.031***	-0.031***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post x 2020 x Cuisine	-0.006***	0.073***	0.039***	0.022***	-0.019***	0.053***	0.027***	0.013***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Post x 2020	-0.145***	-0.145***	-0.145***	-0.145***	-0.147***	-0.147***	-0.149***	-0.148***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Cuisine	-0.048***	-0.033***	0.007**	-0.032***	-0.044***	-0.027***	0.009***	-0.031***
	(0.006)	(0.004)	(0.003)	(0.006)	(0.006)	(0.004)	(0.003)	(0.006)
Cases	-0.059***	-0.059***	-0.059***	-0.059***	-0.069***	-0.068***	-0.069***	-0.069***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Total Obs.	6,456,481	6,682,334	7,718,167	6,555,012	6,456,481	6,682,334	7,718,167	6,555,012
R Squared	0.871	0.869	0.868	0.871	0.872	0.870	0.869	0.872
Tweet Count	108,526	108,526	108,526	108,526	108,526	108,526	108,526	108,526
FE: Restaurant	X	X	X	X	X	X	X	X
FE: County x Year					X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Control Covid Cases	X	X	X	X	X	X	X	X
Controls[4] x Nat. Tweets	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 2.4, representing the main IV specification for covid tweets. The reference group in all models is American restaurants. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. Control variables are also included as interactions with the national time series of tweets to capture the differential impact of the control in the term's salience. All controls are also included as interactions with cuisine. Standard errors clustered at the restaurant level for specifications including restaurant fixed effects, and clustered at the county x year level for specifications including county x year fixed effects.

Table B.1.7: Consumer Visit Response to Placebo Tweets, IV with Alt. Fixed Effects

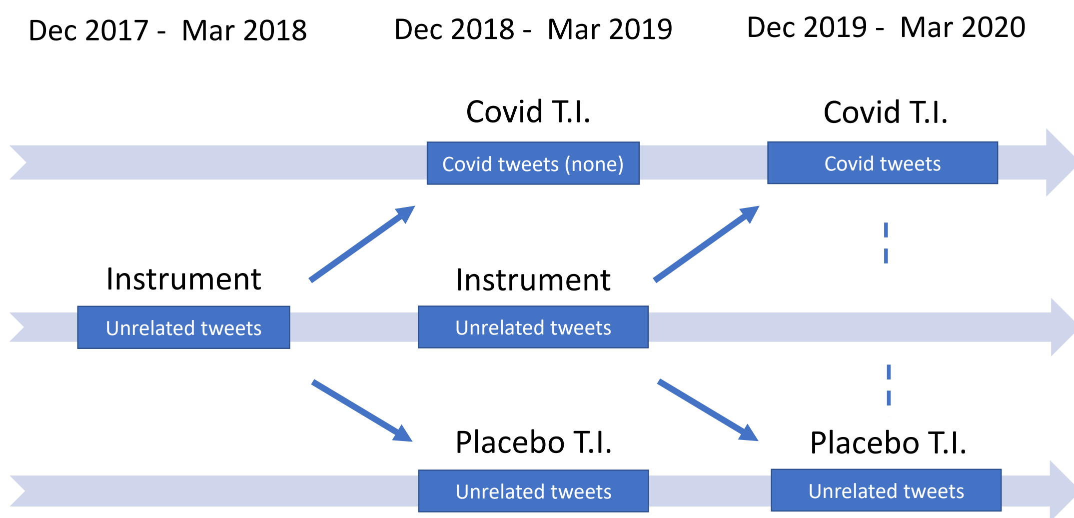
	Dependent Variable: Log Visits							
	Chinese	Asian	Latin	European	Chinese	Asian	Latin	European
Term T. I. x Post x 2020 x Cuisine	0.033**	0.033**	0.003	0.031**	-0.005	-0.010	-0.004	0.000
	(0.015)	(0.014)	(0.008)	(0.013)	(0.015)	(0.014)	(0.008)	(0.013)
Term T. I. x Post x 2020	0.039***	0.039***	0.039***	0.039***	0.066***	0.062***	0.068***	0.068***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Post x 2020 x Cuisine	-0.010***	0.070***	0.037***	0.016***	-0.022***	0.051***	0.025***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Post x 2020	-0.145***	-0.145***	-0.145***	-0.145***	-0.142***	-0.143***	-0.143***	-0.143***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cases x Cuisine	-0.068***	-0.036***	0.014***	-0.035***	-0.064***	-0.030***	0.015***	-0.032***
	(0.006)	(0.004)	(0.003)	(0.006)	(0.006)	(0.004)	(0.003)	(0.006)
Cases	-0.079***	-0.079***	-0.079***	-0.079***	-0.089***	-0.089***	-0.089***	-0.089***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Total Obs.	6,456,481	6,682,334	7,718,167	6,555,012	6,456,481	6,682,334	7,718,167	6,555,012
R Squared	0.870	0.868	0.868	0.870	0.872	0.870	0.869	0.872
Tweet Count	10,677,411	10,677,411	10,677,411	10,677,411	10,677,411	10,677,411	10,677,411	10,677,411
FE: Restaurant	X	X	X	X	X	X	X	X
FE: County x Year					X	X	X	X
Full DDD Controls	X	X	X	X	X	X	X	X
Control Covid Cases	X	X	X	X	X	X	X	X
Controls[4] x Nat. Tweets	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table provides estimates obtained from equation 2.4, representing the main IV specification for placebo tweets. The reference group in all models is American restaurants. The full set of interactions of indicators for Placebo, Post, Year, and cuisine are included in placebo models. Cases is measured as reported covid cases in the prior two weeks. Control variables are included as interactions with Post x Year to capture the differential impact of the control in the post-treatment period. Control variables are also included as interactions with the national time series of tweets to capture the differential impact of the control in the term's salience. All controls are also included as interactions with cuisine. Standard errors clustered at the restaurant level for specifications including restaurant fixed effects, and clustered at the county x year level for specifications including county x year fixed effects.

