

**Institut d'études politiques de Paris  
ÉCOLE DOCTORALE DE SCIENCES PO  
Programme doctoral en économie  
Département d'Économie  
Doctorat en sciences économiques**

**Essays in Economics of Discrimination and  
Diversity**

Vladimir Avetian

Thesis supervised by  
Sergei Guriev, Professeur des universités d'Économie, Sciences Po  
Paris

Jury:

Mr. Eric Brousseau

Professor, University Paris-Dauphine

Mr. Ruben Durante

Research Professor, Universitat Pompeu Fabra (*reviewer*)

Mr. Roberto Galbiati

Directeur de recherche, CNRS-DESP (président)

Mr. Sergei Guriev

Professeur des universités d'Économie, Sciences Po Paris

Mrs. Pamina Koenig

Professeur des universités, Université de Rouen (*reviewer*)

Mr. Benjamin Marx

Assistant Professor, Sciences Po



# Acknowledgements

I want to express my gratitude to my family, friends and colleagues who were there for me while I was working on this dissertation.

I would like to start with my supervisor, Sergei Guriev. I first saw Sergei more than ten years ago, when I participated in a Russian high-school competition in economics. Sergei was giving a lecture in front of us, high school students. I remember him showing us graphs of the GDP of South Korea and Russia. Years later, I was lucky enough to become a PhD student of Sergei. And while the promise of the Russian economy did not come true (today Russia is more imitating South Korea's sibling), another promise has come true: that research is interesting! I am infinitely indebted to Sergei for sustaining my interest in research and helping to develop it into something tangible, these pages. I am just as grateful to him for giving me complete freedom to develop my own ideas, because this freedom is a necessary condition for real scientific work. Without his patient mentoring, support, and wise advice, this dissertation would not have existed.

I am also immensely grateful to all the researchers I met during my studies. I would like to thank Ekaterina Zhuravskaya, with whom I had the opportunity to work as an RA. She gave me an example of a truly enthusiastic researcher who takes on problems of the first order and solves them with rigor and precision. Without her example and always sharp advice, I would not have even passed my first regression. I would like to thank my committee Pierre Philippe Combes and Florian Oswald who were guiding me throughout my PhD and nerve-racking job market. I am grateful to Roland Bénabou, who hosted me during my brief and surreal exchange with Princeton, that also coincided with first wave of Covid. I am grateful to Yann Algan, with whom I also had a chance to work as RA. I would like to thank Sciences Po faculty members who generously commented my work during the seminars and in private talks. In particular, I would like to thank Pierre Cahuc, Jeanne Commault, Clément de Chaisemartin, Golvine de Rochambeau, Michelle Fiorretti, Roberto Galbiati, Emeric Henry, Benjamin Marx, Stefan Pollinger.

I would also like to take this opportunity to thank Eric Brousseau, Ruben Durante, Roberto Galbiati, Pamina Koenig and Benjamin Marx who have agreed to be part of my

jury.

I was also lucky to have a fantastic co-authors on my PhD projects. Each of them is a unique researcher, and I often feel like their student. Stefan Pauly brightened up my gloomy PhD days and gave me an example of a rigorous and versatile researcher who is always willing to try something new. I am grateful to my wonderful team, Annalí, Sulin and Kritika — next generation researchers, dedicated and compassionate.

I also want to say thank you to my PhD friends. My friends Ilya and Masha Eryzhenskiy, who took care of me as if I was their second child. Ilya studied with me every corner of my dissertation, helped me to improve my writing, and even more — went to weekly music sessions with me. From day one my friends Andrei, Mai, and Alex discussed research ideas with me, and we had fun. I want to thank Kirill and Dasha who made my trip to Princeton worth it, even though it was a lockdown. I am proud and happy that I met several PhD cohorts at Sciences Po, Princeton and PSE. I'm afraid of missing someone: Tyler Abbot, Victor Augias, Daniel Barreto, Stephane Benviste, Oliver Cassagneau-Francis, Sophie Cetre, Edoardo Ciscato, Naomi Cohen, Pauline Corblet, Pierre Cotterlaz, Florin Cucu, Nicolo Dalvit, Pierre Deschamps, Samuel Delpeuch, Edgard Dewitte, Mylène Feuillade, Etienne Fize, Emanuele Franceschi, Alexis Ghersengorin, Arthur Guillouzouic Le Corff, Nourhan Hashish, Moritz Hengel, Riddhi Kalsi, Jean-Louis Keene, Gustave Kenedi, Ségal Le Guern, Marco Ranaldi, Léonard Le Roux, Juan Sebastian Ivars, Charles Louis-Sidois, Valentin Marchal, Clément Mazet-Sonilhac, Nikita Melnikov, Alaïs Martin-Baillon, Andreea Minea, Julia Mink, Elisa Mougin, Marco Palladino, Ludovic Panon, Julien Pascal, Aseem Patel, Stefan Pauly, Victor Saint- Jean, Ariane Salem, Jérôme Sansonetti, Sarah Schneider-Strawczynski, Jan Sontag, Camille Urvoy, Lucas Vernet, Pierre Villedieu and Zydney Wong. I was happy to be around these bright people all these years.

I am eternally grateful to my friends who supported me during these not always simple years in Paris: Alex, Andrei, Fabio, François-René, Gulnaz, Houssem, Jenia, Kostia, Mai, Masha, Nastia, Sam, Sebastian, Stefan, Stéphane, Tania, Timur, Tuyana, Vassa, Vika, Zara. I am very grateful to my French teacher and friend Anthony De Zolt.

Thanks to my friends with whom I was separated physically (who were living in Russia or other countries), but who were anyway supporting me a lot. Jenia who always was ready to listen to my woes in our telegram chat. Philipp with whom we managed to create our online music band. My dear friends: Alena, Andrei, Dima, Egor, Eric, Galya, Lecha, Nikita, Oleg, Pasha, Pasha, Rita, Timur, Yulia, Vania, Vlada.

Most importantly, I am infinitely grateful to my family. To my mother, who wanted me to become a medical doctor. With this dissertation, I hope I will half-fulfill her dream. To my father, to my dearest aunt. You were always supporting me and encouraging my

creativity. Without you, I wouldn't exist. I also want to thank my family from the side of my beloved wife: Lyudmila, Sergei, Natasha, Nadezhda, Zinaida. I always happy to be around you.

And finally, you, Lena, my patient wife, I know you will get to this place — you are my best friend, my soulmate. The word 'support' does not fully describe what you are doing to me. I dedicate this work to you.



# Contents

<b>Sommaire</b>	<b>17</b>
<b>Introduction</b>	<b>19</b>
<b>1 Consider the Slavs: Overt Discrimination and Racial Disparities in Rental Housing</b>	<b>31</b>
1 Introduction . . . . .	32
2 Background and Data . . . . .	36
2.1 Cian data . . . . .	37
2.2 Other data . . . . .	39
3 Empirical analysis . . . . .	40
3.1 Estimating equation . . . . .	40
3.2 Main results . . . . .	41
3.3 Results of experiment . . . . .	43
4 Theory . . . . .	44
4.1 Baseline model . . . . .	44
4.2 Search . . . . .	45
4.3 Racial rent differentials in two neighborhoods . . . . .	47
5 Conclusion . . . . .	49
6 Tables . . . . .	58
A Appendix: Design of Correspondence Experiment . . . . .	64
A.1 Messages . . . . .	64
A.2 Names and identities . . . . .	65
A.3 Sending messages . . . . .	66
A.4 Classification of responses . . . . .	66
B Appendix: Empirical Results . . . . .	68
C Appendix: Theory . . . . .	72
C.1 Tenants' problems . . . . .	72

C.2	Optimal Rents and Rent Differential in a Separate Neighborhood . . . . .	72
C.3	Equilibrium . . . . .	73
<b>2</b>	<b>Urban Amenities and Tourism: Evidence from Tripadvisor</b>	<b>75</b>
1	Introduction . . . . .	76
2	Background and Data . . . . .	79
2.1	COVID-19 in Paris . . . . .	79
2.2	Tripadvisor Data . . . . .	79
2.3	Measuring Tourism . . . . .	80
2.4	Content of Reviews . . . . .	81
2.5	Dans Ma Rue . . . . .	82
2.6	Social Connectedness Index . . . . .	82
3	Stylized Facts . . . . .	83
4	Empirical Strategy . . . . .	84
4.1	Restaurant-level Approach . . . . .	84
4.2	Review-level Approach . . . . .	84
5	Results . . . . .	85
6	Robustness & Further Results . . . . .	86
6.1	Neighborhood Complaints . . . . .	86
6.2	Bataclan Attacks . . . . .	86
6.3	Pre-Trends . . . . .	87
6.4	Spillovers . . . . .	88
6.5	Further Robustness Checks . . . . .	88
7	Mechanisms . . . . .	88
7.1	Overcrowding . . . . .	89
7.2	Supply-Side Changes . . . . .	89
7.3	Aversion . . . . .	89
8	Conclusion . . . . .	90
9	Tables . . . . .	99
A	Additional Plots . . . . .	107
B	Robustness Checks . . . . .	108
B.1	Alternative Identification: November 2015 Paris attacks . . . . .	108
B.2	Location-Based Tourism Measure . . . . .	109
B.3	Aggregation of Language-Based Tourism Measure by Different Periods	112
B.4	Clustering . . . . .	113
C	Validation of Tourism Measures . . . . .	114

D	Text Analysis . . . . .	115
<b>3</b>	<b>Going Viral in a Pandemic: Social Media and Allyship in the Black Lives Matter Movement</b>	<b>117</b>
1	Introduction . . . . .	118
2	Background and Data . . . . .	123
2.1	BLM History and Motivating Evidence . . . . .	123
2.2	Main Data Sources . . . . .	125
2.3	Descriptive statistics . . . . .	128
3	Empirical Strategy . . . . .	129
3.1	Baseline Estimating Equation . . . . .	129
3.2	IV Estimation: Super Spreader Events . . . . .	130
4	COVID-19 and BLM . . . . .	133
4.1	Main Results . . . . .	133
4.2	Heterogeneity . . . . .	134
4.3	Alternative Outcomes . . . . .	136
5	Social Media and BLM . . . . .	137
5.1	COVID-19 and the Use of Social Media . . . . .	137
5.2	Twitter and BLM protests . . . . .	140
5.3	News Consumption and Attitudes towards BLM . . . . .	143
6	Competing Mechanisms . . . . .	144
6.1	Broadening versus Scattering of Protest . . . . .	145
6.2	Salience of Racial Inequality . . . . .	147
6.3	Opportunity Cost of Protesting . . . . .	148
6.4	Agitation and Propensity to Protest . . . . .	148
7	Robustness . . . . .	149
7.1	Instrument validity . . . . .	149
7.2	Robustness of main results . . . . .	151
7.3	Alternative Identification Strategies . . . . .	152
8	Conclusion . . . . .	153
9	Figures and Tables . . . . .	162
A	Appendix: Robustness Checks . . . . .	181
A.1	Instrument Robustness . . . . .	181
A.2	Robustness of Main Results . . . . .	183
B	Appendix: Alternative Estimation Strategies . . . . .	191
B.1	Alternative Instrument: Florida Spring Break . . . . .	191

B.2	Difference in Differences: Notable Deaths Sample . . . . .	195
B.3	LASSO Matching: Propensity to Protest . . . . .	198
C	Appendix: Additional Figures and Tables . . . . .	200
D	Data Appendix . . . . .	210
<b>Conclusion</b>		<b>221</b>
<b>Resumé</b>		<b>3</b>

# List of Figures

1.1	Geography of discrimination . . . . .	54
1.2	Daily number of ads posted on the platform . . . . .	54
1.3	Share of discrimination by neighbourhoods on the first and last days of the observational period . . . . .	55
1.4	The racial rent differential by districts ( <i>rayons</i> ) . . . . .	57
2.1	Daily Number of Reviews in Paris (since launch of Tripadvisor) . . . . .	94
2.2	Daily Number of Reviews in Paris . . . . .	94
2.3	Map of restaurants by share of non-french reviews . . . . .	95
2.4	Grid map of restaurants density . . . . .	95
2.5	Tourist Access vs Tourism Proxy . . . . .	96
2.6	Share of French Cuisine . . . . .	96
2.7	Diversity of Cuisine Types . . . . .	97
2.8	Dynamic Effects . . . . .	98
2.A.1	Tripadvisor interface . . . . .	107
2.A.2	Correlating Different Tourism Proxies . . . . .	107
3.1	BLM events over time . . . . .	162
3.2	COVID-19 deaths and timing of GF's murder . . . . .	163
3.3	BLM events and tweets in counties with above and below median COVID-19 deaths per-capita . . . . .	164
3.4	Spatial distribution of US counties based on their BLM protest activities before and after George Floyd's murder . . . . .	165
3.5	Distribution of super-spreader events in the US by their type . . . . .	166
3.6	Window of opportunity for SSEs . . . . .	167
3.7	Timing of SSEs relative to Floyd's murder, protest and COVID-19 deaths . .	168
3.8	Construction of the super-spreading events instrument (example) . . . . .	169
3.9	Geographic distribution of super-spreader events (SSEs) . . . . .	170
3.B.1	Number of devices (log) by US counties pinged during March 1st, 2020 . .	192

3.B.2	Spring Breakers by US counties. Own visualization based on SafeGraph data.	193
3.C.1	Evolution of lockdown stringency index, and masks recommendations . . . . .	200
3.C.2	Evolution of mobility index . . . . .	201

# List of Tables

1.1	Descriptive statistics . . . . .	58
1.2	Main result: The Racial Rent Differential . . . . .	59
1.3	Heterogeneous effects: the Racial Rent Differential and the Share of Discrimination in Neighborhood . . . . .	60
1.4	Heterogeneous Effects: Interactions with Characteristics of Neighborhood . .	61
1.5	Experiment: Main Results . . . . .	62
1.6	Experiment: Subset of ads without overt discrimination . . . . .	63
1.B.1	The Racial Rent Differential: Extended Table . . . . .	68
1.B.2	Placebo: Other Preferences of Landlords . . . . .	69
1.B.3	Robustness: Phone Numbers Fixed Effects . . . . .	69
1.B.4	Increased Search Time: Discrimination and Number of Days before Ad Removed	70
1.B.5	Heterogeneity of Search Time Effect: Interaction with Share of Discrimination in Neighborhood . . . . .	71
1.B.6	Experiments Outcomes . . . . .	71
2.1	Stylized Facts: User Preferences . . . . .	99
2.2	Main Result: Tourism and Restaurant Ratings by Parisians (Restaurant-Level)	100
2.3	Main Result: Tourism and Restaurant Ratings by Parisians (Review-Level) .	101
2.4	Tourism and “Dans Ma Rue” Complaints . . . . .	102
2.5	Textual Outcomes . . . . .	103
2.6	Social Proximity . . . . .	104
2.7	Textual Outcomes and Social Proximity . . . . .	105
2.8	Spillovers . . . . .	106
2.B.1	Tourism and Rating: November 2015 Paris attacks . . . . .	108
2.B.2	Location-Based Measure: Tourism and Restaurant Ratings by Parisians: Restaurant-Level Analysis . . . . .	109
2.B.3	Location-Based Measure: Tourism and Restaurant Ratings by Parisians: Review-Level Analysis . . . . .	110

2.B.4	Location-Based Measure: Textual Outcomes . . . . .	111
2.B.5	Tourism and Ratings: Language-Based Tourism Aggregated by Different Periods	112
2.B.6	Tourism and Ratings: Different Clustering . . . . .	113
2.C.1	Tourist Access . . . . .	114
2.D.1	Dictionary for Text Analysis . . . . .	115
2.D.2	Summary Statistics for Textual Variables . . . . .	115
2.D.3	Ratings and Textual Variables . . . . .	116
3.1	Summary statistics . . . . .	171
3.2	Main Result - COVID exposure and BLM protest . . . . .	172
3.3	Alternative outcomes . . . . .	173
3.4	Heterogeneity by baseline county characteristic . . . . .	174
3.5	COVID-19 exposure and social media use . . . . .	175
3.6	Effect of Twitter on BLM protest . . . . .	176
3.7	Effect of Twitter on BLM protest . . . . .	177
3.8	Survey data: COVID-19, news consumption and attitudes towards BLM, Blacks and COVID-19 . . . . .	178
3.9	Competing Mechanisms: Broadening versus Scattering of Protest . . . . .	179
3.10	Competeting Mechanisms: Salience, Opportunity Cost, and Agitation . . . . .	180
3.A.1	Reduced form: superspread event on the presence of BLM events . . . . .	186
3.A.2	Instrument robustness - SSE timing and distance . . . . .	187
3.A.3	Instrument robustness - SSE definition and weighting . . . . .	188
3.A.4	Robustness of main results - sample composition and variable definition . . . . .	189
3.A.5	Robustness of main results - estimation method, protest propensity, spatial correlation . . . . .	190
3.B.1	Spring breakers IV: Covid-19 deaths on the presence of BLM events, 2SLS . . . . .	194
3.B.2	Notable Deaths Regression . . . . .	197
3.B.3	Matching on past propensity to protest . . . . .	199
3.C.1	Summary statistics - counties with and without prior BLM event . . . . .	203
3.C.2	Summary statistics for super spreading events by their type . . . . .	204
3.C.3	First stage . . . . .	205
3.C.4	COVID deaths interacted with county characteristics - Counties without BLM events before . . . . .	206
3.C.5	Alternative Mechanisms . . . . .	207
3.C.6	Effect of SXSW users on Twitter presence . . . . .	208
3.C.7	Principal component analysis of online presence . . . . .	209

3.D.1	Description of variables and data sources . . . . .	216
3.D.2	Search terms used in indices of search activity . . . . .	217
3.D.3	Summary statistics - counties with and without geo-localized tweets . . . . .	218
3.D.4	Example Tweets . . . . .	219



# Abstract

This dissertation consists of three chapters that examine from three different perspectives how diversity affects the economy. The first chapter focuses on racial discrimination in rental housing. Does discrimination generate a racial gap in housing rents? Usually, discrimination is covert, which makes it difficult to study. In this paper I concentrate on the unique market of Moscow rental housing, where landlords discriminate overtly: on average, 20 percent of ads from a major rental website include racial requirements. Using model with building-level fixed effects, I document that discrimination generates a racial differential in rents: non-discriminatory apartments have a 4% higher price. I also run a correspondence experiment to explore the relationship between overt and subtle forms of discrimination. I find that both forms coexist in the market. The proportion of overt to covert discrimination is stable across neighbourhoods. The average effect is consistent with a random search model with discrimination. However, heterogeneity analysis contradicts some predictions of the model. I show how adding neighbourhood sorting to the model can explain spatial heterogeneity of a racial rent differential. The second chapter is devoted to the competition between residents and tourists for urban amenities. Using TripAdvisor reviews, we construct panel data on tourism and consumption in Paris. We document that during the pandemic a drop in tourism caused an increase in Parisians' satisfaction with restaurants and other amenities. Among three mechanisms – overcrowding, supply-side changes and aversion towards tourists – we only find support for the aversion mechanism. During the pandemic the word ‘tourist’ became less frequent in reviews, while other words relating to food quality, price and overcrowding stay on the same level. The improvement in ratings was stronger in restaurants popular among tourists from countries with a weaker social connection to France measured with Facebook connectedness index. The third chapter explores how contemporary social movements can expand their base. Prompted by the viral video footage of George Floyd’s murder, the Black Lives Matter (BLM) movement gained unprecedented momentum and scope in the spring of 2020. Using Super Spreader Events as a source of plausibly exogenous variation at the county-level, we find that pandemic exposure led to an increase in the likelihood of observing online and offline BLM protests. This effect is most

pronounced in whiter, more affluent and suburban counties. We show that this effect is driven by higher social media take-up among non-traditional users. Specifically, we find that a one standard deviation increase in pandemic exposure led to a doubling of new Twitter accounts in counties with no BLM protest history. Our results suggest that the pandemic acted as a demand shock to social media among non-traditional users, mobilizing new segments of society to join the movement for the first time. We find supporting evidence for this mechanism using individual-level survey data and rule out competing channels, such as pandemic induced salience of racial inequality, lower opportunity cost of protesting or higher overall agitation and propensity to protest.