B.3 Figures

Figure B.2.1: Twitter Timeline, December 2017 - March 2020[3]



[3]The figure displays a timeline of the Twitter data, displaying how each data set is used. Covid tweets are used in two observational years, 2019 - 2020, and have almost no count for the first three pre-treatment periods. Unrelated tweets are obtained for the two years, 2018 - 2019, and are used to construct the intrument. Placebo tweets are obtained for the two observational years, 2019 - 2020, and are used in the placebo specificitons.

Chapter 3 Appendix

C.1 Tables

Table C.1.1: Adding Controls to Baseline

	Dependent Var: log(Asylum Applications Share/Foreign Inflow Share)					
	Baseline	Control 1	Control 2	Control 3	Control 4	Control 5
Grant Rate	0.655^{***}	0.410^{**}	0.366^{**}	0.809^{***}	0.309^{*}	0.769^{***}
	(0.168)	(0.182)	(0.185)	(0.177)	(0.185)	(0.165)
GR x E:P (15+)	-0.111^{***}	-0.064	-0.105^{**}			-0.133^{***}
	(0.035)	(0.046)	(0.044)			(0.041)
GR x Rights	0.584^{*}	1.043^{***}	0.807^{**}	0.095	0.600^{**}	
	(0.312)	(0.359)	(0.334)	(0.305)	(0.261)	
GR x Network	0.311^{***}	0.343^{***}	0.308^{***}	0.341^{***}	0.418^{***}	0.313^{***}
	(0.072)	(0.084)	(0.079)	(0.073)	(0.086)	(0.079)
GR x GDP pc		-0.003				
		(0.014)				
GR x Hhold Cons. pc		-0.048				
		(0.038)				
GR x GDP growth			0.097^{**}			
			(0.043)			
GR x Hhold Cons. growth			-0.002			
			(0.091)			
GR x LFPR				-0.018		
				(0.040)		
GR x Unemp. Rate					0.152^{***}	
					(0.049)	
GR x GNI growth					0.122^{**}	
					(0.056)	
GR x CC						-0.040
						(0.114)
GR x GE						-0.124
						(0.135)
GR x PV						-0.105
						(0.200)
GR x RQ						-0.524^{**}
						(0.258)
GR x RL						0.530^{*}
						(0.293)
GR x VA						0.617
						(0.387)
N	7,132	6,263	6,269	7,132	6,257	7,184
R^2 Apps	0.956	0.94	0.941	0.955	0.941	0.955
R^2 Inflows	0.98	0.983	0.983	0.98	0.983	0.98
FE: Org x Year	X	X	X	X	X	X
FE: Org x Dest	X	X	X	X	X	X
FE: Dest x Year	X	X	X	X	X	X
Weighted	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Models are weighted by the origin-year level total applications lodged and immigrant inflows across destinations (see Equations 3.12 and 3.13). Standard errors are heteroskedasticity-robust. Explanatory variables are demeaned at the origin-year level. Grant Rate uses the 3 year, 1 lag specification, and Network is measured as log(1 + accumulated stock of granted applications since 2008). All other explanatory variables are specified as the (destination-origin) difference and are lagged 1 year.