# Introduction

This thesis combines chapters on three diverse subjects with one thing in common: the subject of diversity. I focus on a particular type of diversity: in race, identity, attitudes and beliefs.

Since Becker (1957) race and identity have become a legit part of economic reasoning. In his work on labor discrimination, Becker considered a situation where workers of two races coexist in the market and some employers have a “distaste” for workers of one race. Becker’s seminal work can be seen as a part of a broader question: “What happens when agents of different races or identities operate in the same economy?” In the three chapters of this thesis I consider three different scenarios that can happen.

The first scenario that has already been mentioned is discrimination – that is, exclusion from the market. The second scenario is a conflict – when neither group is able to exclude the other from the market, but the attitude of the groups is still reflected in their behavior. An example of such a case would be consumer segregation (Davis et al. (2019)). Finally, inclusion is also possible when groups join a coalition, or when cultural transmission occurs. The chapters presented here should be seen as examples, not as generalizations of these scenarios. In the introduction, I will focus on the literature and the issues that surround all three cases.

A key example of exclusion is racial discrimination. A vast economic literature has been developed examining discrimination in various markets and configurations: labor, housing, consumption, credit, schooling, and others<sup>1</sup>.

Two types of discrimination have become the epitome of the theoretical literature: taste-based discrimination and statistical discrimination. Taste-based discrimination is driven by agent preferences (Becker (1957); Arrow (1972); Black (1995)). Statistical discrimination is different. It does not suggest that agents are prejudiced. On the contrary, agents are rational and use the identity of the counterparty as a proxy for its “performance” in a situation of information asymmetry. If discriminated group has a lower performance on average, then discrimination arises as a rational choice. Classical model of statistical discrimination was

---

<sup>1</sup>For extensive reviews of the literature see Lang and Lehmann (2012); Bertrand and Duflo (2017)

proposed by Phelps (1972). More complex setting of this model, as introduced by Tirole (1996), involves a prior stage in which the minority agent can choose how much he or she wants to invest in building the skill that determines future performance. Then the “bad reputation” of the group takes away the agent’s incentive to invest in the skill. It is important to note that both forms of discrimination – statistical or taste-based – meet the UN definition of discrimination and are illegal in many countries<sup>2</sup>.

The frameworks of taste-based and statistical discrimination do not exhaust or represent the multitude of potential mechanisms and institutional settings through which discrimination can occur. Small and Pager (2020) emphasize the importance of other frameworks and show how they can complement and extend traditional approaches. They mention several directions. Some of them have already appeared in the economic literature.

First, people can discriminate without realizing it, a phenomenon that has been called “implicit discrimination” in Bertrand et al. (2005). Second, discrimination can be reinforced through organizational structure even without the intent of individual members. Third, past discrimination (sometimes recorded in law) can have a strong influence on contemporary inequality. For example, Aaronson et al. (2021) show that 1930s “redlining” had long-run socioeconomic effect. Fourth, minor forms of discriminatory behavior can have important consequences. For example, a minority worker may be hired but treated differently in the workplace (he or she has a higher workload, is more closely monitored). Finally, all together, this will also require consideration of a broader set of consequences, such as experienced discrimination and emotional strain.

From the perspective of the empirical literature on discrimination, the main challenge is that discrimination is difficult to observe. In many communities, discrimination is illegal and socially unacceptable. Therefore, in order to study discrimination, we must first learn to detect it. However, this has not always been the case. For example, in the United States before the Civil Rights Act of 1964, racial discrimination was overt and widespread. Job advertisements published in the New York Times regularly contained explicit racial requirements (Darity and Mason (1998)). Housing complexes publicly informed tenants about the “no blacks” policy. But importantly, discrimination in those days was not studied with the statistical tools available today.

One way to identify discrimination is to compare the economic outcomes of different racial groups. This approach has generated a literature that estimates racial gaps using regression decomposition. Racial gaps in the housing market are well documented, with most studies focusing on the United States: Ihlanfeldt and Mayock (2009); Bayer et al. (2017); Yinger (1997); Early et al. (2019). More specifically, for the U.S. rental housing market, Early et al.

---

<sup>2</sup>For the data on the anti-discriminatory laws across countries see Mipex.

(2019) shows that blacks pay 0.6 to 2.4 % percent more than whites for identical housing in identical neighborhoods.

It is debatable, however, whether these results hold when all the necessary controls are included. Neal and Johnson (1996) show that the racial wage gap shrinks or even disappears when a variable measuring a job seeker's cognitive skill is included in the equation<sup>3</sup>. This has led researchers to question: perhaps the gaps previously found in studies are not the result of discrimination, but reflect differences between groups before they enter the market. Following this logic, pre-market differences in human capital can explain racial disparities in wages, and differences in negotiating skills can explain disparities in housing. Relying on regression decomposition, it is difficult to say to what extent racial differences are caused by discrimination. Studies that can address this question in an empirically rigorous way are rare (Fryer et al. (2013)).

Since the beginning of 2000, another strand of the literature has emerged. In order to reveal the existence of differential treatment, researchers began to conduct correspondence experiments. In their seminal work, Bertrand and Mullainathan (2004) sent out pairs of fictitious resumes with Black- or White-sounding names to employers in Boston and Chicago, randomizing other characteristics. This approach allowed them to identify differential treatment: candidates with Black-sounding names were less likely to receive a callback from a potential employer. Correspondence experiments have attracted the close attention of researchers. Baert (2018) discusses its effectiveness and shortcomings. Correspondence experiments have revealed discrimination in many markets, eliminating some of the blind spots characteristic of previous studies of racial discrimination.

At the same time, correspondence experiments do not clearly explore the relationship between discrimination and racial gaps. In the first chapter I identify this link drawing on unique context of Moscow rental housing market, where landlords discriminate overtly: around 20% of Moscow landlords from online marketplace *Cian* include racial requirements to their rental ads. I am going to briefly summarize this chapter further in the introduction.

The second chapter illustrates another common scenario: a conflict between consumers of different groups who meet in the same economic environment without supply-side discrimination.

In this chapter, which is based on joint work Stefan Pauly, we look at intra-city competition between tourists and residents for urban amenities.

As Faber and Gaubert (2019) noted, “tourism involves the export of otherwise non-traded local services by temporarily moving consumers across space, rather than shipping

---

<sup>3</sup>Neal and Johnson (1996) measure skills with Armed Forces Qualification Test (AFQT), a test used to determine qualification for enlistment in the United States Armed Forces

goods”. Based on insights from the trade literature, Faber and Gaubert (2019) conduct a structural analysis of the economic benefits of tourism. Lanzara and Minerva (2019) look at the interactions between tourism and amenities, and consider the welfare consequences. Dissatisfaction with tourism has rarely been explored in the economic literature. Rare exception is Takahashi (2019) who examines the negative effects of tourism from a theoretical perspective.

There are several factors to consider: tourists as imported consumers may have preferences and attitudes that differ from those of residents, they may put additional strain on local infrastructure and services, and finally, residents may have negative attitudes toward tourists. All these aspects are discussed in the second chapter, and a brief summary is presented later in the introduction.

The literature on urban economics has other than tourism examples of conflict between different groups. In many cities different racial groups co-exist, interact and consume in the same environment. Mazzolari and Neumark (2012) observe that diversity among residents correlated with diversity in consumption. This is also consistent with Schiff (2015) evidence about the attractiveness of density in the city. In parallel, it is known that there can be segregation in consumption in the city. Davis et al. (2019) examines segregation in consumption in New York City, adding to the traditional notion of residential segregation in the literature.

The third chapter, which is co-authored with with Annalí Casanueva Artís, Sulin Sardoschau and Kritika Saxena, sheds light on another potential scenario: inclusion. Linked to the political economy of protest, this chapter highlights a crucial aspect of diversity – the ability of different groups to form a coalition to bring political change.

This chapter also stands out from the other two because it relates to the literature examining the role of information and media in the economy. Previous work has shown that social media can solve the collective action and coordination problem for individuals already sympathetic to a political cause: Enikolopov et al. (2018); Manacorda and Tesei (2020). In contrast, we focus on the role of social media as a tool that can expand coalition and mobilize new protesters.

Studies that examine the impact of the Internet and new media tend to use a supply-side shift in the early stages of Internet or social media adaptation: Guriev et al. (2019); Müller and Schwarz (2021); Enikolopov et al. (2018); Manacorda and Tesei (2020). To the best of our knowledge, we are the first to investigate the role of social media in broadening political coalitions through persuasion, rather than mobilizing individuals that are already sympathetic to the movement’s grievances.

Another theme that unites these chapters is that of the digital economy. All chapters

benefit from new data coming from digital platforms. Consumption, housing, transportation have moved online (Goldfarb and Tucker (2019)). Political and socially relevant information is spreading through social media. This creates a digital footprint that can be used by researchers. Economists of the past paid less attention to issues such as inequality, not because these issues were not of social interest. On the contrary, they were always of prime interest, but the data were difficult to obtain.

In the following parts of this introduction I will summarize the main results of each of the chapters of the thesis.

## Chapter 1: Consider the Slavs: Overt Discrimination and Racial Disparities in Rental Housing

Today's discrimination is mostly subtle. This makes its impact hard to measure. This chapter is trying to overcome this challenge drawing on the unique context of Moscow's rental housing market, where landlords discriminate overtly. They include racial requirements to ads, using phrases like "offer is only for slavic tenants", where slavic denotes ethnically Russian tenants or tenants of ethnically Russian appearance.

More specifically, I investigate how discrimination in the market for rental housing can generate a racial rent differential.

I collect new data on rental ads from the major Russian online real estate marketplace cian.ru. The dataset includes all available ads over a period of around six months. I categorise ads by presence of racial requirements and combine it with other observable characteristics of apartments and neighborhoods. Around 20 percent of ads include racial requirements. This setting thus allows me to estimate the effect of discrimination on the racial rent differential. To causally identify this effect, I include building-level fixed effects to the model to absorb any geographic and building-level characteristics.

I find that discrimination generates a significant and sizeable racial rent differential: comparing apartments in the same building with identical observable characteristics, nondiscriminatory apartments have a 4 % higher price. This paper also examines the relationship between overt and subtle forms of discrimination. I conduct classic correspondence experiments, sending messages with non-Russian and Russian-sounding names to a random subset of online ads. This experiment allows me to relate the results obtained from the observational study to the existing body of evidence from the experimental literature. I find that both subtle and overt forms of discrimination coexist on the rental housing market in Moscow. Their relative prevalence is constant across neighbourhoods.

Finally, I borrow a theoretical framework from the literature on labor search with discrimination Black (1995) and apply it to the context of rental housing in Moscow. I demonstrate that the search-based model can explain the existence of the racial rent differential. The intuition is the following: when the search is costly and minorities have higher chances of getting rejected, they are more likely than the majority to accept an unfavorable offer. Then non-discriminating landlords who anticipate it will raise the rent price in equilibrium.

However, the standard search-based model cannot explain the results of the heterogeneity analysis. I find that in neighborhoods (and buildings) with a higher share of discriminating apartments the racial rent differential is lower. At first glance, this contradicts the implication of the model, which says that with a larger proportion of discriminating apartments the gap should expand. However, this view assumes that neighborhoods are different and isolated markets, while in fact potential tenants sort (but not necessarily strongly segregate) between neighborhoods. I include a neighborhood choice stage in the search-based model to explain the results obtained in the heterogeneity analysis.

## Chapter 2: Urban Amenities and Tourism: Evidence from Tripadvisor

*This chapter is co-authored with Stefan Pauly.*

In this paper we estimate the effect of tourism on residents' satisfaction with restaurants and other urban amenities. We use data on restaurant reviews from Tripadvisor – the platform that aggregates user-generated content on restaurant and other travel experiences. We construct unique panel data on consumption and amenities in the city. This data allows us to achieve multiple goals at the same time.

First, we use it to produce a highly granular measure of tourism. The share of non-French among all reviews serves as a close proxy of tourists' presence, which we validate using several other measures. The benefit of this measure is that it can be defined on a very granular level, the restaurant itself. In addition, while many studies focus on the location where tourists stay overnight to study the impact, the measure used here allows to study the location of where tourists consume.

Second, the review data and the ratings given by locals can be used as an indicator of locals' satisfaction with restaurant experience. More generally, it serves as a measure of satisfaction with urban amenities, which varies across space and time. The literature shows that this indicator is meaningful: For example, Kuang (2017) finds that restaurant ratings are highly correlated with real estate prices.

We match restaurant data with another source of information on residents' quality of life: number of complaints on the crowd-sourced platform DansMaRue. The platform is provided by the city hall of Paris. Users can report any problem related to public space (abandoned waste, tags, wild posting, etc.) through the mobile application or the web-site. Then the city administration analyses the reports and try to solve the problems. We treat this disamenity measure as another outcome relevant to our study.

We first document two stylized facts. First, more touristic restaurants receive lower ratings by locals in the cross-section, suggesting a potential disamenity stemming from tourist demand. Second, touristic neighborhoods have a lower variety of amenities which may indicate that tourists value variety less than locals do. Using the pandemic as a source of exogenous variation in international tourist arrivals, we find that the drop in tourism caused an increase in residents' satisfaction with urban amenities, both in terms of restaurant ratings and a decreased number of complaints on DansMaRue. In particular, the average restaurant increases its rating by close to 10 % of a standard deviation in the absence of tourists and the number of complaints in the direct vicinity of the average restaurant decreases by at least 8 %.

Importantly, our effect is not unique to the lockdown-induced tourism decline. We find similar evidence when using the terrorist attacks that took place in November 2015. Our results are also robust to using measures of tourism that are based on the self-declared location of users rather than language.

Next, we consider three potential mechanisms driving our findings: overcrowding, supply side change and residents' aversion towards tourism. Our analysis only finds support for the aversion mechanism. First, we find that the number of reviews explicitly mentioning tourism (which are often negative) declines. Second, relying on a proxy of social connectedness between countries derived from Facebook data, we find that restaurants with a clientele that has little connections to France sees a larger increase in its rating post-lockdown. This suggests that Parisians are less bothered by tourists from countries with which they have strong social ties.

## **Chapter 3: Going Viral in a Pandemic: Social Media and Allyship in the Black Lives Matter Movement**

*This chapter is co-authored with Annalí Casanueva Artís, Sulin Sardoschau and Kritika Saxena.*

What led to the broadening of the Black Lives Matter movement's coalition during the

pandemic? We approach this question in two parts. First, we establish a causal link between exposure to COVID-19 and protest participation at the county level, using Super Spreader Events as a source of exogenous variation. We show that exposure to COVID-19 is associated with an increase in protest behavior but only among those counties that have never protested for a BLM-related cause before.

Second, we develop a novel index of social media penetration at the county level to show that this effect is driven by higher social media take-up during the pandemic but before the protest trigger. While we cannot fully rule out that other mechanisms were at play, we show evidence that alternative explanations such as *i*) a pandemic-induced rise in the salience of racial inequality, *ii*) lower opportunity costs of protesting, *iii*) higher overall propensity to protest and *iv*) a scattering rather than a broadening protest are not driving our results.

Our identification is based on a small window between the end of March and mid April of 2020 during which COVID-19 was prevalent enough but lock-down stringency lax enough to allow for so-called Super Spreader Events (SSE) to occur. These events are characterized by the presence of one highly infectious individual (a super-spreader) and took place mainly at birthday parties, nursing homes or prisons. We exploit cross-sectional variation in the number of SSEs within a 50 kilometer radius from the county border but not within the county 6 weeks prior to the murder of George Floyd to construct our instrument for exposure to COVID-19 at the county level. We include state fixed effects and a vast set of county level controls, most notably the number of historical BLM events between 2014 and 2019, as well as socio-demographic variables and proxies for political leaning and social capital.

We find robust evidence that exposure to COVID-19 increased BLM protest. We estimate that a one standard deviation increase in the number of COVID-19 related deaths in a county at the time of George Floyd's murder (approximately 25 deaths per 100K inhabitants), increases the likelihood of a BLM event occurring in the three weeks following the murder by 5%. Our baseline result is entirely driven by counties with no prior BLM protests and the effect doubles in size and is more precisely estimated for this sub-sample.

In addition, we propose three alternative identification strategies and show that our results replicate. First, using large scale mobile phone mobility data by *SafeGraph*, we instrument pandemic exposure with tourist flows to one of the largest SSEs in the US - Florida spring break in March 2020. Second, we employ a difference in differences approach, for which we scrape information on all similar BLM protest triggers since 2014 to estimate the differential response to a protest trigger before and after the pandemic. Third, we use a LASSO-based matching approach, comparing counties with similar pre-pandemic protest probabilities.

In a next step, we investigate various sources of heterogeneity and show that - in line

with the idea of a broadening movement - our baseline results are driven by whiter, more affluent and sub-urban counties.

In the second part of the paper, we investigate whether the uptake in social media can account for the pandemic-induced broadening of the BLM movement. We start by repeating the above analysis, this time using a novel index of social media penetration as our main outcome variable. We find that the pandemic has a positive and significant effect on our social media index and that this is entirely driven by the sub-sample of counties that have never protested before. For instance, we show that a one standard deviation increase in pandemic exposure led to a doubling of twitter accounts among counties with no prior BLM event, without affecting counties that traditionally protest.

In a next step, we zoom in on the role of twitter in mobilizing BLM protesters. First, we interact baseline twitter penetration (before the pandemic) with exposure to COVID-19. We address the concern that our results could capture underlying factors that drive both Twitter penetration and protest participation, replicating the SXSW instrument for baseline Twitter penetration used by Müller and Schwarz (2020). We show that counties with higher baseline twitter penetration react more to pandemic exposure. Additionally, we interact pandemic exposure with contemporaneous twitter penetration and find that the effect of COVID-19 on protest is entirely driven by counties with higher twitter take-up during the pandemic.

In the last part of our paper, we look at competing mechanisms. Naturally, the pandemic has affected a number of important dimensions that are not limited to higher social media take-up. First, we consider the possibility that our results are driven by a scattering rather than a broadening of BLM protest. More specifically, we verify that the effect is not driven by a substitution away from some locations to others. Second, the pandemic may have increased the overall salience of racial inequality *before* the murder of George Floyd. We test this by interacting COVID-19 with a proxy for disproportional death burden on Blacks and the number of BLM-related search terms on Google before the protest trigger. Third, we investigate whether the pandemic has decreased the opportunity cost of protesting. We interact COVID-19 with the unemployment rate at the county level and stringency at the state level. If individuals choose to protest in lieu of going to work or engage in social activities, we should see a larger effect in counties with higher unemployment rates or stricter stringency measures. Third, we look at the effect of COVID-19 on other protests. If the pandemic increased overall agitation and propensity to protest, then we would expect this to also hold for other causes beyond BLM. We show that these channels are unlikely to drive our results.

# References

- Aaronson, D., Faber, J., Hartley, D., Mazumder, B., and Sharkey, P. (2021). The long-run effects of the 1930s home “redlining” maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, 86:103622.
- Arrow, K. J. (1972). Models of job discrimination. *Racial discrimination in economic life*, 83.
- Baert, S. (2018). Hiring discrimination: an overview of (almost) all correspondence experiments since 2005. In *Audit studies: Behind the scenes with theory, method, and nuance*, pages 63–77. Springer.
- Bayer, P., Casey, M., Ferreira, F., and McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, 102:91–105.
- Becker, G. S. (1957). *The economics of discrimination: an economic view of racial discrimination*. University of Chicago.
- Bertrand, M., Chugh, D., and Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, 95(2):94–98.
- Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 1:309–393.
- Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.
- Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, 13(2):309–334.
- Darity, W. A. and Mason, P. L. (1998). Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives*, 12(2):63–90.