Table C.1.2: Different Grant Rate Specifications

	Dependent Var: log(Asylum Applications Share/Foreign Inflow Share)								
	1Y-1L	1Y-2L	1Y-3L	2Y-1L	2Y-2L	2Y-3L	3Y-1L	3Y-2L	3Y-3L
Grant Rate	0.509^{***}	0.495^{***}	0.239^{*}	0.634^{***}	0.518^{***}	0.373^{**}	0.655^{***}	0.642^{***}	0.460^{***}
	(0.131)	(0.135)	(0.136)	(0.147)	(0.155)	(0.148)	(0.168)	(0.165)	(0.178)
GR x E:P (15-64)	−0.084	−0.060	0.038	−0.129	−0.039	0.060^{**}	−0.111	−0.045	0.054^{*}
	(0.028)	(0.028)	(0.028)	(0.031)	(0.032)	(0.029)	(0.035)	(0.033)	(0.032)
GR x Rights	0.229	0.341	−0.263	0.621^{**}	0.286	−0.250	0.584^{*}	0.424	0.110
	(0.249)	(0.265)	(0.263)	(0.281)	(0.278)	(0.292)	(0.312)	(0.309)	(0.322)
GR x Network	0.139	0.198	0.241	0.228	0.280	0.264	0.311	0.347	0.151
	(0.558)	(0.745)	(0.914)	(0.811)	(1.005)	(0.904)	(0.998)	(1.122)	(0.469)
N	6,440	6,295	6,158	6,988	6,840	6,148	7,132	6,427	5,673
R^2 Apps	0.952	0.954	0.954	0.955	0.954	0.954	0.956	0.955	0.956
R^2 Inflows	0.980	0.980	0.980	0.980	0.980	0.982	0.980	0.981	0.984
FE: Org x Year	X	X	X	X	X	X	X	X	X
FE: Org x Dest	X	X	X	X	X	X	X	X	X
FE: Dest x Year	X	X	X	X	X	X	X	X	X
Weighted	X	X	X	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Models are weighted by the origin-year level total applications lodged and immigrant inflows across destinations (see Equations 3.12 and 3.13). Standard errors are heteroskedasticity-robust. Models provide baseline estimates with different specifications for the grant rate by window and lag. Explanatory variables are demeaned at the origin-year level. E:P (15-64) and Rights refer to the difference in the employment to population ratio (destination - origin) for those 15 to 64 and the difference in rights index (destination - origin), respectively, which are both lagged 1 year. Network is measured as log(1 + accumulated stock of granted applications since 2008).

Table C.1.3: Robustness Checks

	Dependent Var: log(Asylum Applications Share/Foreign Inflow Share)									
	Baseline	Geneva	No Germany	No 2015-16 Syria	No Greece	No Hungary	Add USA	Occ Inflows	Edu Inflows	Fam Inflows
Grant Rate	0.655***	0.739***	0.604***	0.545***	0.593***	0.604***	0.750***	0.612**	1.263***	0.757***
	(0.168)	(0.180)	(0.170)	(0.161)	(0.166)	(0.160)	(0.162)	(0.246)	(0.225)	(0.171)
GR x E:P (15-64)	−0.111***	−0.133***	−0.124***	−0.104***	−0.023	−0.130***	−0.088**	−0.032	−0.180***	−0.143***
	(0.035)	(0.041)	(0.037)	(0.035)	(0.034)	(0.035)	(0.036)	(0.053)	(0.057)	(0.035)
GR x Rights	0.584*	0.825**	0.852***	0.806***	0.381	0.724**	0.501*	0.097	0.600	1.128***
	(0.312)	(0.327)	(0.317)	(0.299)	(0.303)	(0.305)	(0.279)	(0.410)	(0.396)	(0.312)
GR x Network	0.311***	0.383***	0.296***	0.303***	0.349***	0.342***	0.331***	0.204*	0.327***	0.217***
	(0.072)	(0.089)	(0.075)	(0.070)	(0.066)	(0.071)	(0.065)	(0.105)	(0.094)	(0.072)
N	7,132	6,491	6,491	7,086	6,887	7,030	7,926	4,175	4,283	6,581
R^2 Apps	0.956	0.944	0.944	0.945	0.957	0.957	0.952	0.967	0.968	0.955
R^2 Inflows	0.980	0.981	0.981	0.981	0.984	0.980	0.982	0.976	0.987	0.983
FE: Org x Year	X	X	X	X	X	X	X	X	X	X
FE: Org x Dest	X	X	X	X	X	X	X	X	X	X
FE: Dest x Year	X	X	X	X	X	X	X	X	X	X
Weighted	X	X	X	X	X	X	X	X	X	X

*p < 0.1, **p < 0.05, ***p < 0.01

Note: Models are weighted by the origin-year level total applications lodged and immigrant inflows across destinations (see Equations 3.12 and 3.13). Standard errors are heteroskedasticity-robust. Each alternate specification is an adjustment from the baseline model. Geneva model presents the baseline model where grant rates include accepted applications only from those granted asylum status and do not include humanitarian status. Columns 3-6 remove Germany observations, 2015-2016 Syria observations, Greece observations, and Hungary observations, respectively. Column 7 adds US observations from UNHCR. "Occ", "Edu", and "Fam" models use only permits awarded for each of those reasons as the non-asylum inflow. Explanatory variables are demeaned at the origin-year level. E:P (15-64) and Rights refer to the difference in the employment to population ratio (destination - origin) for those 15 to 64 and the difference in rights index (destination - origin), respectively, which are both lagged 1 year. Grant rate uses the 3 year, 1 lag specification. Network is measured as log(1 + accumulated stock of granted applications since 2008).