- Davis, D. R., Dingel, J. I., Monras, J., and Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, 127(4):000–000.
- Early, D. W., Carrillo, P. E., and Olsen, E. O. (2019). Racial rent differences in us housing markets: Evidence from the housing voucher program. *Journal of Regional Science*, 59(4):669–700.
- Enikolopov, R., Petrova, M., and Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics*, 10(1):150–74.
- Faber, B. and Gaubert, C. (2019). Tourism and economic development: Evidence from mexico's coastline. *American Economic Review*, 109(6):2245–93.
- Fryer, R. G., Pager, D., and Spenkuch, J. L. (2013). Racial disparities in job finding and offered wages. *The Journal of Law and Economics*, 56(3):633–689.
- Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.
- Guriev, S., Melnikov, N., and Zhuravskaya, E. (2019). 3g internet and confidence in government. Available at SSRN 3456747.
- Ihlanfeldt, K. and Mayock, T. (2009). Price discrimination in the housing market. *Journal of Urban Economics*, 66(2):125–140.
- Kuang, C. (2017). Does quality matter in local consumption amenities? an empirical investigation with yelp. *Journal of Urban Economics*, 100:1–18.
- Lang, K. and Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4):959–1006.
- Lanzara, G. and Minerva, G. A. (2019). Tourism, amenities, and welfare in an urban setting. *Journal of Regional Science*, 59(3):452–479.
- Manacorda, M. and Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567.
- Mazzolari, F. and Neumark, D. (2012). Immigration and product diversity. *Journal of Population Economics*, 25(3):1107–1137.
- Müller, K. and Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4):2131–2167.

- Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, 104(5):869–895.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661.
- Schiff, N. (2015). Cities and product variety: evidence from restaurants. *Journal of Economic Geography*, 15(6):1085–1123.
- Small, M. L. and Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, 34(2):49–67.
- Takahashi, T. (2019). Battles between residents and tourists: On the welfare effects of growing tourism.
- Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.
- Yinger, J. (1997). Cash in your face: The cost of racial and ethnic discrimination in housing. *Journal of Urban Economics*, 42(3):339–365.

# Chapter 1

## Consider the Slavs: Overt Discrimination and Racial Disparities in Rental Housing

### Abstract

Does discrimination generate a racial gap in housing rents? Usually, discrimination is covert, which makes it difficult to study. In this paper I concentrate on the unique market of Moscow rental housing, where landlords discriminate overtly: on average, 20 percent of ads from a major rental website include racial requirements. Using model with building-level fixed effects, I document that discrimination generates a racial differential in rents: non-discriminatory apartments have a 4% higher price. I also run a correspondence experiment to explore the relationship between overt and subtle forms of discrimination. I find that both forms coexist in the market. The proportion of overt to covert discrimination is stable across neighbourhoods. The average effect is consistent with a random search model with discrimination. However, heterogeneity analysis contradicts some predictions of the model. I show how adding neighbourhood sorting to the model can explain spatial heterogeneity of a racial rent differential.

# 1. Introduction

Racial discrimination is usually hidden from public view. Aiming to reveal the very fact of discrimination, economists mainly resort to one of two approaches. The first type is observational studies that estimate racial gaps in economic outcomes like wages and rents. The second type is correspondence experiments that uncover the differential treatment. As a result, both racial gaps and discrimination are well-documented in many markets and countries<sup>1</sup>. However, there are few pieces of evidence on the link between the two, so it is still under discussion: to what extent does discrimination generate racial gaps?

Economists have repeatedly questioned the contribution of discrimination to racial gaps, pointing out to the premarket factors (education, social capital, culture) as the main drivers (Neal and Johnson (1996), Heckman (1998)). At the same time, the systematic evidence on this link is hard to obtain mainly due to the private nature of discrimination. The rare exception is Fryer et al. (2013) who show that in the US labor market at least one-third of the black-white wage gap can be attributed to discrimination.

While it is rare nowadays, overt discrimination has been widespread in the past. Writing on the United States before the Civil Right Act of 1964, Arrow (1998) noted:

The presence of racial discrimination throughout American society was, to use the words of Samuel Johnson, a fact *too evident for detection and too gross for aggravation*. To establish the existence of discrimination, estimating wage equations would have been beside the point. Of course, society and scholars would want to know the quantitative implications of discrimination for income as well as other indices of well-being. But the fact of discrimination would not have needed testing.

Today's discrimination is mostly subtle. This makes its impact hard to measure. This paper is trying to overcome this challenge drawing on the unique context of Moscow's rental housing market, where landlords discriminate overtly. They include racial requirements to ads, using phrases like "offer is only for slavic tenants", where *slavic* denotes ethnically Russian tenants or tenants of ethnically Russian appearance.

More specifically, I investigate how discrimination in the market for rental housing can generate a racial rent differential.

I collect new data on rental ads from the major Russian online real estate marketplace *cian.ru*. The dataset includes all available ads over a period of around six months. I categor-

---

<sup>1</sup>See Bertrand and Duflo (2017) for an extensive review of empirical studies on discrimination. It also discusses the methodological difference between regression decompositions and field experiments, as well as other original lines of research.

ise ads by presence of racial requirements and combine it with other observable characteristics of apartments and neighborhoods. Around 20 percent of ads include racial requirements. This setting thus allows me to estimate the effect of discrimination on the racial rent differential. To causally identify this effect, I include building-level fixed effects to the model to absorb any geographic and building-level characteristics.

I find that discrimination generates a significant and sizeable racial rent differential: comparing apartments in the same building with identical observable characteristics, non-discriminatory apartments have a 4% higher price.

This paper also examines the relationship between overt and subtle forms of discrimination. I conduct classic correspondence experiments, sending messages with *non-Russian* and *Russian-sounding* names to a random subset of online ads. This experiment allows me to relate the results obtained from the observational study to the existing body of evidence from the experimental literature. I find that both subtle and overt forms of discrimination coexist on the rental housing market in Moscow. Their relative prevalence is constant across neighbourhoods.

Finally, I borrow a theoretical framework from the literature on labor search with discrimination Black (1995) and apply it to the context of rental housing in Moscow. I demonstrate that the search-based model can explain the existence of the racial rent differential. The intuition is the following: when the search is costly and minorities have higher chances of getting rejected, they are more likely than the majority to accept an unfavorable offer. Then non-discriminating landlords who anticipate it will raise the rent price in equilibrium.

However, the standard search-based model cannot explain the results of the heterogeneity analysis. I find that in neighborhoods (and buildings) with a higher share of discriminating apartments the racial rent differential is lower. At first glance, this contradicts the implication of the model, which says that with a larger proportion of discriminating apartments the gap should expand. However, this view assumes that neighborhoods are different and isolated markets, while in fact potential tenants sort (but not necessarily strongly segregate) between neighborhoods. I include a neighborhood choice stage in the search-based model to explain the results obtained in the heterogeneity analysis.

Racial gaps in the housing market are well documented, with most studies focusing on the United States: Ihlanfeldt and Mayock (2009), Bayer et al. (2017), Yinger (1997); Early et al. (2019). More specifically, for the US rental market Early et al. (2019) show that Blacks pay 0.6 - 2.4 % higher rent price than Whites for identical housing in identical neighborhoods. From the landlord's point of view these results suggest lost profits. There are few papers that investigate the trade-off between decision to discriminate and lost profits. Hedegaard and Tyran (2014) conduct field experiments to measure the sensitivity of discrimination to

changes in opportunity cost. Finally, in a simultaneous and independent research project Veterinarov and Ivanov (2018) perform similar analysis using data on overt discrimination from Russian online marketplace and find the set of similar empirical results. In contrast to Veterinarov and Ivanov (2018) my paper proposes different theoretical mechanism and introduce the analysis of interaction between overt and subtle types of discrimination. It is crucial to note that reproduction of the same observational study using different empirical strategies increases the reliability of the existence of the racial rent differential.

There are numerous studies that document racial discrimination on the housing market with the help of correspondence and audit experiments: Yinger (1986), Carpusor and Loges (2006), Hanson and Hawley (2011) in the US, Ahmed and Hammarstedt (2008) in Sweden, Acolin et al. (2016) in France. When it comes to the labor market, explicit racial requirements are rather rare in Russia: Bessudnov and Shcherbak (2018) conduct a correspondence experiment and document substantial and statistically significant differences in callbacks between majorities and minorities.

This study contributes to an emerging body of literature exploiting user-generated content and text analysis. As an example, Stephens-Davidowitz (2014) uses Google search data as a proxy for racial animus. Closest to my paper is Kuhn and Shen (2012) who study overt gender discrimination in Chinese online job listings, however, they do not estimate the effect on prices, but instead try to determine the causes of discrimination. A detailed review of the methods used for text analysis can be found in Gentzkow et al. (2017).

The link between overt and subtle forms of discrimination is a recurring theme in the sociological literature Small and Pager (2020), Pager (2007). The subtle form has several notable features. First, the discriminating person can either be aware or unaware that he or she is discriminating. “*Unconscious*” discrimination was conceptualised by psychologists and economists as an *implicit discrimination* Bertrand et al. (2005). Second, the analysis of subtle discrimination blurs the line between statistical and taste-based discrimination: the qualitative studies show that employers narrate their prejudiced attitudes using “statistical” arguments, but fail to update their beliefs when facing contradicting information Pager and Karafin (2009). This also corresponds to the observation that locals in many countries highly overestimate the number of immigrants and perceive imprecisely their characteristics Alesina et al. (2018).

Overt discrimination is often regarded as a pure manifestation of racial animus. At the same time, anecdotal evidence suggests, that overt discrimination observed in the rental housing in Moscow has a lot in common with typical subtle discrimination, where landlords do not consider their behavior as discriminating.

The theoretical section of this paper is related to literature that implements taste-based

discrimination to search models. Since the interest of this paper leans towards the impact of discrimination and not its causes, it is reasonable to concentrate on a competitive taste-based framework. Thereby, we leave aside the question of the rationality of landlords' beliefs and assume that landlords have an exogenous distaste of minorities.

A standard Beckerian perfect competition framework (Arrow (1972), Becker (2010)) does not explain the existence of the cost of discrimination. Such an effect would persist if and only if two markets would fully separate between the majorities and the minorities. It implies that the majority rent only discriminating apartments, while discriminating apartments make up only 20 percentage of the rental market. In a more realistic scenario perfect competition leads to a unique price.

Racial discrimination on the labor market has been studied more extensively than discrimination on the housing market<sup>2</sup>. Following insights from the labor literature, I adapt the search model proposed in Black (1995) to the context of rental housing in Moscow. In this model discriminating landlords refuse to accept minorities at any price, which makes search more costly for minorities. Therefore, landlords who do not discriminate increase their rent, since minority tenants with increased search costs tend to accept more expensive offers.

Other important models of random search with discrimination are proposed in Bowlus and Eckstein (2002) and Rosén (1997). Directed search with discrimination is presented in Lang et al. (2005). When it comes to the rental housing market, search models with discrimination are less common. A notable exclusion is an early model proposed by Courant (1978), which has a lot of similarities with Black (1995). Another original mechanism of discrimination during the search, which is called "neighbour discrimination", was proposed by Combes et al. (2018). It captures the situation when landlords who own more than one apartment in a building can discriminate minorities even if they do not have a distaste for them. When a landlord rents an apartment to minority tenants, he or she reduces the attractiveness of his or her other property, because other potential tenants on the market are prejudiced against minorities. There are also several papers that study search and matching on the housing market regardless of the discrimination context: Albrecht et al. (2016), Carrillo (2012), Ngai and Tenreyro (2014).

The paper is organized as follows. Section 2 describes the data and background of the online housing marketplace. Section 3 presents the major empirical findings on racial rent differentials and the results of a correspondence experiment. Section 4 examines a theoretical framework that sheds light on the mechanism of existence of the racial rent differential and tries to explain the heterogeneity of this effect.

---

<sup>2</sup>See Lang and Lehmann (2012) for an extensive literature review on the topic of racial discrimination on the labor market

## 2. Background and Data

Russia is a multinational state: 19% of the population are not ethnic Russians (Census, 2010). There is also a large population of immigrants. According to UN data, around 11 millions immigrants resided in Russia in 2019 (8% of the total population), which made Russia the second country in the world by the population of immigrants after the US. It is important to note that the overwhelming majority of immigrants residing in Russia are citizens of the former Soviet Union or their descendants. Among the largest “non-slavic” ethnic groups residing in Moscow, there are Tatars, Bashkir, Chuvashs, Chechens, Armenians, Avars, Mordvins, Kazakhs, Azerbaijanis, Uzbeks, Kyrgyz, Tadzhiks to name a few.

Xenophobic attitudes are rather common in Russia. According to Russian independent polling organisation Levada Center, 63 percent of Moscow respondents are permissive about discriminating rental advertisements. Every second respondent approve the political slogan “Rossiya dlya Russkikh”, which can be translated as “Russia should be for ethnic Russians”. These attitudes have historical roots. The Soviet Union pursued a complex and controversial ethnic policy, blending anti-discriminatory and discriminatory interventions, such as: vigorous anti-racism propaganda, harsh control of the population mobility (restrictions on mobility, or, on the contrary, waves of forced migration) and promotion of local languages and cultures Martin et al. (2001). Dissolution of the Soviet Union stimulated nationalist movements and ethnic violence both among Russian and non-Russian populations.

Modern Russia pursues an ambivalent anti-discrimination policy. On the one hand, the number of those convicted of hate speech has increased from 149 to 604 from 2011 to 2017<sup>3</sup>. On the other hand, the judicial practice is poor when it comes to actual discrimination in the labor and housing markets<sup>4</sup>. In particular, a discriminating landlord does not pay any fees and has no other constraints for including racial preferences in apartments ads.

While people of many ethnicities reside in Moscow, there is no evidence of apparent racial segregation comparable to the one found in American and European cities Vendina (2002); Vendina et al. (2019). The census also does not show signs of strong segregation (Figure 1.4(a)). At the same time, the share of non-Russian residents is higher in the city center – the more prestigious part of Moscow, where overt discrimination is rare. The lack of strong segregation in Moscow is probably a heritage of the strict housing regulation imposed in the Soviet Union.

---

<sup>3</sup>According to the Judicial Department at the Supreme Court of the Russian Federation. The statistics was published by newspaper Kommersant

<sup>4</sup>For the legal practices on discrimination in Russia see journalistic investigation by online newspaper Meduza

The empirical part of this paper benefits from the structure of the Russian housing stock: it allows me to introduce building-level fixed effects to the model. The state of modern mass housing in Russia is largely determined by Soviet post-war housing policy. Two crucial features of this policy should be noted: the housing stock was state-owned and dwelling allocation was state controlled. Since the 1970s, urban development has been focused on 9 and 16-storey buildings. The new private wave of development inherits the Soviet housing approach of multi-story community blocks. The data used in this paper shows: the median building is 12-storey with around 200 apartments. In addition, apartments in the same building are usually homogeneous in quality.

## 2.1. *Cian* data

Every day the web-site *cian.ru* posts around two thousand rental offers, around two thousand offers disappear from the site, and around 28 thousand offers remain available. According to user statistics *cian.ru* is the biggest online platform to search for long-term rentals in Russia. Over the last decade the property market has almost entirely gone online. Therefore, data collected from *cian.ru* is the most feasible and complete representation of rental supply in Moscow.

Potential tenants get access to the platform through the search interface, where they can specify desired characteristics of the apartment: expected rent price, location, number of rooms, surface area, layout. Then users can browse the list of search results. If a user is interested in the offer, he or she can respond through an online form or call the given phone number.

Each ad consists of the basic apartment's characteristics, a text description and a set of images. Descriptive statistics of ads are reported in Panel A of Table 1.1. For most apartments, the exact address is indicated. I geocoded addresses, calculated distances between buildings and the city center, distances between buildings and closest metro stations. Location data also allows to group apartments at the building level, district level (12 *okrugs*, according to Moscow administrative division) and subdistrict level (146 *raions* and settlements). Descriptive statistics of buildings, districts and subdistricts characteristics are presented in Panels B, C and D of Table 1.1.

The main observation period lasted from May 27 to November 11, 2018. There is also a stand alone one-day snapshot, which was collected on April 2, 2017. Data were scraped from the site every midnight Moscow time, when users are supposedly least active. There were few days when it was not possible to collect data – I exclude these days from analysis. The final dataset consists of 117 daily snapshots. Figure 1.2 shows that the number of posted ads

is a seasonal variable. It varies between 22 thousands and 35 thousands, increases in summer and decreases in autumn. This fluctuation can be explained with seasonality of demand.

Figure 1.1 (a) reports the map of Moscow, where each dot corresponds to an observed building and the color indicates the share of discriminating apartments in each building. It is clear that discrimination is uneven throughout Moscow. The city center and southwest area are associated with low levels of discrimination, whereas outskirts tend to be most discriminating. The map of discrimination aggregated by subdistricts is presented in the Figure 1.1 (b). It can be seen that in some subdistricts the share of discriminating apartments can reach as much as 54 %. The spatial pattern of discrimination is highly stable (see Figure 1.3).

The resulting panel consists of 213 thousands ads that appeared on the site during the observation period. Using this data one can see how rent prices have been changing during the observation period. Two groups of observations stand out: first, around 80 percent of offers that have not changed rent price during the whole period, and, second, the group of offers that decreased the rent price. This pattern motivates the use of the latest rent prices in estimation of the cost of discrimination — these rent prices are closer to equilibrium prices.

The supply side is represented by two types of actors: landlords and agents. They both can directly access the platform. Agents are licensed specialists hired by landlords who take on the job of finding a reliable tenant at an optimal rent price. Anecdotal evidence suggests that, when it comes to ethnic requirements, agents transmit preference of landlords with whom they work. Both agents and landlords leave their phone numbers in rental ads, but it is not always possible to distinguish whether the counterparty is the landlord or the agent.

Using accompanying ads' texts, I was able to identify the presence of racial discrimination. For the baseline analysis, I resorted to a dictionary approach<sup>5</sup>. The algorithm consists of several steps: first, I calculate frequencies of all unigrams, bigrams and trigrams, then examine them manually to reveal the ones related to ethnicity of tenant and, finally, flagged ads containing these n-grams. Discrimination in ads is manifested in a highly uniform way: most of discriminating landlords use the phrase "Slavs only". The rest of discriminating landlords use words with roots: *slav-*, *russ-*, *caucas-*, *asia-*. For the key phrases, few instances of reversed use were detected and excluded (for example, preceding "not only", or following "are allowed"). There are also specific inclusive phrases in the data, such as "all ethnicities are allowed".

In each specification controls for the individual characteristics of apartments are added. Surface area, layout, floor number are explicit characteristics of apartment. To proxy for more ambiguous characteristics, I construct two variables: the length of announcement in

---

<sup>5</sup>See Gentzkow et al. (2017) for the review of various approaches in text analysis.

characters and the number of photos attached.

## 2.2. Other data

I complement the user-generated data from *cian.ru* with socio-economic data from the Russian Census (2010). Data on population, ethnic composition, level of education, fluency in Russian is grouped on *rayon* (subdistrict) level. I also use electoral statistics from the 2018 Russian presidential elections. This data is provided by the Central Election Commission of the Russian Federation.

In Appendix A I report the design of a correspondence experiment. I respond to a sample of ads through the online form and manipulate the names of potential tenants such that one group of names could be perceived as “Russian-sounding” and another group as “non-Russian-sounding”. There are no public data on birth names in Russia, so I construct an approximate ranking of names using data from the Russian social network *vk.com*. I use the data on the city of residence to make a rating of the most popular names in Moscow and Makhachkala — a multi-ethnic city where Russians make up only 5.4 percent.

### 3. Empirical analysis

#### 3.1. Estimating equation

The Moscow housing stock consists of multi-storey buildings with large number of apartments. The median building is 12-storey and multiple apartments are often exposed in one building.

When calculated for the entire observation period, the median building has around 12 apartments exposed. Apartments in the same building are usually of a similar quality, and “vertical” or in-building segregation is uncommon in Moscow. This structure of the housing stock is beneficial for my analysis: I employ a model with building level fixed effects to estimate the racial rent differential. The baseline specification is:

$$\log(RentPrice_{ib\tau}) = \alpha Discrim_{ib\tau} + X'_{ib\tau}\gamma + \sigma_b + \phi_\tau + \epsilon_{ib\tau} \quad (1.1)$$

Each observation is an ad that was posted within the observation period. Subscript  $i$  denotes a posted offer,  $b$  is an index of building and  $\tau$  is an index of the day when the offer was posted. *Discrim* is a dummy variable of interest that indicates the presence of discrimination in ad’s text.  $\sigma_b$  and  $\phi_\tau$  are building and day of posting fixed effects.

Building fixed effects allow to absorb the spatial and building specific variations. Coefficient of interest  $\alpha$  is an estimate of the cost of discrimination. It reflects the difference in the rent prices between discriminating and non-discriminating apartments. I also control for apartments’ individual characteristics: the set of controls  $X_{ib\tau}$ . The characteristics of the apartment are divided into two types: one that can be measured directly, such as surface area and apartment layout, and once that cannot be measured directly, such as general cleanliness, quality of repair, lack of dysfunctions. I try to control for these “soft” features using length of advertisement in characters and number of attached photos.

Less restrictive specifications were also tested: the model with *rayon* level fixed effects and the model with *okrug* level fixed effects. Both of these specifications include controls for distances to the city center and to the closest metro station.

This identification strategy holds several assumptions. First, I assume that discrimination in the ad is a direct reflection of real intention of landlord to discriminate. In latter part of this paper I also test the Moscow rental market for the presence of covert discrimination.

Second, I assume that the number of photos and length of text are good proxies for quality of apartment. I include other text-based measures of apartment quality for robustness.

I also explore how the racial rent differential depends on neighborhood characteristics, including the average level of discrimination in the neighborhood. The heterogeneity of the

effect is crucial for understanding the mechanism of the racial rent differential – theoretical discussion of the mechanism is presented in section 4. To do the heterogeneity analysis I interact the discrimination dummy with the share of discrimination in neighborhood and building:

$$\log(RentPrice_{ib\tau}) = \alpha Discrim_{ib\tau} + \beta Discrim_{ib\tau} \times DiscrRate_{iu} + X'_{ib\tau} \gamma + \sigma_b + \phi_\tau + \epsilon_{ib\tau} \quad (1.2)$$

For both neighborhoods and buildings the discrimination rates are calculated as a share of discriminating ads in total number of ads that were posted during the observation period excluding the contribution of interacted observation. Maps of discrimination rate calculated for buildings and subdistricts are shown in Figure 1.1.

$DiscrimRate_u$  is the surrounding discrimination rate for offer  $i$  in the unit  $u$ . This specification is tested for discrimination rates on different levels: buildings, *rayons* and *okrugs*.

### 3.2. Main results

#### 3.2.1. Racial rent differential

Table 1.2 presents the estimations of the racial rent differential. The extended table can be found in Table 1.B.1 in Appendix. The results bring out a strong and negative effect of discrimination on the price. The first column shows the results of the preferred specification: the one that includes building level fixed effects. I also include to the model time fixed effects (through variables that indicate the day when the ad appeared on the site) which helps to eliminate the impact of seasonality associated with the housing market. This specification also includes controls for individual characteristics of the apartment. Standard errors are clustered at the building level. This result indicates sizeable racial rent differential – around 4% of apartment's rent price.

Column two and three presents the results of the models with *rayon* and *okrug* level fixed effect correspondingly. These specifications also includes controls for logarithms of distances to the city center and the closest metro station. The fourth column presents results of the OLS regression without location-based fixed effects. It can be seen that the coefficient of interest increases from the first to the fourth specification. It can be explained by the fact that on average buildings and districts with less expensive property are also associated with discrimination.

### 3.2.2. Placebo and robustness

I estimate several placebo regressions that have the same equation as in the main specification presented in column 1 of Table 1.2. Instead of the discrimination variable I introduce two different text-based variables that also indicate preferences of the landlord: preference for tenants without kids and preference for tenants without pets. Results are presented in the Table 1.B.2. The coefficient for “no kids” variable is not significant, whereas the coefficient for “no pets” is significant, but relatively small – around 0.5% – and positive (unlike the main result obtained for the discrimination dummy). This positive effect for apartments that do not accept tenants with pets can be explained: potentially, landlords that historically did not accept tenants with pets were able to keep their property in better condition. I also repeat the main specification which is presented in the Table 1.2, but with text-based dummies from the placebo analysis as controls: the main result remains robust. Finally, I estimate the main specification including phone numbers fixed effects to absorb the variation in counterparty identities (however, phone variable does not allow to distinguish between landlords and agents). The coefficient decreases but not drastically – it stays around 3% (Table 1.B.3).

### 3.2.3. Heterogeneity analysis

The racial rent differential is not uniform across Moscow neighborhoods. To investigate how it changes, I perform heterogeneity analysis. Table 1.3 indicates that in neighborhoods with higher prevalence of discrimination the rent differential is smaller than in neighborhoods where discrimination is relatively rare. The same is true for the level of building. A higher share of discriminating apartments in a building is associated with a lower rent differential.

When it comes to other socio-economic characteristics of neighborhoods, we observe the following: the racial rent differential is *higher* in neighborhoods with a higher share of non-Russian residents, with a higher selling prices in housing, with a higher share of residents with higher education, with a higher share of votes for presidential candidates in ‘opposition’ to Vladimir Putin (Table 1.4).

As a result, we see that both distributions of frequency of discrimination and of the value of racial rent differential have the same center-periphery structure, but other meaningful variables also have a similar spatial distribution: education, population, average rent and purchase price of real estate, share of non-Russian residents.<sup>6</sup>

---

<sup>6</sup>See maps in section 5 and Figure 1.4

### *3.2.4. Impact of discrimination on search time*

The landlords' disadvantage from discriminating behaviour manifests itself through the increased search time.<sup>7</sup> Extra days spent on the market waiting should naturally be considered as a part of cost of discrimination. Table 1.B.4 presents the estimated effect of discrimination on the number of days offers have been exposed on the platform. The data used in this analysis do not include observations that were available on the first day and observations that stay on the site on the last day of the observation period. Specifications in Table 1.B.4 are similar to the ones from Table 1.2, but with the logarithm of number of days in exposure in left-hand side. In each regression I control for logarithm of apartment's rent price.<sup>8</sup>

An apartment that does not accept non-slavic tenants remains on the market 10 % longer. This effect is not particularly large if we take into account that for an average ad it turns into one extra day. Though it is a costly delay, but one that landlords suffer only occasionally — in contrast to the monthly rental discount.

### *3.3. Results of experiment*

The design of an experiment is presented Appendix A in Appendix. Table 1.5 presents the results of an experiment. Each column presents the outcomes of a probit regression where the dependent variable is an answer dummy: one, if counterparty replied to the message and answered the question, and otherwise – zero. This experiment provides us with several important results. First, indeed, applicants with non-Russian sounding names have significantly lower probability of receiving benevolent response from apartments' accounts that have racial preferences in ads. At the same, it is also true to a certain degree for non-discriminating accounts: non-Russian applicants have a lower chance to receive a reply than Russian applicants even from accounts that have no racial preferences in ads (Table 1.5). This result speaks in favor of coexistence of overt and subtle forms of discrimination in the Moscow rental housing. There is another important result, which can be seen in the Table 1.6. This table presents subsample analysis: it takes ads without racial preferences and splits the sample by neighborhoods. The city center is notable for the low level of overt discrimination, however, one could suggest that landlords in this elite neighborhood switch from overt to subtle discrimination. The experiment's results do not support this hypothesis. Subtle discrimination is more prevalent in the outskirts, so, on the average, subtle discrimination is proportional to neighborhood's overt discrimination.

---

<sup>7</sup>However, despite the fact that it is impossible to observe whether the apartment is really rented out, the date when the offer disappears from the platform can be used as the best possible approximation.

<sup>8</sup>Prices on the last day are used here.

## 4. Theory

The Beckerian neoclassical framework fails to explain the persistence of the cost of discrimination. In this setting both landlords and tenants are price-takers. Two markets, discriminating and equally accessible, exists with two rents respectively:  $p_d$  and  $p_{nd}$ .

Assume that predictions of the model are in line with the empirical findings and  $p_d^* < p_{nd}^*$ . This scenario intends full market segregation. Otherwise, the majority from the discriminating market will move to another market until rents equalize. However, the full segregation is implausible since it means that majority constitutes only 20% of the rental housing market.

Literature on discrimination in the labor market solves this issue by introducing frictional environment. The notable contributions in this direction were made by Black (1995), Rosén (1997), Bowlus and Eckstein (2002), Lang et al. (2005).

### 4.1. Baseline model

In this section I adapt the random search model from Black (1995) to the context of Moscow rental housing. To take into account the heterogeneous structure of the Moscow housing market, I consider the model with two "neighborhoods" between which potential tenants are sorted.

There are two neighborhoods  $A$  and  $B$ . Both of them are functioning as independent rental housing markets. There are two types of landlords in both neighborhoods: discriminating (those who refuse to rent an apartment to a *non-slavic* tenant at any price), and non-discriminating (those who are indifferent of tenant's race). The share of discriminating landlords in the neighborhood  $i$  is  $\theta_i$ . I assume that the neighborhood  $B$  is more discriminating, i.e.  $\theta_B > \theta_A$ .

#### 4.1.1. Sorting

There are two types of tenants: *slavic* and *non-slavic*. The share of *slavic* tenants is  $\pi$ , and the share of *non-slavic* tenants is  $1 - \pi$ . Each *slavic* and *non-slavic* tenant chooses the probability of entering the neighborhood  $A$  with probabilities  $q_s$  and  $q_{ns}$  respectively, and of entering the neighborhood  $B$  with probabilities  $1 - q_s$  and  $1 - q_{ns}$ . As a result, the shares of *slavic* tenants in the neighborhoods  $A$  and  $B$  are:

$$\begin{aligned}\pi_A &= \frac{q_s\pi}{q_s\pi + q_{ns}(1 - \pi)} \\ \pi_B &= \frac{(1 - q_s)\pi}{(1 - q_s)\pi + (1 - q_{ns})(1 - \pi)}\end{aligned}$$

*Slavic* and *non-slavic* tenants extract reservation utilities  $V_s^i$  and  $V_{ns}^i$  respectively from the rental housing market. These reservation utilities will be described below.

In a general setting, when residents decide where to live, they take many factors into account: prices, access to schools, proximity to workplace, amenities and more. While this paper does not aim to model the sorting process in an extensive way, it is still important to introduce to the model motives not related to rental housing. In this stylized model I assume that neighborhood with a lower share of discrimination  $A$  is also a central district with rich amenities and better access to work and schooling (which correspond to the Moscow context). Assume, there are shares of both *slavic* and *non-slavic* potential tenants who are attached to the central district  $A$ ,  $\mu_s \leq q_s$  and  $\mu_{ns} \leq q_{ns}$ . They do not choose between neighborhoods and search apartments in  $A$  by default. After “mobile” tenants choose their neighborhoods, tenants of all types start apartment search in their respective neighborhoods.

## 4.2. Search

Within each neighborhood tenants of both types sequentially search for an apartment paying  $k$  for each period of the search. When a tenant finds and rents an apartment, he or she stops searching and lives in this apartment forever.

Tenants learn three features during the visit of the apartment online page: how much they value this apartment –  $\alpha$ , the type of landlord and the rent  $p$  that was set in advance by the landlord. While this mechanism does not fully take into account the informational structure of the online platform, it approximates the search process online: tenants need to invest their time and effort in studying ads. The individual value of apartment  $\alpha$  is randomly distributed with distribution function  $F(\alpha)$  and density function  $f(\alpha)$ . Following Black I assume  $F(\alpha)$  is strictly log-concave.

There is an important deviation from Black (1995) when it comes to price setting. The main interest of Black’s model is the racial wage gap, where employers can set different wages for individual members of minorities and non-minorities. In my model I assume that non-discriminating landlord sets a unique rent price for both *slavic* and *non-slavic* tenants, and a discriminating landlord sets a price for *slavic* tenants and do not accept *non-slavic* tenants at any price.

### 4.2.1. Tenants’ problems

Tenants’ equilibrium strategies can be described with reservation utilities such that tenants are indifferent between renting an apartment and continuing the search. Two options available for *slavic* tenants: renting an apartment from a discriminating landlord and renting

an apartment from a non-discriminating. This leads to the following dynamic equation:

$$V^s = \theta \mathbb{E} \max\{\alpha - p_d, V^s\} + (1 - \theta) \mathbb{E} \max\{\alpha - p_{nd}, V^s\} - k \quad (1.3)$$

Minorities' problem looks different: with probability  $\theta$  they meet a discriminating landlord and, therefore, they cannot rent this apartment and receive their reservation utility.

$$V^{ns} = \theta V^{ns} + (1 - \theta) \mathbb{E} \max\{\alpha - p_{nd}, V^{ns}\} - k \quad (1.4)$$

#### 4.2.2. Landlords' problem

Each landlord behaves as a monopsonistic competitor. Therefore, they maximize the rent, considering probabilities of tenants' acceptance. Discriminating landlords rent an apartment if and only if tenant is *slavic*. Thus, their expected utility can be written as:

$$\mathbb{E} u_d = (1 - F(V^s + p_d))p_d \quad (1.5)$$

Non-discriminating landlords accept tenants of both types and they set a unique price to tenants of both types.

$$\mathbb{E} u_{nd} = p_{nd}(\pi(1 - F(V^s + p_d)) + (1 - \pi)(1 - F(V^{ns} + p_{nd}))) \quad (1.6)$$

#### 4.2.3. The Optimal Rents and the Racial Rent Differential in a Separate Neighborhood

Assume that  $\alpha$  is drawn from uniform distribution on interval  $[0, \beta]$ . Then the equilibrium rent prices of both discriminating and non-discriminating apartments are defined by a system of two equations. For a neighborhood  $i \in \{A, B\}$  this system can be written as:

$$2k\beta = \theta^i(p_d^i)^2 + (1 - \theta^i)(2p_d^i - p_{nd}^i)^2 \quad (1.7)$$

$$p_{nd}^i = \frac{1 - \pi^i}{1 + \pi^i} \sqrt{\frac{2\beta k}{1 - \theta^i}} + \frac{2\pi^i}{1 + \pi^i} p_d^i \quad (1.8)$$

, where  $p_{nd}^i$  and  $p_d^i$  are rent prices of discriminating and non-discriminating apartments in neighborhood  $i$ ,  $\theta^i$  is a share of discriminating landlords in neighborhood  $i$  and  $\pi^i$  is a share of *slavic* tenants in neighborhood  $i$ .

Several facts follow from of this system. First, it shows the existence of the racial rent differential presented in the empirical part of this paper (Section 3).

**Proposition 1.**  $\Delta = p_{nd} - p_d > 0$  for any value of  $\theta$  and  $\pi$  when non-slavic tenants participate in a search, i.e.  $V_d(\theta, \pi) > 0$ .

Second, it can be shown that, consistently with the empirical findings,  $\Delta^i(\theta^i, \pi^i)$  is decreasing with an increase of  $\pi^i$ , share of potential *slavic* tenants in the neighborhood  $i$ . However, conflicting with the evidence I found,  $\Delta^i(\theta^i, \pi^i)$  is increasing with the share of discriminating apartments  $\theta^i$ .

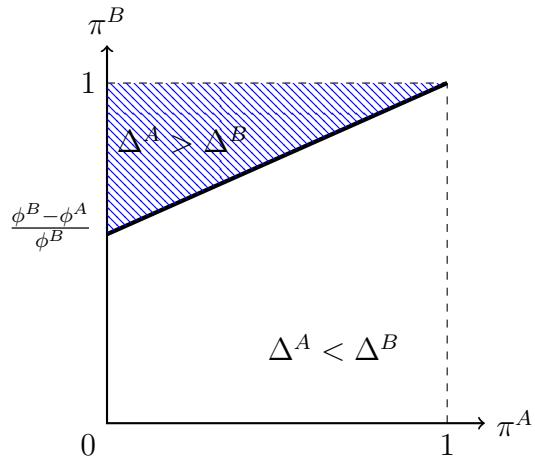
**Proposition 2.** For any given  $\theta \in (0, 1)$   $\Delta(\theta, \pi)$  is decreasing with  $\pi$ . For any given  $\pi \in (0, 1)$   $\Delta(\theta, \pi)$  is decreasing with  $\theta$ .

The interpretation of this relationship is as follows: with an increase of the share discriminating apartment frictions for *non-slavic* tenants increase and non-discriminating landlords respond with increased rent prices, therefore the differential increases.

However, in this setting it is still possible that the neighborhood with a higher share of discriminating apartments has a higher racial rent differential, because the differential also depends on the share of *slavic* tenants in the neighborhood.

#### 4.3. Racial rent differentials in two neighborhoods

Suppose, there are two neighborhoods  $A$  and  $B$ , such that  $\theta^B > \theta^A$ . Assume that the shares of discriminating apartments  $\theta^i$  are exogenous characteristics of a neighborhood. It can be shown that in an interval  $\pi^i \in (0, 1)$  function  $\Delta(\pi^i)$  can be well-approximated with a linear function  $\Delta(\pi^i) = -\phi^i(\theta^i)\pi^i + \phi^i(\theta^i)$ , where  $\phi^i(\theta^i)$  is a coefficient that depends on a share of discrimination in neighborhood  $\theta^i$ . Therefore, it can be shown that for neighborhoods  $A$  and  $B$  two spaces consisting of pairs  $(\pi^A, \pi^B)$  exist: one, for which  $\Delta^A > \Delta^B$ , and one, for which  $\Delta^A < \Delta^B$ .



**Proposition 3.** *The city economy can reach such equilibrium that  $\Delta^A > \Delta^B$  when*

$$(\pi^A, \pi^B) = \left( \frac{\mu_s \pi}{\mu_s \pi + \mu_{ns}(1 - pi)}, \frac{(1 - \mu_s)\pi}{1 - \mu_s \pi + (1 - \mu_{ns})(1 - \pi)} \right)$$

In this case, both *slavic* and *non-slavic* mobile tenants will sort to the neighborhood  $B$ . For such equilibrium to appear we should assume sufficiently large share of non-mobile *non-slavic* tenants, which in reality can be interpret as either high attachment to services accessible in the city center or high attachment to non-discriminating environment.

Despite the fact that this model is highly stylized, it still shows how heterogeneous effects found in empirical section of this paper can emerge. It also corresponds to the fact that the share of non-Russian residents is higher in the Moscow city center than on the outskirts, according to the Census (2010).

## 5. Conclusion

Racial discrimination can generate significant racial disparities in economic outcomes: I find that an apartment with a discriminatory ad has 4% lower rent price than an identical, but non-discriminating apartment in the same building. This result complements well-established theoretical insights on how differential treatment can generate racial differentials in the housing market. While there are many channels through which racial differentials can occur, pure discrimination in the market remains important and requires further research.

This paper touches on the uncovered topic of the relationship between overt and subtle forms of discrimination. I analyse unique data from the Moscow rental housing, where landlords do not hide their racial preferences. I show that overt and subtle forms of discrimination are closely related. I find that they coexist in Moscow rental housing market and that their relative prevalence is stable across neighborhoods.