Destination Countries:

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Germany
- Denmark
- Estonia
- Finland
- France
- Great Britain
- Greece
- Hungary
- Iceland
- Italy
- Liechtenstein
- Lithuania
- Luxembourg
- Malta
- Netherlands
- Norway
- Poland
- Portugal
- Romania
- Slovenia
- Slovakia
- Spain
- Sweden
- Switzerland

Origin Countries:

- Afghanistan
- Albania
- Algeria
- Angola
- Argentina
- Armenia
- Azerbaijan
- Bangladesh
- Belarus
- Benin
- Bosnia and Herz.
- Brazil
- Burkina Faso
- Burundi
- Cameroon
- Central African Rep.
- Chad
- China
- Colombia
- Comoros
- Congo
- Côte d'Ivoire
- Cuba
- Djibouti
- DR of Congo
- Ecuador
- Egypt
- El Salvador
- Eritrea
- Ethiopia
- Gabon
- Gambia
- Georgia
- Ghana
- Guatemala
- Guinea
- Guinea-Bissau
- Haiti
- Honduras
- India
- Indonesia
- Iran
- Iraq
- Jamaica
- Jordan
- Kazakhstan
- Kenya
- Kuwait
- Kyrgyzstan
- Lebanon
- Liberia
- Libya
- Malawi
- Malaysia
- Mali
- Mauritania
- Mexico
- Moldova
- Mongolia
- Montenegro
- Morocco
- Myanmar
- Nepal
- Nicaragua
- Niger

- Nigeria
- North Macedonia
- Pakistan
- Peru
- Phillippines
- Russia
- Rwanda
- Saudi Arabia
- Senegal
- Serbia
- Sierra Leone
- South Africa
- Sri Lanka
- Sudan
- Syria
- Tajikistan
- Tanzania
- Togo
- Tunisia
- Turkey
- Turkmenistan
- Uganda
- Ukraine
- Uzbekistan
- Venezuela
- Vietnam
- Yemen
- Zimbabwe

Bibliography

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of The Nazis in Prewar Germany." *The Quarterly Journal of Economics*, 130(4): 1885–1939.

Agan, Amanda, and Sonja Starr. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *The Quarterly Journal of Economics*, 133(1): 191–235.

Albert, Christoph, and Joan Monras. 2018. "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin." C.E.P.R. Discussion Papers 12842.

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*, 31(2): 211–236.

Allport, Gordon. 1954. *The Nature of Prejudice.* Oxford:Addison-Wesley.

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics*, 133(4): 1885–1932.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives*, 32(3): 259–280.

Bartoš, Vojtěch, Michal Bauer, Jana Cahlíková, and Julie Chytilová. 2021. "Covid-19 Crisis and Hostility against Foreigners." *European Economic Review*, 137: 103818.

Basavaraju, Sridhar V, Monica E Patton, Kacie Grimm, Mohammed Ata Ur Rasheed, Sandra Lester, Lisa Mills, Megan Stumpf, Brandi Freeman, Azaibi Tamin, Jennifer Harcourt, Jarad Schiffer, Vera Semenova, Han Li, Bailey Alston, Muyiwa Ategbole, Shanna Bolcen, Darbi Boulay, Peter Browning, Li Cronin, Ebenezer David, Rita Desai, Monica Epperson, Yamini Gorantla, Tao Jia, Panagiotis Maniatis, Kimberly Moss, Kristina Ortiz, So Hee Park, Palak Patel, Yunlong Qin, Evelene Steward-Clark, Heather Tatum, Andrew Vogan, Briana Zellner, Jan Drobeniuc, Matthew R P Sapiano, Fiona Havers, Carrie Reed, Susan Gerber, Natalie J Thornburg, and Susan L Stramer. 2021. "Serologic Testing of US Blood Donations to Identify Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)–Reactive Antibodies: December 2019–January 2020." *Clinical Infectious Diseases*, 72(12): 1004–1009.

Basso, Gaetano, Francesco D'Amuri, and Giovanni Peri. 2019. "Immigrants, Labor Market Dynamics and Adjustment to Shocks in the Euro Area." *IMF Economic Review*, 67(3): 528–572.

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.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Bertoli, Simone, Herbert Brücker, and Jesús Fernández-Huertas Moraga. 2022. "Do Applications Respond to Changes in Asylum Policies in European Countries?" *Regional Science and Urban Economics*, 93: 103771.

Bertrand, M., and E. Duflo. 2017. "Chapter 8 - Field Experiments on Discrimination." *Handbook of Economic Field Experiments*, 1: 309–393.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2022. "Inaccurate Statistical Discrimination: An Identification Problem." *Review of Economics and Statistics*.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2019. "Beliefs about Gender." *American Economic Review*, 109(3): 739–773.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." National Bureau of Economic Research w2248.

Borjas, George J. 2001. "Does Immigration Grease the Wheels of the Labor Market?" *Brookings Papers on Economic Activity*, 2001(1): 69–133.

Braghieri, Luca, Ro'ee Levy, and Alexey Makarin. 2022. "Social Media and Mental Health." *American Economic Review*, 112(11): 3660–3693.

Brekke, Jan-Paul, and Grete Brochmann. 2015. "Stuck in Transit: Secondary Migration of Asylum Seekers in Europe, National Differences, and the Dublin Regulation." *Journal of Refugee Studies*, 28(2): 145–162.

Brekke, Jan-Paul, Marianne Roed, and Pål Schøne. 2016. "Reduction or Deflection? The Effect of Policy on Interconnected Asylum Flows." Social Science Research Network SSRN Scholarly Paper ID 2753100, Rochester, NY.

Bursztyn, Leonardo, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott. 2022*a*. “Opinions as Facts.” *ECONtribute Discussion Paper No. 159.*

Bursztyn, Leonardo, Georgy Egorov, Ingar Haaland, Aakaash Rao, and Christopher Roth. 2022*b*. “Scapegoating during Crises.” *AEA Papers and Proceedings*, 112: 151–155.

Bursztyn, Leonardo, Georgy Egorov, Ruben Enikolopov, and Maria Petrova. 2019. “Social Media and Xenophobia: Evidence from Russia.” *NBER Working Paper Series No. w26567.*

Cable News Network, CNN. 2020. “Racist attacks on Asians spreading faster than coronavirus in US.” *https://www.cnn.com/2020/02/20/us/coronavirus-racist-attacks-against-asian-americans/index.html*, accessed 02/21/2020.

Cadena, Brian C. 2013. “Native Competition and Low-Skilled Immigrant Inflows.” *Journal of Human Resources*, 48(4): 910–944.

Cadena, Brian C., and Brian K. Kovak. 2016. “Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession.” *American Economic Journal: Applied Economics*, 8(1): 257–290.

Cao, Andy, Jason M. Lindo, and Jiee Zhong. 2022. “Can Social Media Rhetoric Incite Hate Incidents? Evidence from Trump’s ”Chinese Virus” Tweets.” *NBER Working Paper Series No. w30588.*

Card, David. 2001. “Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration.” *Journal of Labor Economics*, 19(1): 22–64.

Card, David, and Ethan G. Lewis. 2007. “The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts.” National Bureau of Economic Research c0095.

Carlana, Michela. 2019. “Implicit Stereotypes: Evidence from Teachers’ Gender Bias.” *The Quarterly Journal of Economics*, 134(3): 1163–1224.

Carrell, Scott E., Mark Hoekstra, and James E. West. 2019. “The Impact of College Diversity on Behavior toward Minorities.” *American Economic Journal: Economic Policy*, 11(4): 159–182.

Center for Disease Control, CDC. 2020. “Coronavirus Disease 2019.” *https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html*, accessed 03/12/2023.

Chen, Jiafeng, and Jonathan Roth. 2023. “Logs with Zeros? Some Problems and Solutions.” *Quarterly Journal of Economics*, The Quarterly Journal of Economics.

Chong, Eugene. 2021. “Simple Methods for Normalizing SafeGraph Patterns Data Over Time.” *https://colab.research.google.com/drive/1lTi8JXfX9rh2mnuFjgYgKsIcMcDI3EG_?usp=sharing#scrollTo=useful-margin*.

Cinelli, Matteo, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. 2020. “The COVID-19 Social Media Infodemic.” *Scientific Reports*, 10(1): 16598.

Coffman, Katherine B., Christine L. Exley, and Muriel Niederle. 2021. “The Role of Beliefs in Driving Gender Discrimination.” *Management Science*, 67: 3321–3984, iii–iv.

Connor, Phillip. 2010. "Explaining the Refugee Gap: Economic Outcomes of Refugees versus Other Immigrants." *Journal of Refugee Studies*, 23(3): 377–397.