Finally, I borrow theoretical framework from the literature on labor search with discrimination and show how the racial rent differential can occur. I do heterogeneity analysis and find that the racial rent differential is higher in neighborhoods with a lower share of discriminating landlords. I show that this result can coincide with a random search model with discrimination by introducing the stylized version of neighborhood sorting.

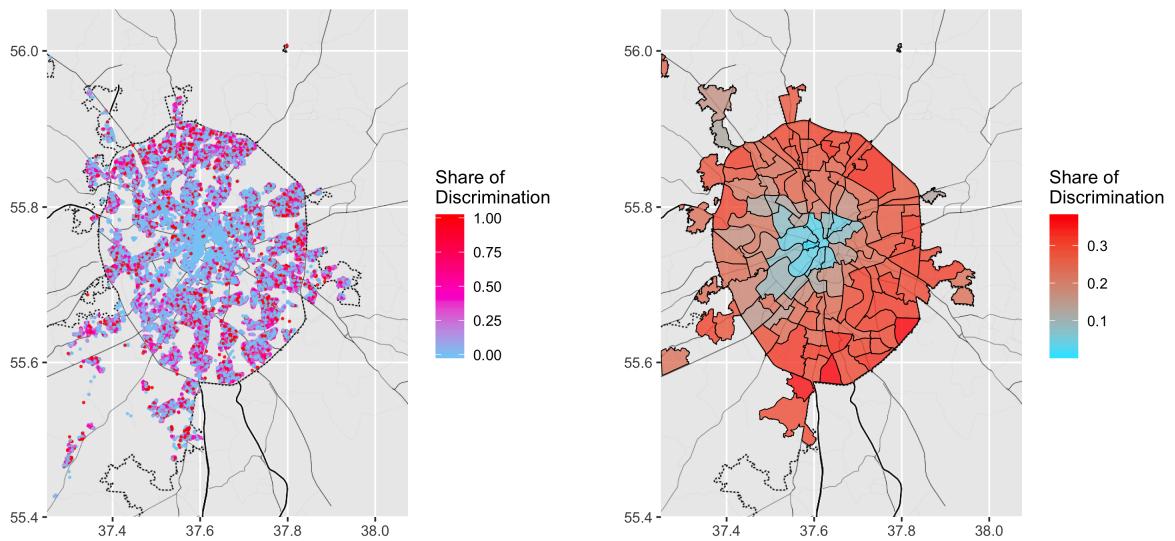
# References

- Acolin, A., Bostic, R., and Painter, G. (2016). A field study of rental market discrimination across origins in france. *Journal of Urban Economics*, 95:49–63.
- Ahmed, A. M. and Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the internet. *Journal of Urban Economics*, 64(2):362–372.
- Albrecht, J., Gautier, P. A., and Vroman, S. (2016). Directed search in the housing market. *Review of Economic Dynamics*, 19:218–231.
- Alesina, A., Miano, A., and Stantcheva, S. (2018). Immigration and redistribution. Technical report, National Bureau of Economic Research.
- Arrow, K. J. (1972). Models of job discrimination. *Racial discrimination in economic life*, 83.
- Arrow, K. J. (1998). What has economics to say about racial discrimination? *Journal of economic perspectives*, 12(2):91–100.
- Baert, S. (2018). Hiring discrimination: an overview of (almost) all correspondence experiments since 2005. In *Audit studies: Behind the scenes with theory, method, and nuance*, pages 63–77. Springer.
- Bayer, P., Casey, M., Ferreira, F., and McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, 102:91–105.
- Becker, G. S. (2010). *The economics of discrimination*. University of Chicago press.
- Bertrand, M., Chugh, D., and Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, 95(2):94–98.
- Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 1:309–393.

- Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.
- Bessudnov, A. and Shcherbak, A. (2018). Ethnic hierarchy in the russian labour market: A field experiment.
- Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, 13(2):309–334.
- Bowlus, A. J. and Eckstein, Z. (2002). Discrimination and skill differences in an equilibrium search model. *International Economic Review*, 43(4):1309–1345.
- Carpusor, A. G. and Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology*, 36(4):934–952.
- Carrillo, P. E. (2012). An empirical stationary equilibrium search model of the housing market. *International Economic Review*, 53(1):203–234.
- Combes, P.-P., Decreuse, B., Schmutz, B., and Trannoy, A. (2018). Neighbor discrimination theory and evidence from the french rental market. *Journal of Urban Economics*, 104:104–123.
- Courant, P. N. (1978). Racial prejudice in a search model of the urban housing market. *Journal of Urban Economics*, 5(3):329–345.
- Early, D. W., Carrillo, P. E., and Olsen, E. O. (2019). Racial rent differences in us housing markets: Evidence from the housing voucher program. *Journal of Regional Science*, 59(4):669–700.
- Fryer, R. G., Pager, D., and Spenkuch, J. L. (2013). Racial disparities in job finding and offered wages. *The Journal of Law and Economics*, 56(3):633–689.
- Gentzkow, M., Kelly, B. T., and Taddy, M. (2017). Text as data. Technical report, National Bureau of Economic Research.
- Hanson, A. and Hawley, Z. (2011). Do landlords discriminate in the rental housing market? evidence from an internet field experiment in us cities. *Journal of Urban Economics*, 70(2-3):99–114.
- Heckman, J. J. (1998). Detecting discrimination. *Journal of economic perspectives*, 12(2):101–116.

- Hedegaard, M. and Tyran, J.-R. (2014). The price of prejudice.
- Ihlanfeldt, K. and Mayock, T. (2009). Price discrimination in the housing market. *Journal of Urban Economics*, 66(2):125–140.
- Kuhn, P. and Shen, K. (2012). Gender discrimination in job ads: Evidence from china. *The Quarterly Journal of Economics*, 128(1):287–336.
- Lang, K. and Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4):959–1006.
- Lang, K., Manove, M., and Dickens, W. T. (2005). Racial discrimination in labor markets with posted wage offers. *American Economic Review*, 95(4):1327–1340.
- Martin, T. D. et al. (2001). *The affirmative action empire: nations and nationalism in the Soviet Union, 1923-1939*. Cornell University Press.
- Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, 104(5):869–895.
- Ngai, L. R. and Tenreyro, S. (2014). Hot and cold seasons in the housing market. *American Economic Review*, 104(12):3991–4026.
- Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science*, 609(1):104–133.
- Pager, D. and Karafin, D. (2009). Bayesian bigot? statistical discrimination, stereotypes, and employer decision making. *The Annals of the American Academy of Political and Social Science*, 621(1):70–93.
- Rosén, Å. (1997). An equilibrium search-matching model of discrimination. *European Economic Review*, 41(8):1589–1613.
- Small, M. L. and Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, 34(2):49–67.
- Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics*, 118:26–40.
- Vendina, O. (2002). Social polarization and ethnic segregation in moscow. *Eurasian Geography and Economics*, 43(3):216–243.

- Vendina, O., Panin, A., and Tikunov, V. (2019). The moscow social space: features and structure. *Regional Research of Russia*, 9(4):383–395.
- Veterinarov, V. and Ivanov, V. (2018). Slavs only: Ethnic discrimination and rental prices. Available at SSRN 3249624.
- Yinger, J. (1986). Measuring racial discrimination with fair housing audits: Caught in the act. *The American Economic Review*, pages 881–893.
- Yinger, J. (1997). Cash in your face: The cost of racial and ethnic discrimination in housing. *Journal of Urban Economics*, 42(3):339–365.

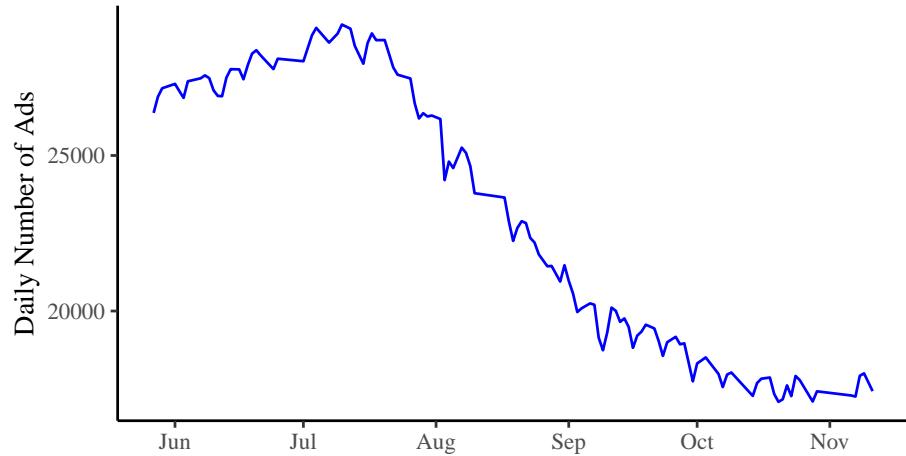


(a) Map of discrimination by buildings

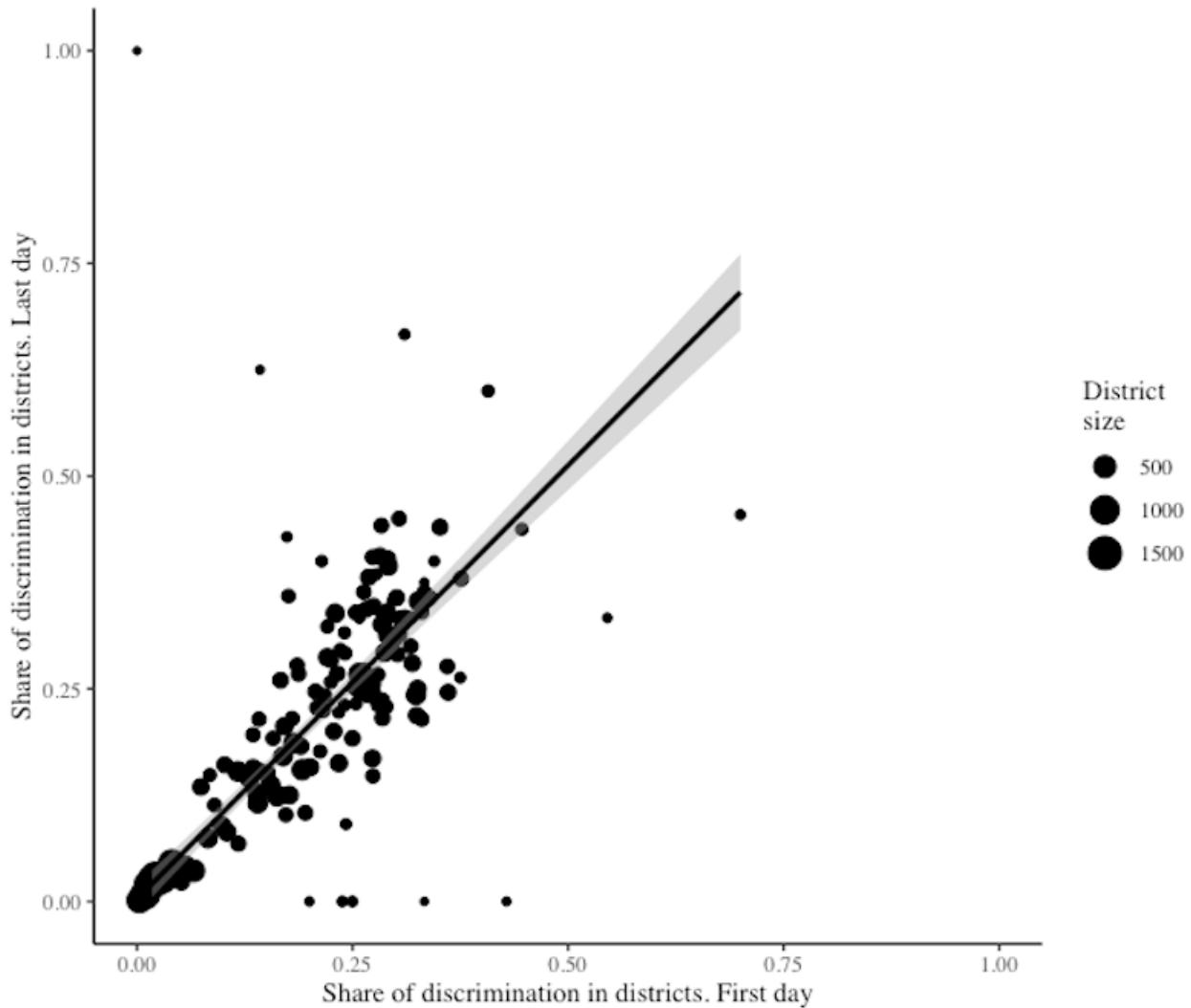
(b) Map of discrimination by subdistricts

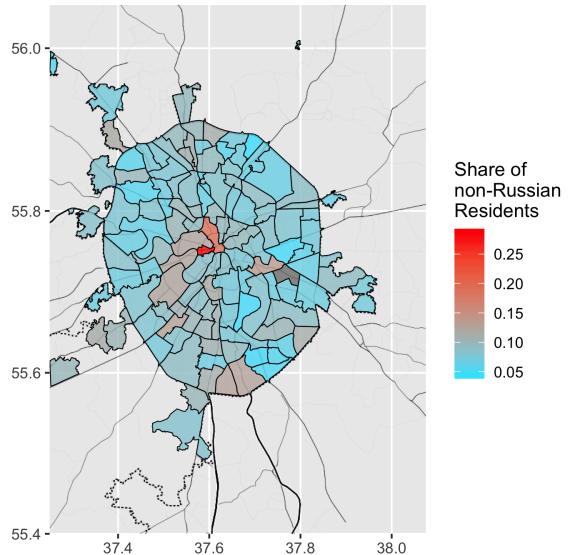
### Geography of discrimination

Daily number of ads posted on the platform

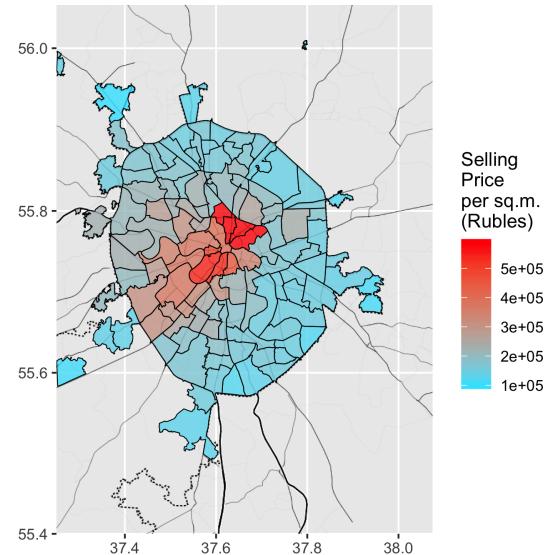


Share of discrimination by neighbourhoods on the first and last days of the observational period

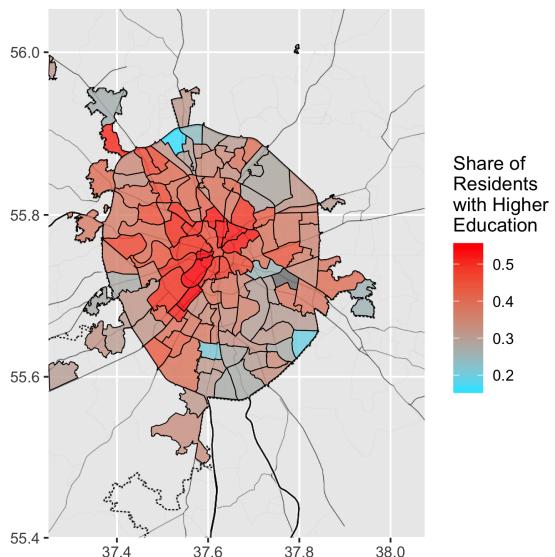




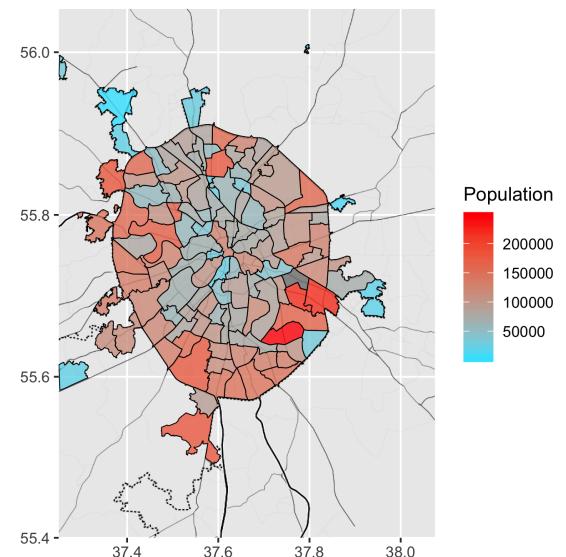
(a) Share of Non-Russian Residents



(b) Rent Price per sq. m.

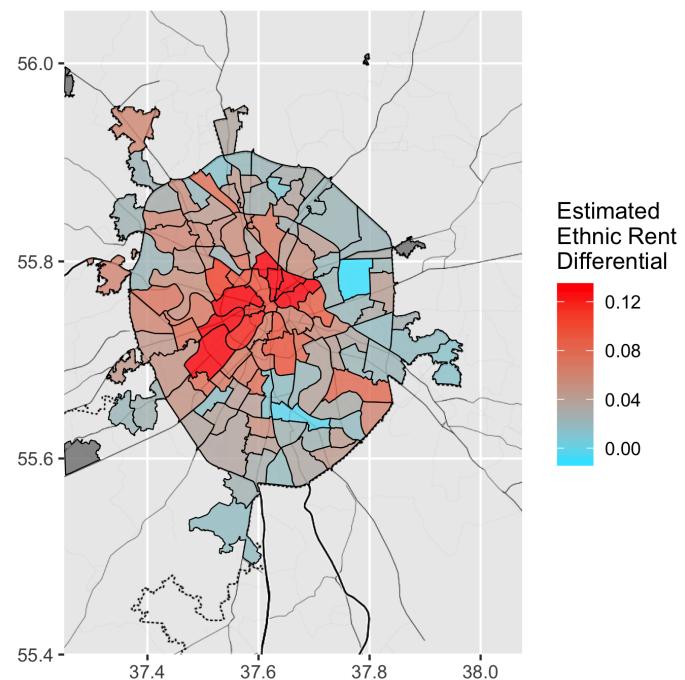


(c) Share of Residents with Higher Education



(d) Population (thousands)

The racial rent differential by districts (*rayons*)



## 6. Tables

Table 1.1: Descriptive statistics

Panel A. Apartments exposed during the observation period

	Obs	Mean	Std. Dev	Min	Max
Price (rubles)	139,965	72,190	92,962	14,500	1,024,106
Kitchen area (sq.m.)	139,965	10.27	5.42	1	160
Living area (sq.m.)	139,965	38.14	27.58	.9	450
Total area (sq.m.)	139,965	62.65	41.00	10	500
Floor number	139,965	7.06	5.74	1	85
Days in exposure	139,965	18.48	29.76	0	168
Length of text (symbols)	139,965	800.19	527.51	52	3743
Number of photos	139,965	12.09	7.59	0	50
Declare descrimination	139,965	.20	.40	0	1
Declare inclusivity	139,965	.005	.07	0	1

Panel B. Buildings' characteristics

Number of floors	20,417	10.27	5.42	1	160
Distance to city center (km)	20,417	11.59	5.85	.24	59.80
Distance to closest metro (km)	20,417	1.36	2.21	.005	55.89
Share of discriminating apartments	20,417	.24	.28	0	1

Panel C. Subdistricts' characteristics

Share of discriminating apartments	140	.23	.08	.009	.54
Population (thousands)	125	92	43	3	247
Share of non-Russian	125	.08	.02	.04	.28
Share of Central Asian population	124	.007	.006	.002	.03
Share of North Caucasian population	122	.004	.002	.001	.02
Share of Jewish population	125	.005	.003	.0008	.02
Price per sq. m. (rubles)	140	886	267	443	1863

Panel C. Districts' characteristics

Share of discriminating apartments	12	.23	.06	.05	.33
---------------------------------------	----	-----	-----	-----	-----

Table 1.2: Main result: The Racial Rent Differential

	Dep. Var.: Logarithm of rent price			
	(1)	(2)	(3)	(4)
Discrimination dummy	-0.0409*** (0.001)	-0.0638*** (0.004)	-0.0670*** (0.008)	-0.0743*** (0.003)
Observations	139,965	139,965	139,965	139,965
Building FE	Yes			
Subdistrict FE		Yes		
District FE			Yes	
Day of posting FE	Yes	Yes	Yes	Yes
Controls (apartment char.)	Yes	Yes	Yes	Yes
Controls (building char.)		Yes	Yes	Yes

*Note:* Estimation of the effect of overt discrimination in the ad on the rent price. Each observation is an individual ad posted on the website *cian.ru* during the observation period from May 27 to November 11, 2018. Standard errors are clustered on the building, *rayon* and *okrug* levels in specifications (1), (2) and (3) correspondingly. Standard errors in parenthesis.

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

Table 1.3: Heterogeneous effects: the Racial Rent Differential and the Share of Discrimination in Neighborhood

	Dep. Var.: Logarithm of Rent Price			
	(1)	(2)	(3)	(4)
Discrimination	-0.0409*** (0.001)	-0.0488*** (0.002)	-0.1009*** (0.006)	-0.1030*** (0.007)
discrimination dummy		0.0339*** (0.007)		
Discrimination dummy × Share of discrimination in building				
Discrimination dummy × Share of discrimination in subdistrict			0.2463*** (0.022)	
Discrimination dummy × Share of discrimination in district				0.2660*** (0.029)
Average of interacting variable	.074	.052	.050	
Maximum of interacting variable	1	.52	.33	
Observations	139,965	139,965	139,965	139,965
Building FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

*Note:* Estimation of the heterogeneous effect of overt discrimination in the ad on the rent price. Interaction terms are dummy for discrimination interacted with shares of discrimination in buildings, subdistricts and districts. Each observation corresponds to an individual ad posted on the website *cian.ru* during the observation period from May 27 to November 11, 2018. Standard errors are clustered on the level of buildings. Standard errors in parenthesis.

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

Table 1.4: Heterogeneous Effects: Interactions with Characteristics of Neighborhood

	Dependent variable: Logarithm of rent price			
	(1)	(2)	(3)	(4)
Discrimination	0.7024***	0.0214***	0.0112**	-0.0168***
dummy	(0.061)	(0.007)	(0.005)	(0.006)
Discrimination	-0.0613***			
dummy ×	(0.005)			
<i>Housing selling</i>				
<i>price in district</i>				
Discrimination		-0.1739***		
dummy × <i>Higher</i>		(0.021)		
<i>education in</i>				
<i>district</i>				
Discrimination			-0.5560***	
dummy × <i>Votes</i>			(0.053)	
<i>for 'liberals'</i>				
Discrimination				-0.2927***
dummy × <i>Share</i>				(0.069)
<i>of 'non-Russians'</i>				
Observations	139,965	139,965	139,965	139,965
Building FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

*Note:* Estimation of the heterogeneous effect of overt discrimination in the ad on the rent price. Interaction terms are dummy for discrimination interacted with characteristics of neighborhoods. Each observation corresponds to an individual ad posted on the website *cian.ru* during the observation period from May 27 to November 11, 2018. Standard errors are clustered on the level of buildings. Standard errors in parenthesis.

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

Table 1.5: Experiment: Main Results

	Dependent variable: Reply rate (dummy)		
	All ads	Ads without discrimination	Ads with discrimination
	(1)	(2)	(3)
Non-Russian name	-0.5511*** (0.091)	-0.3596*** (0.130)	-0.7631*** (0.130)
Observations	874	444	430
Order dummy	Yes	Yes	Yes
Text dummy	Yes	Yes	Yes
Price (log)	Yes	Yes	Yes
Total area (log)	Yes	Yes	Yes
Length of text (log)	Yes	Yes	Yes
Ground floor	Yes	Yes	Yes
Last floor	Yes	Yes	Yes

*Note:* Each column gives the results of a probit regression where the dependent variable is the answer dummy: one denotes benevolent reply from agent/landlord and zero denotes non-response (while message has been read) or refusal. Robust standard errors in parenthesis.

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

Table 1.6: Experiment: Subset of ads without overt discrimination

	Dependent variable: Reply rate (dummy)		
	All districts	Less discriminating districts	More discriminating districts
	(1)	(2)	(3)
Non-Russian name	-0.3596*** (0.130)	-0.3079* (0.168)	-0.4923** (0.209)
Observations	444	272	172
Order dummy	Yes	Yes	Yes
Text dummy	Yes	Yes	Yes
Price (log)	Yes	Yes	Yes
Total area (log)	Yes	Yes	Yes
Length of text (log)	Yes	Yes	Yes
Ground floor	Yes	Yes	Yes
Last floor	Yes	Yes	Yes

*Note:* Each column gives the results of a probit regression where the dependent variable is the answer dummy: one denotes benevolent reply from agent/landlord and zero denotes non-response (while message has been read) or refusal. The sample consists of only ads without overt discrimination. Robust standard errors in parenthesis.

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

## A. Appendix: Design of Correspondence Experiment

Moscow landlords and agents explicitly discriminate against minorities in rental ads. However, it is not entirely clear whether discrimination in ads really turns into active discrimination in marketplace. It is also not necessary that landlords, who do not use language of discrimination, do not discriminate privately. In this section I explore these possibilities with help of correspondence experiment.

Since seminal paper by Bertrand and Mullainathan (2004) economists extensively use approach of correspondence study to reveal racial, ethnic or gender discrimination on various markets.<sup>9</sup> This approach is based on direct manipulation of applicants characteristics, specifically names, when it comes to the subject of racial discrimination. In this way, Bertrand and Mullainathan randomly assigned African-American sounding names to job applicant's resumes, send these resumes to real employers in Boston and Chicago and compared call backs rates of two racial groups. This study revealed that applicants with African-American names have statistically and economically significantly lower probability of call back.

I conduct correspondence experiment using online contact form which is available on the platform and which allows to reach a person behind the ad. I use design of paired-matched applications and send couples of short messages with Russian and non-Russian identities. Experiment was conducted in two separate rounds.

### A.1. Messages

The platform provides users who are looking for apartments, two alternative ways to contact landlords or agents: via a public mobile phone or through an online form. The second is intended to ask the landlord or agent a short clarifying question about the proposal. The online form was chosen as the communication device for the experiment for technical reasons.

Following the way the online form is organized, I built two simple questions that were used as the basis for the intervention. Translations of these two questions are following:

**Q1.** *Hello, I'm interested in your apartment. May I contact you tonight? [First name]*

**Q2.** *Good afternoon, your offer interested me. I would like to ask a clarifying question. When could one move to an apartment? [First name]*

---

<sup>9</sup>See Baert (2018) for review of correspondence experiments

As can be seen, the topics of the questions are not related to the topic of ethnic discrimination. The sole purpose of these questions is to enable landlords (or agents) to react to the name of the applicant. The online form is not the main means of communication: its role is to be an intermediate stage before a telephone conversation, which in itself is an intermediate stage before a personal visit to the apartment. As a rule, the online form is not used to conclude transactions or discuss conditions. Therefore, the experiment was designed in such a way that the landlords could ignore the messages of the applicants with non-Russian names and, thus, disrupt the interaction at the first stage.

## A.2. Names and identities

When the applicant submits his message through the form, the landlords can observe only the message itself. Despite this, separate accounts with realistic email addresses were created for each identity.

The variation of perceived ethnicity of names is a treatment of the experiment. Two rounds of experiment were conducted. They are different in terms of name selection approaches. It is important to note here that in Russia there is no common dataset on birth names. For the first round of the experiment, only two names were chosen: the Russian-speaking name Andrei and the Turkic name Arslan. Both names are popular and recognisable in Russia.

In the second round, a more rigorous approach to names selection was used. Between the first and second stages of the experiment, I created an original set of data on names in Russia, using account statistics collected from the popular Russian social network *vk.com*. Ratings of names by popularity for each Russian city was constructed.

Two cities were selected among the entire set: Moscow and Makhachkala. The first is a city in which the majority of the population is Russian: around 90 percent according to 2010 Russian Census. The second is plural city with only 6.3 percent of Russian residents. The largest ethnic groups in this region are among the most discriminated groups in the Moscow housing market and labor market.<sup>10</sup> Most of the representatives of these ethnic groups are citizens of Russia.

I take the 10 most popular names in Moscow and the 10 most popular names in Makhachkala, excluding the first places in the ranking and the names used in the first round of the experiment. The resulting set of names was used in the second round.

---

<sup>10</sup>Bessudnov and Shcherbak (2018) find that Chechen job seekers have one of the lowest callback rates. Given that the set of names of largest ethnic groups in Dagestan intersects widely with the set of Chechen names, this result is valid for the most popular names of Makhachkala residents.

### *A.3. Sending messages*

The experiment was conducted in two rounds: June 20-21, 2018 and December 13-14, 2019. The design of the second round was changed due to the fact that statistics on names became available. In this section, I describe the procedure and schedule of the first round of experiment and difference between first and second round.

The sample was constructed from the set of new offers that become available on the platform during the night 19-20 June, 2018. To identify these offers, I select those ones that appeared this night and were not available on previous days.

The next step, I randomly remove from the sample all offers with duplicate phone numbers, except one. Landlords or agents with duplicate phone numbers are coordinating the rental processes of more than one apartment. By design of experiment it is necessary not to contact one person through several different offers' pages. Such messages can be perceived as conspicuous and can bias results of experiment.

At this stage, 291 new discriminating offers were obtained. I randomly select other 291 offers among non-discriminating set. The resulting 582 observations become the sample of the first round of experiment.

As a final preparatory phase, texts of messages and identities for the first request were randomly independently attached to each offer. For the second paired message another text and alternative identity were used.

Finally, during the day of June 20, I manually sent the first message through the form of each offer. The process of sending messages is difficult to automate, because the platform prevents such interventions. The next day, requests with alternative texts and names were sent via forms with the same offers. The one day period was chosen as long enough to be realistic and short enough to decrease the number of cases when offers are no longer available to the time of second message.

Thanks to the randomization of the order and message texts, the influence of these two factors do not influence results.

During the second round names of two groups were randomized.

### *A.4. Classification of responses*

Landlords or agents can reply in free form, however several basic types were identified. Classification is following:

1. Answer question or ask to call
2. Ask extended identification of potential tenant/ explicitly ask about ethnicity

3. “Already rented”
4. Message was not read
5. Read, but not answered
6. Rejects, motivating this with the tenant’s ethnicity
7. Rejects, motivating this with the tenant’s gender

Landlords or agents do not have other ways to communicate with potential tenant, therefore there are no other possible response ways to be coded.

In analysis of experiment’s outputs, this classification was simplified. Point 1 was considered as “likely non-discriminating”, points 2, 3, 5, 6, 7 is combined in one category “likely discriminating”. Observations with point 4 replies were excluded from the analysis.

## B. Appendix: Empirical Results

Table 1.B.1: The Racial Rent Differential: Extended Table

	Dependent variable: Logarithm of rent price			
	(1)	(2)	(3)	(4)
Discrimination dummy	-0.0409*** (0.001)	-0.0638*** (0.004)	-0.0670*** (0.008)	-0.0743*** (0.003)
Log total surface	0.7091*** (0.007)	0.8817*** (0.025)	0.8972*** (0.052)	0.9204*** (0.010)
LivingArea / TotalArea	0.1964*** (0.013)	0.1918*** (0.037)	0.2224*** (0.027)	0.2023*** (0.026)
Number of floors		0.0095*** (0.001)	0.0101*** (0.000)	0.0106*** (0.001)
Ground floor	-0.0198*** (0.003)	-0.0078 (0.005)	-0.0022 (0.007)	-0.0040 (0.006)
Last floor	0.0139*** (0.003)	0.0057 (0.005)	0.0062 (0.004)	0.0060 (0.005)
Log dist. to center		-0.2741*** (0.029)	-0.3069*** (0.018)	-0.3383*** (0.006)
Log dist. to metro		-0.0296*** (0.005)	-0.0400*** (0.005)	-0.0390*** (0.003)
Log(number of photo + 1)	0.0084*** (0.001)	0.0134*** (0.002)	0.0144*** (0.002)	0.0168*** (0.001)
Log length of text (10 chars)	0.0280*** (0.001)	0.0432*** (0.002)	0.0443*** (0.003)	0.0468*** (0.002)
Log days in exposure	0.0148*** (0.001)	0.0217*** (0.001)	0.0217*** (0.003)	0.0229*** (0.001)
Constant	7.7410*** (0.023)	7.4413*** (0.141)	7.4171*** (0.260)	7.3820*** (0.037)
Observations	139,965	139,965	139,965	139,965
R-squared	0.952	0.890	0.882	0.876
Building FE	Yes			
Subdistrict FE		Yes		
District FE			Yes	
Day of posting FE	Yes	Yes	Yes	Yes

*Note:* The sample consists of all ads posted on the web-site during the observation period. Standard errors are clustered on the level of buildings, subdistricts and districts in specifications (1), (2) and (3) correspondingly.  
Standard errors in brackets.

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

Table 1.B.2: Placebo: Other Preferences of Landlords

	Dependent variable: Logarithm of rent price		
	(1)	(2)	(3)
No animals	0.0050** (0.002)		0.0164*** (0.002)
No kids		-0.0020 (0.002)	0.0048** (0.002)
Only for Slavs			-0.0430*** (0.001)
Observations	139,965	139,965	139,965
Building FE	Yes	Yes	Yes
Day of posting FE	Yes	Yes	Yes
Controls (apartment char.)	Yes	Yes	Yes

*Note:* Standard errors are clustered on the level of buildings. Standard errors in parenthesis. \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1

Table 1.B.3: Robustness: Phone Numbers Fixed Effects

	Dependent variable: Logarithm of rent price			
	(1)	(2)	(3)	(4)
Discrimination dummy	-0.0315*** (0.002)	-0.0483*** (0.003)	-0.0506*** (0.005)	-0.0547*** (0.002)
Observations	130,179	125,191	125,192	125,194
Building FE	Yes			
Phone FE	Yes	Yes	Yes	Yes
Subdistrict FE		Yes		
District FE			Yes	
Day of posting FE	Yes	Yes	Yes	Yes
Controls (apartment char.)	Yes	Yes	Yes	Yes
Controls (building char.)		Yes	Yes	Yes

*Note:* Standard errors are clustered on the level of buildings, subdistricts and districts in specifications (1), (2) and (3) correspondingly. Standard errors in parenthesis. \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1

Table 1.B.4: Increased Search Time: Discrimination and Number of Days before Ad Removed

	Dependent variable: # of days before ad removed (log)			
	(1)	(2)	(3)	(4)
Discrimination dummy	0.1060*** (0.011)	0.1025*** (0.014)	0.0996*** (0.016)	0.1002*** (0.012)
Log total surface	0.1065*** (0.028)	0.1167*** (0.029)	0.1420*** (0.026)	0.1493*** (0.025)
LivingArea / TotalArea	-0.1014* (0.053)	-0.0025 (0.064)	-0.0188 (0.075)	-0.0225 (0.051)
Number of floors		-0.0027*** (0.001)	-0.0033** (0.001)	-0.0032*** (0.001)
Ground floor	0.0270 (0.020)	0.0376* (0.019)	0.0320** (0.013)	0.0319* (0.018)
Last floor	-0.0035 (0.017)	0.0231 (0.016)	0.0221* (0.011)	0.0224 (0.016)
Log dist. to center		-0.0506 (0.042)	0.0327 (0.042)	0.0035 (0.012)
Log dist. to metro		0.0399*** (0.009)	0.0502*** (0.012)	0.0543*** (0.006)
Log(number of photo + 1)	0.1239*** (0.006)	0.1292*** (0.007)	0.1293*** (0.007)	0.1288*** (0.006)
Log lenght of text (10 chars)	0.0253*** (0.005)	0.0267*** (0.006)	0.0295** (0.010)	0.0297*** (0.005)
Log price	0.6007*** (0.030)	0.5011*** (0.028)	0.4730*** (0.035)	0.4659*** (0.022)
Constant	-5.1956*** (0.251)	-4.0956*** (0.283)	-4.0736*** (0.423)	-3.9579*** (0.185)
Observations	116,278	112,497	112,498	112,498
Building FE	Yes	No	No	No
Subdisctrict FE	No	Yes	No	No
District FE	No	No	Yes	No
Day of posting FE	Yes	Yes	Yes	Yes
Controls (apartment char.)	Yes	Yes	Yes	Yes
Controls (building char.)		Yes	Yes	Yes

*Note:* The Sample consists of ads posted on the web-site during the observation period excluding ads that were available on the first and last days of the observations period. Standard errors are clustered on the level of buildings, subdistricts and districts in specifications (1), (2) and (3) correspondingly.

Standard errors in brackets.

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

Table 1.B.5: Heterogeneity of Search Time Effect: Interaction with Share of Discrimination in Neighborhood

	Dependent variable: Number of days in exposure (log)			
	(1)	(2)	(3)	(4)
Discrimination dummy	0.1060*** (0.011)	0.2455*** (0.017)	0.1090*** (0.036)	0.0768* (0.045)
Discrimination dummy × <i>Share of discrimination in building</i>		-0.5873*** (0.062)		
Discrimination dummy × <i>Share of discrimination in subdistrict</i>			-0.0122 (0.145)	
Discrimination dummy × <i>Share of discrimination in district</i>				0.1250 (0.186)
Observations	116,278	116,278	116,278	116,278
R-squared	0.396	0.397	0.396	0.396
Building FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

*Note:* The sample consists of ads posted on the web-site during the observation period. Standard errors are clustered on the level of buildings.