Corno, Lucia, Eliana La Ferrara, and Justine Burns. 2022. "Interaction, Stereotypes, and Performance: Evidence from South Africa." *American Economic Review*, 112(12): 3848–3875.

DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya. 2014. "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia." *American Economic Journal: Applied Economics*, 6(3): 103–132.

Department of Justice, DOJ. 2021. "2020 FBI Hate Crimes Statistics." *https://www.justice.gov/crs/highlights/2020-hate-crimes-statistics*, Accessed 01/30/2023.

Diop-Christensen, Anna, and Lanciné Eric Diop. 2021. "What Do Asylum Seekers Prioritise—Safety or Welfare Benefits? The Influence of Policies on Asylum Flows to the EU15 Countries." *Journal of Refugee Studies*, feab077.

Dipoppa, Gemma, Guy Grossman, and Stephanie Zonszein. 2021. "Locked Down, Lashing Out: Situational Triggers and Hateful Behavior Towards Minority Ethnic Immigrants." *SSRN Working Paper No. 3789339.*

Doleac, Jennifer L., and Luke C.D. Stein. 2013. "The Visible Hand: Race and Online Market Outcomes." *The Economic Journal*, 123(572): F469–F492.

Dulleck, Uwe, Jonas Fooken, and Yumei He. 2020. "Hukou Status and Individual-Level Labor Market Discrimination: An Experiment in China." *ILR Review*, 73(3): 628–649.

Economist Intelligence Unit, EIU. 2021. "Economist Intelligence Unit - Democracy Index Annual Reports." *https://www.eiu.com/n/campaigns/democracy-index-2020/*, accessed 11/21/2021.

Esponda, Ignacio, Ryan Oprea, and Sevgi Yuksel. 2023. "Seeing What Is Representative." *The Quarterly Journal of Economics*, 138(4): 2607–2657.

European Commission, European Commission. 2016. "European Commission - PRESS RELEASES - Press Release - Questions & Answers: Recommendation on the Conditions for Resuming Dublin Transfers of Asylum Seekers to Greece (2016)." *https://ec.europa.eu/commission/presscorner/detail/en/MEMO_16_4253*.

European Commission, European Commission. 2020. "Common European Asylum System." *https://ec.europa.eu/home-affairs/policies/migration-and-asylum/common-european-asylum-system_en*.

European Commission, Eurostat. 2021. "Eurostat." *https://ec.europa.eu/eurostat/web/main/data/database*, accessed 11/21/2021.

Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *The Quarterly Journal of Economics*, 116(1): 351–377.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

Google. 2022. "Google Trends Keyword Search: coronavirus." *retrieved November 6, 2022 via R package gtrendsR.*

Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *The American Economic Review*, 84(3): 622–641.

Guevara, C. Angelo, Alejandro Tirachini, Ricardo Hurtubia, and Thijs Dekker. 2020. "Correcting for Endogeneity Due to Omitted Crowding in Public Transport Choice Using the Multiple Indicator Solution (MIS) Method." *Transportation Research Part A: Policy and Practice*, 137: 472–484.

Hatton, Timothy J. 2005. "European Asylum Policy." *National Institute Economic Review*, 194(1): 106–119.

Hatton, Timothy J. 2009. "The Rise and Fall of Asylum: What Happened and Why?" *The Economic Journal*, 119(535): F183–F213.

Hedegaard, Morten Størling, and Jean-Robert Tyran. 2018. "The Price of Prejudice." *American Economic Journal: Applied Economics*, 10(1): 40–63.

Holzer, Harry J., and Keith R. Ihlanfeldt. 1998. "Customer Discrimination and Employment Outcomes for Minority Workers." *The Quarterly Journal of Economics*, 113(3): 835–867.

Holzer, Thomas, Gerald Schneider, and Thomas Widmer. 2000. "The Impact of Legislative Deterrence Measures on the Number of Asylum Applications in Switzerland (1986–1995)." *International Migration Review*, 34(4): 1182–1216.

Huang, Justin T., Masha Krupenkin, David Rothschild, and Julia Lee Cunningham. 2023. "The Cost of Anti-Asian Racism during the COVID-19 Pandemic." *Nature Human Behaviour*, 1–14.

Jaeger, David A. 2007. “Green Cards and the Location Choices of Immigrants in the United States, 1971–2000.” In *Immigration*. Vol. 27 of *Research in Labor Economics*, , ed. Barry R. Chiswick, 131–183. Emerald Group Publishing Limited.

Jaeger, David A., Theodore J. Joyce, and Robert Kaestner. 2020. “A Cautionary Tale of Evaluating Identifying Assumptions: Did Reality TV Really Cause a Decline in Teenage Childbearing?” *Journal of Business & Economic Statistics*, 38(2): 317–326.