Standard errors in parenthesis. \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1

Table 1.B.6: Experiments Outcomes

Non slavic names	Slavic names					
	Answer back	Ask id	Is rented	Not read	Read, no answer	Total
Answer back	162	2	0	0	18	182
Ask id	12	1	0	0	3	16
Is rented	0	0	1	0	0	1
Not read	2	0	0	63	3	68
Read, no answer	77	1	3	4	142	227
Reject (due to ethnicity)	13	1	0	0	0	14
Reject (due to gender)	1	0	0	0	0	1
Total	267	5	4	67	166	509

## C. Appendix: Theory

### C.1. Tenants' problems

$$Emax\{\alpha - p_{nd}, V^{ns}\} = P(\alpha - p_{nd} > V^{ns}) \times E(\alpha - p_{nd}) + P(\alpha - p_{nd} < V^{ns}) \times V^{ns} =$$

$$\int_{V^{ns}+p_{nd}}^{\infty} f(\alpha) d\alpha \times E(\alpha - p_{nd}) + (1 - \int_{V^{ns}+p_{nd}}^{\infty} f(\alpha) d\alpha) \times V^{ns} =$$

$$\int_{V^{ns}+p_{nd}}^{\infty} (\alpha - p_{nd} - V^{ns}) f(\alpha) d\alpha + V^{ns}$$

$$V^{ns} - \theta V^{ns} = (1 - \theta) \left( \int_{V^{ns}}^{\infty} (\alpha - p_{nd} - V^{ns}) f(\alpha) d\alpha + V^{ns} \right) - k$$

$$\frac{k}{1 - \theta} = \int_{V^{ns}+p_{nd}}^{\infty} (\alpha - p_{nd} - V^{ns}) f(\alpha) d\alpha$$

$$Emax\{\alpha - p_i, V^s\} = \int_{V^s+p_i}^{\infty} (\alpha - p_i - V^s) f(\alpha) d\alpha + V^s$$

*Non-slavic* tenants' problem when  $\alpha$  is distributed uniformly:

$$\frac{k}{1 - \theta} = \int_{V^{ns}+p_{nd}}^{\beta} \frac{\alpha - p_{nd} - V^{ns}}{\beta} d\alpha = \frac{(\beta - p_{nd} - V^{ns})^2}{2\beta}$$

*Slavic* tenants' problem when  $\alpha$  is distributed uniformly:

$$2k\beta = \theta(\beta - p_d - V^s)^2 + (1 - \theta)(\beta - p_{nd} - V^s)^2$$

### C.2. Optimal Rents and Rent Differential in a Separate Neighborhood

Tenants problems can be rearranged such that (1.3) and (1.4) respectively become:

$$k = \theta \int_{V^s+p_s}^{\infty} (\alpha - p_d - V^s) f(\alpha) d\alpha + (1 - \theta) \int_{V^s+p_{nd}}^{\infty} (\alpha - p_{nd} - V^s) f(\alpha) d\alpha \quad (1.9)$$

$$\frac{k}{1 - \theta} = \int_{V^{ns}+p_{nd}}^{\infty} (\alpha - p_{nd} - V^{ns}) f(\alpha) d\alpha \quad (1.10)$$

Then assume that  $\alpha$  is drawn from uniform distribution on interval  $[0, \beta]$ . The equations can be rewritten as:

$$2k\beta = \theta(\beta - p_d - V^s)^2 + (1 - \theta)(\beta - p_{nd} - V^s)^2 \quad (1.11)$$

$$V^{ns} = \beta - p_{nd} - \sqrt{\frac{2\beta k}{1 - \theta}} \quad (1.12)$$

With  $\beta$  both mean and variance of  $\alpha$  increase. The parameter  $\beta$  can be interpret as likelihood of finding tenant who values the apartment highly.

First order conditions for landlords problems (1.5) and (1.6) respectively are :

$$p_d = \frac{1 - F(V^s + p_d)}{f(V^s + p_d)} \quad (1.13)$$

$$\pi(p_{nd} - p_{nd}F(V^s + p_{nd}) + (1 - \pi)(p_{nd} - p_{nd}F(V^{ns} + p_{nd})) = 0 \quad (1.14)$$

In the same way as in tenants' problems assumption on uniform distribution is imposed. Hence the equations appear as follows:

$$p_d = \frac{1}{2}(\beta - V^s) \quad (1.15)$$

$$p_{nd} = \frac{1}{2}(\beta - (\pi V^s + (1 - \pi)V^{ns})) \quad (1.16)$$

Four equations (first-order conditions of two tenants' and two landlords problems) contains four unknown variables: prices and reservation values. Therefore, together these equations define equilibrium. With simple rearrangements this system can be reduced to two equations that bind two prices: on discriminating and non-discriminating markets.

$$2k\beta = \theta p_d^2 + (1 - \theta)(2p_d - p_{nd})^2 \quad (1.17)$$

$$p_{nd} = \frac{1 - \pi}{1 + \pi} \sqrt{\frac{2\beta k}{1 - \theta}} + \frac{2\pi}{1 + \pi} p_d \quad (1.18)$$

### C.3. Equilibrium

The model can be defined with four equations:

$$\begin{cases} 2k\beta = \theta(\beta - p_d - V^s)^2 + (1 - \theta)(\beta - p_{nd} - V^s)^2 \\ V^{ns} = \beta - p_{nd} - \sqrt{\frac{2\beta k}{1-\theta}} \\ p_d = \frac{1}{2}(\beta - V^s) \\ p_{nd} = \frac{1}{2}(\beta - (\pi V^s + (1 - \pi)V^{ns})) \end{cases}$$

This can be reduced to the system of two equations that define optimal rent sums:

$$\begin{cases} 2k\beta = \theta(\beta - p_d - V^s)^2 + (1 - \theta)(\beta - p_{nd} - V^s)^2 \\ p_{nd} = \frac{1-\pi}{1+\pi} \sqrt{\frac{2\beta k}{1-\theta}} + \frac{2\pi}{1+\pi} p_d \end{cases}$$

The fact that rent differential is positive in optimum ( $p_{nd} - p_d > 0$ ) can be proved geometrically. The first equation is equation of ellipse sloped to the right, and the second equation defines straight line with slope that equals to  $\frac{2\pi}{1+\pi}$ . For any  $\pi$  this line is less steep than line  $p_{nd} = p_d$ . The point of intersection of ellipse and axis  $p_{nd}$  is  $\sqrt{\frac{2\beta k}{1-\theta}}$ , whereas the point of intersection of straight line given by second equation and axis  $p_{nd}$  is  $\sqrt{2\beta k}$ , which is less than  $\sqrt{\frac{2\beta k}{1-\theta}}$ . Therefore, for any values of parameters  $p_{nd} - p_d > 0$ .

# Chapter 2

## Urban Amenities and Tourism: Evidence from Tripadvisor

### Abstract

Using TripAdvisor reviews, we construct panel data on tourism and consumption in Paris. We document that during the pandemic a drop in tourism caused an increase in Parisians' satisfaction with restaurants and other amenities. Among three mechanisms — overcrowding, supply-side changes and aversion towards tourists — we only find support for the aversion mechanism. During the pandemic the word ‘tourist’ became less frequent in reviews, while other words relating to food quality, price and overcrowding stay on the same level. The improvement in ratings was stronger in restaurants popular among tourists from countries with a weaker social connection to France measured with Facebook connectedness index.

# 1. Introduction

“Are there too many tourists in Paris?” – was a title of the conference organised by the city hall of Paris on June 24, 2019. While the speakers of the conference agreed that overtourism in Paris has not yet reached the same scale as in Amsterdam or Barcelona, they also admitted that “rapid and poorly regulated growth” of tourism can be harmful to the city<sup>1</sup>. There were reasons for concern. The number of foreign tourists to France has more than doubled over the previous 15 years. In 2019 France was the most visited country in the world, and Paris was the third most visited city. During the year 35.4 million tourist stayed in the city’s hotels, which is approximately 16 times more than the population of Paris.

In the years preceding the pandemic, concerns about tourism have became common in Europe.<sup>2</sup> Anti-tourist protests took place in Barcelona, San Sebastián, Mallorca, Venice and other European cities. Anti-tourist graffities, typically saying “tourist go home”, were spreading across the cities including Paris.

However, during the summer of 2020, there were no crowds of tourists in Paris. The problem of overtourism raised at the city hall conference faded into the background, when the COVID-19 pandemic and the stringency measures, imposed by the governments, disrupted tourist inflows, causing, as was coined by the World Tourism Organization, “the worst year in tourism history”.

It is still unclear what the tourism industry will face after the pandemic: whether it will continue to grow at the pre-pandemic rate, slow down or start to shrink. While the industry is on hold, the questions posed by researchers and policy-makers before the pandemic remain relevant and open. What is an optimal level of tourism? What are its costs and benefits? At the same time, the unexpected shock in tourism created a proper setting to explore the question: “What would life be for residents of Paris if there were no tourists?” In fact, during the summer of 2020 Parisians were not bothered by an excess of tourists, while restaurants and other urban amenities remained accessible, and COVID-19 cases and deaths were relatively low, as the first pandemic wave was fading out. In addition, restaurants were kept open artificially through heavy government subsidies, providing a unique setting to study demand-related factors without an endogenous adjustment of supply.

In this paper we estimate the effect of tourism on residents’ satisfaction with restaurants and other urban amenities. We use data on restaurant reviews from Tripadvisor – the

---

<sup>1</sup>See CNews. The World Tourism Organization (UNWTO) defines overtourism as “the impact of tourism on a destination, or parts thereof, that excessively influences perceived quality of life of citizens and/or quality of visitor experiences in a negative way” Carvão et al. (2018). For a review on overtourism from the tourism management literature see Capocchi et al. (2019).

<sup>2</sup>See the Guardian

platform that aggregates user-generated content on restaurant and other travel experiences. We construct unique panel data on consumption and amenities in the city. This data allows us to achieve multiple goals at the same time.

First, we use it to produce a highly granular measure of tourism. The share of non-French among all reviews serves as a close proxy of tourists' presence, which we validate using several other measures. The benefit of this measure is that it can be defined on a very granular level, the restaurant itself. In addition, while many studies focus on the location where tourists stay overnight to study the impact, the measure used here allows to study the location of where tourists consume.

Second, the review data and the ratings given by locals can be used as an indicator of locals' satisfaction with restaurant experience. More generally, it serves as a measure of satisfaction with urban amenities, which varies across space and time. The literature shows that this indicator is meaningful: For example, Kuang (2017) finds that restaurant ratings are highly correlated with real estate prices.

We match restaurant data with another source of information on residents' quality of life: number of complaints on the crowd-sourced platform *DansMaRue*. The platform is provided by the city hall of Paris. Users can report any problem related to public space (abandoned waste, tags, wild posting, etc.) through the mobile application or the web-site. Then the city administration analyses the reports and try to solve the problems. We treat this disamenity measure as another outcome relevant to our study.

We first document two stylized facts. First, more touristic restaurants receive lower ratings by locals in the cross-section, suggesting a potential disamenity stemming from tourist demand. Second, touristic neighborhoods have a lower variety of amenities which may indicate that tourists value variety less than locals do.

Using the pandemic as a source of exogenous variation in international tourist arrivals, we find that the drop in tourism caused an increase in residents' satisfaction with urban amenities, both in terms of restaurant ratings and a decreased number of complaints on *DansMaRue*. In particular, the average restaurant increases its rating by close to 10% of a standard deviation in the absence of tourists and the number of complaints in the direct vicinity of the average restaurant decreases by at least 8%.

Importantly, our effect is not unique to the lockdown-induced tourism decline. We find similar evidence when using the terrorist attacks that took place in November 2015. Our results are also robust to using measures of tourism that are based on the self-declared location of users rather than language.

Next, we consider three potential mechanisms driving our findings: overcrowding, supply-side change and residents' aversion towards tourism. Our analysis only finds support for the

aversion mechanism. First, we find that the number of reviews explicitly mentioning tourism (which are often negative) declines. Second, relying on a proxy of social connectedness between countries derived from Facebook data, we find that restaurants with a clientele that has little connections to France sees a larger increase in its rating post-lockdown. This suggests that Parisians are less bothered by tourists from countries with which they have strong social ties.

This study is most closely related to a growing literature studying the interaction of tourism and local amenities. Allen et al. (2020) study the effects of tourism on residents' welfare in Barcelona. Building on a quantitative spatial model and credit card expenditure data, they derive the incidence of tourism on locals' welfare and find a largely heterogeneous impact which negatively affects those living in the center, while resulting in welfare gains for those living in less central parts of the city. While they are able to quantify the welfare effects of tourism, our paper focuses on how tourism affects the reported satisfaction with the quality of specific amenities and highlights the channels through which tourism operates.

This paper is also related to the literature on endogenous amenities. In contrast to historical sites and natural landmarks, endogenous amenities such as restaurants and bars are reactive to demand. In particular, Almagro and Dominguez-Iino (2019) study how amenities and location sorting by residents endogenously adjust to a large increase in tourist demand, focusing on the city of Amsterdam. Relative to their paper, we focus on relatively short-term effects where amenities and residence location are essentially fixed.<sup>3</sup>

More generally, our paper builds on the literature emphasizing the importance of amenities. In their seminal paper Glaeser et al. (2001) explore the role of cities as centres of consumption. They show that high-amenity cities have been growing faster than low-amenity cities, highlighting the importance of amenities for location choices. Generally, on the importance of urban amenities for attracting residents see also Carlino and Saiz (2019), Lee (2010) and Couture and Handbury (2020).

This paper is not the first to use data on restaurant reviews to study urban amenities. Kuang (2017) argues that quality of urban amenities are important for city residents, which is revealed in real estate prices. She measures the quality of amenities using restaurant ratings posted by users on Yelp.

It is worth noting that tourism can have a substantial positive economic impact, and tourism suspension causes deep economic damage to the local economy (see e.g. Faber and Gaubert (2019)). This paper does not focus on the direct effects of tourism on the local economy, but rather its impact on local amenities.

Finally, this paper belongs to the growing and diverse literature on the COVID-19 pan-

---

<sup>3</sup>The government was essentially freezing the local economy through heavy subsidies.

demic and its interaction with the urban structure Gupta et al. (2021); Althoff et al. (2020); De Fraja et al. (2020); Miyauchi et al. (2021); Couture et al. (2021); Gupta et al. (2020); Coven et al. (2020).

## 2. Background and Data

In this section we first discuss how in the summer of 2020 the Covid-19 pandemic led to a sharp drop of tourists coming to Paris, while there were few restrictions for locals. Next, we discuss our main dataset on restaurant reviews that were collected from the website *Tripadvisor* and additional datasets from other sources that we use.

### 2.1. COVID-19 in Paris

The first restrictions related to Covid-19 took effect in early 2020. On March 12, Emmanuel Macron announced in a televised address that all schools and universities across France would be closed. On March 13, 2020, Prime Minister Edouard Philippe announced the closure of all pubs, restaurants, cinemas and nightclubs. After three months of strict lockdown measures, on June 14, cafes, restaurants and pubs reopened in Paris.

While the restaurant sector returned to normality, tourism remained heavily affected by the global pandemic. The Ile-de-France region which encompasses Paris was especially heavily hit. Relative to July 2019, it saw a drop of 70.8% in overnight stays in its hotels in July 2020<sup>4</sup>. The following months saw a similar drop in demand in the hospitality sector. This drop was especially pronounced among tourists not residing in France. Compared to 2019, France saw 71.8% less non-residents in overnight stays in 2020, whereas overnight stays by residents declined only by 10.5%. To summarize, Paris saw a large drop in tourism in the summer of 2020 which was mainly concentrated in international arrivals.

### 2.2. Tripadvisor Data

*Tripadvisor* is a user-generated social media review site, which publishes user reviews on restaurants, hotels and other attractions. We collected data on all Parisian restaurants that were listed on the site on November 17, 2020.<sup>5</sup> We obtained information on restaurant characteristics, such as the type of cuisine and the address, and individual review data, including the review's date, text, language, user, user location and rating. We geocode

---

<sup>4</sup>See INSEE FOCUS No. 235 here <https://www.insee.fr/fr/statistiques/5369851#consulter>

<sup>5</sup>In this analysis, we restrict ourselves to restaurants located in Paris *intra-muros* – the city of Paris that consists of 20 municipal arrondissements and excludes the surrounding Greater Paris area.

restaurants' addresses. We leverage the data on review's language and user location to separate consumption of residents and tourists. As a result we construct unique and highly detailed panel that reflects city's restaurant consumption across space and time.

Figure 2.1 presents the daily number of reviews of the roughly 15,000 Parisian restaurants, cafes and bars left on the platform since its launch. The time trends are represented by smoothing splines. Reviews are split into two categories: reviews written in French and written in other languages. The figure shows both the process of technology adoption and the fluctuations in restaurant consumption. French users began adapting the platform in 2007, and their usage peaked in 2017.

Figure 2.2 zooms in the same time series to a period starting from 2018 when the platform's penetration is relatively stable. The beginning and the end of the "first-wave" lockdown imposed by the French government are marked with a blue dotted line. During the lockdown both French and non-French reviews dropped to near zero. Then, starting in June, French reviews revived, but foreign reviews remained on a negligible level. The observational period ends with both French and non-French review numbers going back to zero due to the introduction of a second wave of restrictions. As a whole, these figures demonstrate that the review data allows us to differentiate between demand by residents and tourists.

### 2.3. Measuring Tourism

In this paper we use review data to construct a highly granular measure of tourism at the restaurant level. Importantly, it gives us an indicator of where tourists consume in the city rather than where they stay. Our preferred proxy of tourism is constructed as a share of reviews written in languages other than French. In Section B.2 in the Appendix we repeat our analysis using an alternative measure of tourism based on users' home locations.

The Figure 2.3 shows a map of our tourism measure. A lighter color indicates a higher share of non-French reviews. As expected, restaurants with the highest levels of tourism are located in the areas known for Paris' major attractions: the Eiffel tower, Montmartre, Notre-Dame de Paris and the Arc de Triomphe.

To validate our proxy for tourism more formally, we use data from the *Enquêtes de fréquentation des sites culturels* provided by the *Observatoire économique du tourisme parisien* (Observatory of the Parisian tourism economy). This survey contains the share among all tourists coming to Paris visiting different tourist attractions. We consider tourists visiting from 2015 to 2019 and geocode the 18 attractions that are located intra-muros contained in the survey. Then, we construct a measure for demand by tourists that follows the market access framework widely used in the economic geography literature:

$$\text{Tourist Access}_i = \sum_j \frac{\text{Visitors}_j}{\text{Distance}_{ij}}$$

Note that we are implicitly assuming a distance elasticity of tourist consumption trips of -1. While we are not aware of a paper estimating this parameter specifically for demand by tourists, Miyauchi et al. (2021) look at the distance elasticity of location choice for consumption trips. They find a value of -1.09 and thus close to -1.

Next, we correlate our tourism proxy with the tourist demand measure. As Figure Figure 2.5 shows, we find a strong positive correlation between the two (the  $R^2$  of a linear regression is 0.19). The correlation is robust to controlling for quartier fixed effects, meaning that, even after controlling for a relatively fine-grained spatial unit, the remaining variation in our tourism proxy is correlated with tourist access (see Table Table 2.C.1). Together, this shows that our proxy for tourism correlates strongly with other, external measures of tourism.

Finally, to further corroborate our proxy for tourism, we rely on user location information. In particular, we compute the share of users by restaurant who indicate a location in a country other than France. As figure Figure 2.A.2 shows, the two measures are highly correlated (the  $R^2$  of a linear regression is around 0.77).

#### 2.4. Content of Reviews

We perform text analysis of reviews to better understand users' concerns. We distinguish five topics that are relevant to the mechanisms we want to test for: discussion on tourism, concerns about low food quality, high price, long waiting time and noisy environment.

The mapping of the review texts to topics is determined by manually constructed dictionaries. The procedure of constructing the dictionary is the following. First, we examined around one thousand randomly selected reviews to find a sample of words that relates to the topic in a non-ambiguous way. Second, we validate these terms searching for counter-examples in the corpus – the “false-positives”, the reviews where these terms are mentioned, but in fact these reviews are not related to the topic. Third, we extend our dictionary with common misspellings of the selected terms. We also take partial forms of the words. Lastly, we create a list of ‘minus’ phrases, so that wordings such as “*pas cher*” (not expensive) will not be flagged as “*cher*” (expensive).

Overall , our approach minimises *false positives* (the probability that the text is attributed to the topic, when in fact it is not related to the topic), but is does not minimise *false negatives* (the probability that the text is not attributed to the topic, when in fact it is related to the topic). The short version (without misspellings and versions) of our dictionary is presented in

the Table 2.D.1. The summary statistic of topics is presented in the Table 2.D.2. Notably, all topics occur with relatively similar frequency (between 2% and 6%) and thus allow a meaningful comparison.

## 2.5. *Dans Ma Rue*

Most of our analysis is based on the TripAdvisor data. To externally validate that the presence of tourists affects locals' satisfaction with amenities, we draw on an additional dataset from the application *Dans Ma Rue* created by the Municipality of Paris. With the help of this application, citizens can register and geolocalise 'anomalies' observed in public space in Paris.<sup>6</sup> Users upload the complaints directly from their smartphones, specifying the location, date and the subject. The aim of the application is to improve the quality of Parisian public space by giving access of user-generated data on 'anomalies' to municipal service. The application was launched in 2012. For our analysis we focus on complaints about commercial activity which is the category most related to restaurant activity.