Kanazawa, Mark T., and Jonas P. Funk. 2001. “Racial Discrimination in Professional Basketball: Evidence from Nielsen Ratings.” *Economic Inquiry*, 39(4): 599–608.

Kearney, Melissa S., and Phillip B. Levine. 2015. “Media Influences on Social Outcomes: The Impact of MTV’s ”16 and Pregnant” on Teen Childbearing.” *American Economic Review*, 105(12): 3597–3632.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. “Systemic Discrimination Among Large U.S. Employers*.” *The Quarterly Journal of Economics*, 137(4): 1963–2036.

Latoja, Hill, and Samantha Artiga. 2022. “COVID-19 Cases and Deaths by Race/Ethnicity: Current Data and Changes Over Time.”

Levy, Ro’ee. 2021. “Social Media, News Consumption, and Polarization: Evidence from a Field Experiment.” *American Economic Review*, 111(3): 831–870.

List, John A. 2004. “The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field.” *The Quarterly Journal of Economics*, 119(1): 49–89.

Lu, Runjing, and Sophie Yanying Sheng. 2022. "How Racial Animus Forms and Spreads: Evidence from the Coronavirus Pandemic." *Journal of Economic Behavior & Organization*, 200: 82–98.

Luca, Michael, Elizaveta Pronkina, and Michelangelo Rossi. 2022. "Scapegoating and Discrimination in Times of Crisis: Evidence from Airbnb." *NBER Working Paper Series No. w30344*.

Marmaros, David, and Bruce Sacerdote. 2006. "How Do Friendships Form?" *The Quarterly Journal of Economics*, 121(1): 79–119.

McFadden, Daniel. 1973. *Conditional Logit Analysis of Qualitative Choice Behavior.* Institute of Urban and Regional Development, University of California.

Mengel, Friederike, Jan Sauermann, and Ulf Zolitz. 2019. "Gender Bias in Teaching Evaluations." *Journal of the European Economic Association*, 17(2): 535–566.

MIT Election Data and Science via Harvard Dataverse, MIT. 2023. "County Presidential Election Returns 2000-2020." *https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ*, accessed 11/18/2023.

Mobius, Markus M., and Tanya S. Rosenblat. 2006. "Why Beauty Matters." *American Economic Review*, 96(1): 222–235.

Müller, Karsten, and Carlo Schwarz. 2021. "Fanning the Flames of Hate: Social Media and Hate Crime." *Journal of the European Economic Association*, 19(4): 2131–2167.

Müller, Karsten, and Carlo Schwarz. 2023. "From Hashtag to Hate Crime: Twitter and Antiminority Sentiment." *American Economic Journal: Applied Economics*, 15(3): 270–312.

Nardotto, Mattia, and Sandra Sequeira. 2021. "Identity, Media and Consumer Behavior." *CEPR Press Discussion Paper No. 15765*.

National Public Radio, NPR. 2020. "Coronavirus: New York Creates 'Containment Area' Around Cluster In New Rochelle." *https://www.npr.org/sections/health-shots/2020/03/10/814099444/new-york-creates-containment-area-around-cluster-in-new-rochelle*, accessed 02/12/2023.

Neumayer, Eric. 2004. "Asylum Destination Choice: What Makes Some West European Countries More Attractive Than Others?" *European Union Politics*, 5(2): 155–180.

New York Times, NYT. 2022. "New York Times Coronavirus Cases and Deaths Archive." *https://github.com/nytimes/covid-19-data*, accessed 10/12/2022.

Olden, Andreas, and Jarle Møen. 2022. "The Triple Difference Estimator." *The Econometrics Journal*, 25(3): 531–553.

Pandya, Sonal S., and Rajkumar Venkatesan. 2016. "French Roast: Consumer Response to International Conflict—Evidence from Supermarket Scanner Data." *Review of Economics and Statistics*, 98(1): 42–56.

Petrin, Amil, and Kenneth Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research*, 47(1): 3–13.

Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology*, 90: 751–783.

Pew Research Center, (Pew). 2013. "Social Networking Site Users." *https://www.pewresearch.org/internet/2013/02/14/social-networking-site-users/.*

Pew Research Center, (Pew). 2020. "Differences in How Democrats and Republicans Behave on Twitter."

Pew Research Center, (Pew). 2021. "The Views and Experiences of U.S. Adult Twitter Users."

Pew Research Center, (Pew). 2022. "Social Media and News Fact Sheet | Pew Research Center." *https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/.*

Romano, Sebastian D. 2021. "Trends in Racial and Ethnic Disparities in COVID-19 Hospitalizations, by Region — United States, March–December 2020." *MMWR. Morbidity and Mortality Weekly Report*, 70.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. 2023. "ACS 2019 5-Year Sample." `https://doi.org/10.18128/D010.V14.0`, IPUMS USA: Version 14.0. Minneapolis, MN: IPUMS, 2023.

SafeGraph. 2022. "SafeGraph: Understand Consumer Behavior With Precise Foot Traffic Data." `https://www.safegraph.com/products/patterns`, retrieved April 9, 2022.

Scanlon, Dennis P., Michael Chernew, Catherine McLaughlin, and Gary Solon. 2002. "The Impact of Health Plan Report Cards on Managed Care Enrollment." *Journal of Health Economics*, 21(1): 19–41.

Simonov, Andrey, Szymon K. Sacher, Jean-Pierre H. Dubé, and Shirsho Biswas. 2020. "The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the Covid-19 Pandemic." *NBER Working Paper Series No. w27237*.

Squire, Ryan Fox. 2019. "What about bias in your dataset? Quantifying Sampling Bias in SafeGraph Patterns." *https://colab.research.google.com/drive/1u15afRytJMsizySFqA2EPlXSh3KTmNTQ#sandboxMode=true&scrollTo=oBADlZgHTsWg*.

Takhteyev, Yuri, Anatoliy Gruzd, and Barry Wellman. 2012. "Geography of Twitter Networks." *Social Networks*, 34(1): 73–81.

The World Bank, The World Bank. 2021. "World Bank Databank: World Development Indicators." *https://databank.worldbank.org/source/world-development-indicators#*, accessed 11/21/2021.

Thielemann, Eiko R. 2006. "The Effectiveness of Governments' Attempts to Control Unwanted Migration." In *Immigration and the Transformation of Europe.* , ed. Craig A. Parsons and Timothy M. Smeeding, 442–472. Cambridge:Cambridge University Press.

Time Magazine, Time Magazine. 2020. "As Coronavirus Spreads, So Does Xenophobia and Racism — Time." *https://time.com/5797836/coronavirus-racism-stereotypes-attacks/*, accessed 04/24/2023.

Toshkov, Dimiter, and Laura de Haan. 2013. "The Europeanization of Asylum Policy: An Assessment of the EU Impact on Asylum Applications and Recognitions Rates." *Journal of European Public Policy*, 20(5): 661–683.

Toshkov, Dimiter Doychinov. 2014. "The Dynamic Relationship between Asylum Applications and Recognition Rates in Europe (1987–2010)." *European Union Politics*, 15(2): 192–214.

United Nations High Commissioner for Refugees, UNHCR. 2017. "UNHCR urges suspension of transfers of asylum-seekers to Hungary under Dublin." *https://www.unhcr.org/en-us/news/press/2017/4/58eb7e454/unhcr-urges-suspension-transfers-asylum-seekers-hungary-under-dublin.html*.

United Nations High Commissioner for Refugees, UNHCR. 2021. "UNHCR Refugee Statistics." *https://www.unhcr.org/refugee-statistics/download/?url=o5c9VE*, accessed 11/21/2021.

Uruci, Xhulio. 2024a. "Animus or Beliefs? Consumer Discrimination In Uncertain Times."

US Census. 2019*a*. "U.S. Census Bureau's 2018 – 2019 Population Estimates Program (PEP)." *https://www.census.gov/programs-surveys/popest/data/tables.html*, accessed November 10, 2023.

US Census. 2019*b*. "U.S. Census Bureau's 2019 Statistics of U.S. Businesses (SUSB)." *https://www.census.gov/data/tables/2019/econ/susb/2019-susb-annual.html*, accessed November 23, 2023.

Vink, Maarten, and Frits Meijerink. 2003. "Asylum Applications and Recognition Rates in EU Member States 1982–2001: A Quantitative Analysis." *Journal of Refugee Studies*, 16(3): 297–315.

Wang, Tianyi. 2021. "Media, Pulpit, and Populist Persuasion: Evidence from Father Coughlin." *American Economic Review*, 111(9): 3064–3092.

WHO. 2023. "United States of America: WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data." *https://covid19.who.int/*.

World Health Organization, WHO. 2023. "United States of America: WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data." *https://covid19.who.int/*, accessed 01/30/2023.

Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." *The Quarterly Journal of Economics*, 129(4): 1947–1994.

Yi, Guanting. 2023. "The Economic Effect of Discrimination: Evidence from Restaurant Sector." *SSRN Working Paper No. 4214241*.

Zorlu, Aslan, and Clara H. Mulder. 2008. "Initial and Subsequent Location Choices of Immigrants to the Netherlands." *Regional Studies*, 42(2): 245–264.

Zussman, Asaf. 2023. "Discrimination in Times of Crises and the Role of the Media." *American Economic Journal: Applied Economics*.