The high resolution of the data allows us to only consider complaints that are possibly related to a particular restaurant. We assign complaints to a given restaurant within a 100m radius.

## 2.6. *Social Connectedness Index*

Below we want to test whether the origin of tourists has an impact on locals' perception of them. To proxy for cultural proximity between foreign countries and France we rely on the Social Connectedness Index (SCI) published by Facebook.<sup>7</sup> It is based on the number of Facebook friendships between users located in a pair of countries. More precisely, it is computed as

$$\text{Social Connectedness}_{ij} = \frac{\text{FB Friends}_{ij}}{\text{FB Users}_i \times \text{FB Users}_j}$$

where  $\text{FB Friends}_{ij}$  are the number of friendships between users residing in countries  $i$  and  $j$  and  $\text{FB Users}_i$  the number of users in country  $i$ . For further details on the methodology see Bailey et al. (2018). Relying again on the information on users' origin, we compute the average social connectedness between the French population and the non-French customers of a particular restaurant.

---

<sup>6</sup>The set of potential 'anomalies' includes overflowing litter bins, illegal graffiti, abandoned objects, road damage and many others.

<sup>7</sup>The version we use dates from October 2021.

### 3. Stylized Facts

This section presents stylized facts about the geography of tourism in Paris.

*More touristic restaurants receive lower ratings.*

To compare the perceived value of more and less touristic places, we run the following regression at the review level

$$\text{Rating}_{rij} = \beta \text{Tourism}_j + X_j + \gamma_i + \epsilon_{rij} \quad (2.1)$$

where  $\text{Rating}_{rij}$  is the rating given by user  $i$  for restaurant  $j$  in review  $r$ . Our variable of interest is  $\text{Tourism}_j$  which is a measure of how touristic restaurant  $j$  is. We add other controls at the restaurant level ( $X_j$ ) and control for user-level fixed effects ( $\gamma_i$ ). This means we are comparing different reviews made by the same user, controlling for all unobservables at the level of the user. We also estimate a variation of this specification with quartier fixed effects. This captures any geographic amenity shifter, e.g. restaurants located along the river Seine receiving systematically higher ratings because of a nice view. We cluster standard errors at the restaurant level.

Table 2.1 displays the results of estimating equation Eq. (2.1). The estimation is based on pre-Covid data in order to avoid any confounding effects. We estimate the regression separately for Parisians only, since we are interested in the value of amenities for the local population. We find that overall more touristic places receive lower ratings ( $\beta < 0$ ), after controlling for the (log) number of reviews received by the restaurant and for user and grid cell fixed effects. Using the most stringent specification with quartier-level fixed effects in column 3, we find that an increase in tourism demand by one standard deviation is associated with a rating that is around 2% lower<sup>8</sup>.

*More touristic neighborhoods have less diverse restaurants*

While more touristic venues seem to receive lower ratings, we also find that tourism systematically correlates with other characteristics of neighborhood amenities. We start from the idea that tourists often visit foreign places to get an impression of the local culture. Thus, local businesses may cater to this demand by offering a version of French culture that is particularly appealing to tourists. Indeed we find that the share of restaurants offering French cuisine is much higher than in neighborhoods more dominated by locals (see Figure 2.6).

---

<sup>8</sup>The standard deviation of tourism intensity is around 0.125 and the mean rating is around 3.82

To capture diversity more broadly, we compute the market share of each cuisine type (weighted by the number of reviews). We then compute the Herfindahl index and show that more touristic areas have a systematically more concentrated market for restaurants (see Figure 2.7). This illustrates that tourism is associated with a less diverse set of amenities.

## 4. Empirical Strategy

We employ a standard difference-in-differences framework at two different levels of aggregation to study the impact of the absence of tourists on locals' valuation of amenities. First, a restaurant-level approach gives us a broad picture of whether more and less touristic venues evolved differently over time. Second, review-level regressions allow us to assess whether the same users evaluated initially more touristic restaurant differently when borders were closed.

### 4.1. Restaurant-level Approach

At the restaurant level, we use the following specification

$$Y_{jt} = \beta \times \text{Post-Lockdown}_t \times \text{Tourism}_j + \gamma_j + \delta_t + \epsilon_{jt} \quad (2.2)$$

where  $Y_{jt}$  is an outcome of restaurant  $j$  in month  $t$ .  $\text{Post-Lockdown}_t$  is a binary variable indicating whether month  $t$  belongs to the post-lockdown period.  $\text{Tourism}_j$  measures to what extent restaurant  $j$  is frequented by tourists. We include restaurant fixed effects ( $\gamma_j$ ) and month fixed effects ( $\delta_t$ ). In a more stringent variation of this specification we also include quartier-time fixed effects. This controls for any unobserved time-varying factors at the neighborhood level, such as an increased share of remote working that may affect residential neighborhoods differently than the business district. Standard errors are clustered at the quartier level.

Below we will focus on one main outcome. We look at the average rating that restaurant  $j$  receives in month  $t$ , only looking at reviews by local residents. Our hypothesis is that tourism lowers the utility locals derive from amenities (visiting a restaurant in our case). We thus expect  $\beta > 0$ .

### 4.2. Review-level Approach

At the review level, we use the following specification

$$Y_{ijt} = \beta \times \text{Post-Lockdown}_t \times \text{Tourism}_j + \gamma_j + \delta_t + \mu_i + \epsilon_{ijt} \quad (2.3)$$

where  $Y_{ijt}$  is a rating by user  $i$  for restaurant  $j$  in month  $t$ . As above,  $\text{Post-Lockdown}_t$  is a binary variable indicating whether month  $t$  belongs to the post-lockdown period.  $\text{Tourism}_j$  measures to what extent restaurant  $j$  is frequented by tourists. In addition to restaurant and month fixed effects ( $\gamma_j, \delta_t$ ), we also include user fixed effects, relying on within-user changes pre- to post-lockdown. Again, we cluster standard errors at the quartier level.

While including user fixed effects is already restrictive, identification can still come from comparing the magnitude of within-user changes across users, depending on whether they visited a touristic restaurant or not. If e.g. an increased life satisfaction post-lockdown and the propensity to visit more touristic restaurants were both determined by an unobserved third factor, our findings would be spurious. We thus, in a final step, interact user fixed effects with a post-lockdown dummy. This restricts identification to users who review at least two restaurants either before or after the lockdown. Intuitively, this specification captures whether the penalty for more touristic places decreased after the lockdown relying only on different ratings for more or less touristic restaurants by a user in the same period.

Our parameter of interest is  $\beta$ . Our hypothesis is that tourism is bad for locals' utility derived from a restaurant visit. Hence, we should observe that post-lockdown, when restaurants were open, but tourists were not present, initially touristic places start receiving higher ratings ( $\beta > 0$ ).

## 5. Results

Table 2.2 shows the results of estimating equation Eq. (2.2) using the average monthly rating by Parisians at the restaurant level as the outcome variable.<sup>9</sup> We find that initially more touristic venues receive higher ratings when tourists are no longer around. Importantly, the effect is not driven by neighborhood-level trends as including quartier-time fixed effects only marginally changes the coefficient.

The magnitude of the coefficient can be best understood when considering the average tourism share of around 31.6%. The estimate in column 2 then implies that in Paris without tourists, which comes close to the reality of the post-lockdown summer, locals rate the average restaurant around 0.1 (or around 8% of a standard deviation) higher. At the 90th percentile of the tourism share this estimate more than doubles to around .22 (or around 17% of a standard deviation).

---

<sup>9</sup>Note that the sample is thus constrained to restaurants that receive at least one rating by a Parisian in a given month.

Table 2.3 shows the results of a user-level estimation (see equation Eq. (2.3)). Importantly, this econometric approach allows us to exploit within-user changes in behavior while holding fixed time-invariant characteristics, such as preferences for certain types of neighborhoods or restaurant types. We confirm our results at the user level, i.e. Parisians rate their experience higher in places previously frequented by many reviewers not from Paris. The coefficient is of similar magnitude as at the restaurant level.

## 6. Robustness & Further Results

In this section we first present results using the data on neighborhood complaints as a different measure of disamenities. Then, we show that our result is not specific to the pandemic-induced shock to tourism, not driven by pre-trends, not affected by spillovers and present minor robustness exercises such as different levels of clustering.

### 6.1. Neighborhood Complaints

So far we have focused only on data coming from *Tripadvisor*. To provide further evidence that the lower influx of tourists improved locals' perceived satisfaction with local amenities, we analyze data on complaints registered within 100m of the restaurants in our sample by local residents (see section 2.2 for a detailed description). The goal of this exercise to show that tourism not only affects people going to restaurants but also local residents.

We estimate equation Eq. (2.2), replacing the average rating of the restaurant with the number of complaints in the vicinity of a restaurant within a given month. As this is a count variable which contains zeros, we use a Poisson model to estimate this equation.

Table 2.4 presents the results. We find that complaints around touristic restaurants decline relative to less touristic ones. Using the most conservative estimate in column 2, complaints around a restaurant with an average share of tourists among its customers decrease by around 8%.<sup>10</sup>

The positive impact of a decrease in the arrival of tourists is thus not only reflected in restaurant ratings, but also confirmed by an entirely external data source, namely crowd-sourced complaints that are used to improve municipal services.

### 6.2. Bataclan Attacks

We exploit the Covid-19 pandemic as an exogenous shock to tourism. However, the pandemic also affected the mobility of residents and thus the spatial mobility patterns in

---

<sup>10</sup>We use the average tourism share of 31.6% and multiply it with the coefficient in column 2.

the city. While there were little restrictions in place during the summer of 2020, some people continued to work from home. In the empirical analysis above we control for trends that happen at the level of neighborhoods. Thus, a general shift in where the working population consumes is controlled for. In addition, we present results at the user level, thereby abstracting from compositional changes in the restaurants' visitors.

Still the pandemic may have affected restaurants in ways that are unobservable to us and correlated with our measures of tourism. For example, restaurants with larger outdoor facilities may have benefited most after the lockdown was lifted, as people continued to be cautious because of the risk to get infected. If the availability of outdoor facilities is correlated with our measure of tourism, we are wrongly attributing the observed changes in ratings and demand to tourism.

To alleviate concerns related to the specific nature of the pandemic, we instead use the terrorist attacks that took place on November 13, 2015 as an exogenous shock to tourism. Three groups launched a total of six attacks that day in Paris, killing 130 people. These gruesome attacks shocked France and were widely covered in the international press. In the months that followed, Paris saw a strong decline in tourism. Occupancy rates were down by 13.1% in the three months following the attacks compared to the same period in the year before.<sup>11</sup>

Table 2.B.1 display the results of estimating equation Eq. (2.2) using reviews from January 2015 to June 2016 and defining tourism intensity based on data from 2014. November 2015 is dropped from the analysis and December 2015 onwards is defined as post-Bataclan. We find that initially more touristic restaurants received better ratings by Parisians after the November attacks. Compared to Table 2.2, the coefficient is substantially smaller which is in line with a lower drop in tourism arrivals than during the summer of 2020. Overall, this very different natural experiment lends support to our hypothesis that tourism negatively affects the quality of amenities as perceived by locals. This does not seem to be driven by factors specific to the pandemic.

In addition, the November attacks allow us to look at the reaction of reviewers that are not from Paris. Interestingly, there is no effect on their ratings of touristic places. This suggests that the externalities caused by tourism specifically affect locals.

### 6.3. Pre-Trends

In order to asses the timing of the effect that we find, we estimate equation Eq. (2.2) allowing for  $\beta$  to be time-varying. In particular, we estimate one coefficient per quarter

---

<sup>11</sup>See <https://www.costar.com/article/724916287> for reporting on the impact of terrorist attacks on hotel occupancy rates.

and set the first quarter of 2020 as reference group. If the effect is driven by the sudden and unexpected absence of tourists due to the pandemic, we should observe no differential trends for more touristic restaurants prior to the outbreak of Covid-19. Figure 2.8 plots the estimated coefficients along with 90% confidence intervals. The figure shows that prior to the Covid-19 outbreak coefficients are close to and not statistically different from zero. Then, in Q3 and Q4 of 2020 coefficients are positive and statistically different from zero. This lends further support to the interpretation that Covid-19 led to a shift in locals' ratings of touristic venues.

#### 6.4. Spillovers

The analysis is focused on tourists visiting a particular restaurant. We thus far have not tested if this effects spills over to restaurants located close by. In this case the effect of tourism would be further amplified. We thus include in our baseline specification, equation Eq. (2.2), measures of many tourists visit restaurants in the surrounding area. As Table 2.8 shows, using different distances, we do not find strong evidence for that. The impact of a reduced influx of tourists seems to be mostly limited to the restaurant itself.

#### 6.5. Further Robustness Checks

In order to lend further credibility to our main result we perform several robustness exercises. First, we report our main result clustering standard errors at different levels. As Table 2.B.6 shows, clustering at the quartier level as done throughout our analysis is on the conservative side. Second, we use different measures of tourism. In Table 2.B.5 we vary the period over which we compute the initial tourism share. Again, our results are robust to these different permutations. Third, we use the share of reviews left by non-Parisians instead of the share of reviews not written in French. As Table 2.B.2 illustrates, using this different proxy results in a qualitatively similar effect.<sup>12</sup>

## 7. Mechanisms

To get at the mechanism, we use two different approaches. First, we use the text-based classification of reviews described in section 2.2. In particular, we estimate the following equation

---

<sup>12</sup>Note that this measure likely also captures domestic tourism. Since travel restrictions mainly applied to international visitors, we focus on the share of non-French reviews below.

$$\text{Share Reviews}_{jt} = \beta \times \text{Post-Lockdown}_t \times \text{Tourism}_j + \gamma_j + \delta_t + \epsilon_{jt} \quad (2.4)$$

where  $\text{Share Reviews}_{jt}$  is the share of reviews of restaurant  $j$  in month  $t$  referring to a particular type of topic, such as overcrowding.<sup>13</sup> The rest of the specification is as described in section 4. We also estimate a review-level version of this specification. The results are displayed in Table 2.5.

Second, we split the coefficient on the tourism-post interaction by variables defined at the restaurant level. This allows us to see if the effect is driven by certain types of restaurants.

Below, we will discuss three main mechanisms: overcrowding, supply-side changes and a direct aversion against the presence of tourists.

### 7.1. Overcrowding

A long waiting time and a noisy environment are distinctive features of overcrowding. Congestion caused by tourists should lead to an increase of frequencies of these topics. As Table 2.5 shows, we find no evidence pointing in this direction. More touristic restaurants did not receive relatively less reviews mentioning a long wait or noise after the lockdown. We interpret this as congestion not being a major driver of our results.

### 7.2. Supply-Side Changes

Low quality of food can be associated with the supply-side mechanism. According to this mechanism, restaurants change their technology when they are oriented to the tourist market – automatize the production, but also decrease the quality perceived by residents, since in this case the restaurants face lower incentives to provide consistent quality (tourists are not repeat consumers). This tendency should reflect in reviews left by residents. A similar logic can be applied to the concerns of too high prices. When consumers say that the price is too high, it likely means that price does not correspond to the perceived quality of the product.

### 7.3. Aversion

Another driver of our results could just be a direct, taste-based aversion of locals against tourists, closely linked and probably not distinguishable of xenophobia. As Table 2.5 shows, the only reviews that explicitly mention tourists appear significantly less after the lockdown

---

<sup>13</sup>Similarly, we estimate equation Eq. (2.3) with a dummy as dependent variable indicating whether a topic is mentioned in the review or not.

in initially touristic places. This suggests that it is something about the presence of tourists themselves rather than perceived overcrowding or decreases in quality.

To further test whether a direct aversion against the presence of tourists is at play, we test whether the increase in ratings is higher when the tourists are socially more distant to the local population. In particular, we exploit the information on users' origin provided in their profile. This allows us to compute for each restaurant the share of reviewers from a given country of origin. We combine this with the Social Connectedness Index (SCI) to compute the average SCI between restaurants' foreign reviewers and France.<sup>14</sup>

If Parisians have a distaste for foreigners from less familiar countries, we should see a higher increase in satisfaction for restaurants with many visitors from these countries. We thus estimate the treatment effect separately for restaurants with above and below-median SCI value. Table 2.6 shows that the increase in ratings of touristic places is indeed driven by low-SCI restaurants. For example, in column 4, the treatment effect for high-SCI is close to and not statistically different from zero. The coefficient for low-SCI places on the other hand suggests that touristic, low-SCI restaurants increased their average rating by around 0.13. This evidence is thus consistent with homophily among locals.

One concern might be that social connectedness is correlated with actual tourist arrivals from a country during the post-lockdown summer. However, the nature of the shock is such that arrivals from all countries drop to almost zero. Identification is thus almost entirely based on the pre-Covid exposure to tourism. In unreported results we control for differential changes in demand by nationality using a Bartik-style shock and find almost no change in our estimates.

## 8. Conclusion

This paper studies the impact of tourism on urban amenities. Exploiting a large decline in international travel during the COVID-19 pandemic, we find that tourism decreases the perceived quality of restaurants among locals. We find suggestive evidence that the negative effect of tourism operates through direct aversion against the presence of tourists, rather than overcrowding or supply-side changes. The effect is concentrated in restaurants where the tourist clientele was from countries that have few social ties with the French population.

This paper contributes to an emerging literature on the effects of tourism on locals' welfare. While the existing literature emphasizes price channels, i.e. tourists driving up prices Allen et al. (2020) and endogenous adjustment of amenities Almagro and Dominguez-Iino (2019), we show that tourism has an additional effect on existing amenities which

---

<sup>14</sup>See section 2.6 for a description of the SCI.

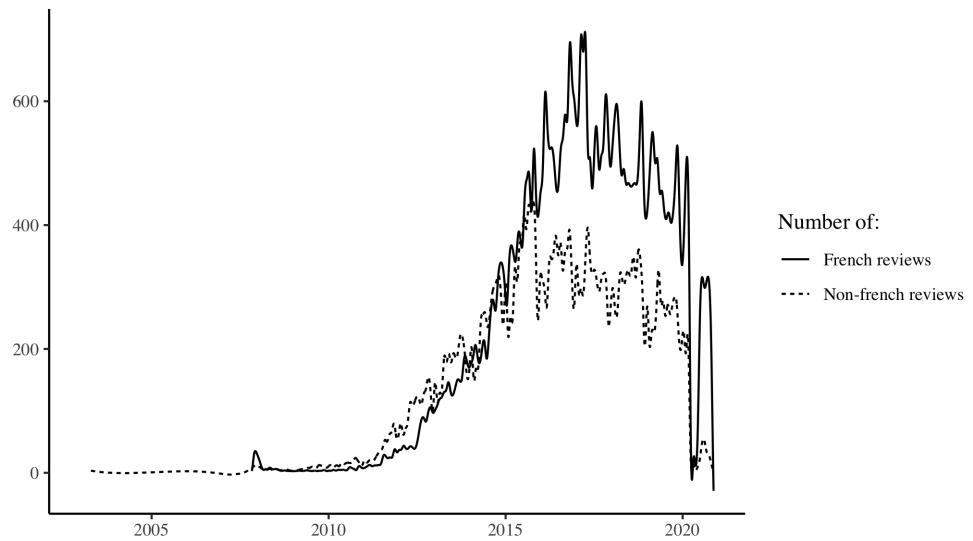
lowers their experienced quality. While we do not aim to evaluate the overall welfare impact of tourism in this paper, we highlight an additional source of discontent that can be caused by tourism. This adds to the debate preceding the pandemic on limiting tourism inflows in some of the most popular tourist destinations. It remains an open question whether tourism will rebound to its pre-pandemic levels. If it does not, our paper provides a preview how persistently lower inflows may affect locals' quality of life.

# References

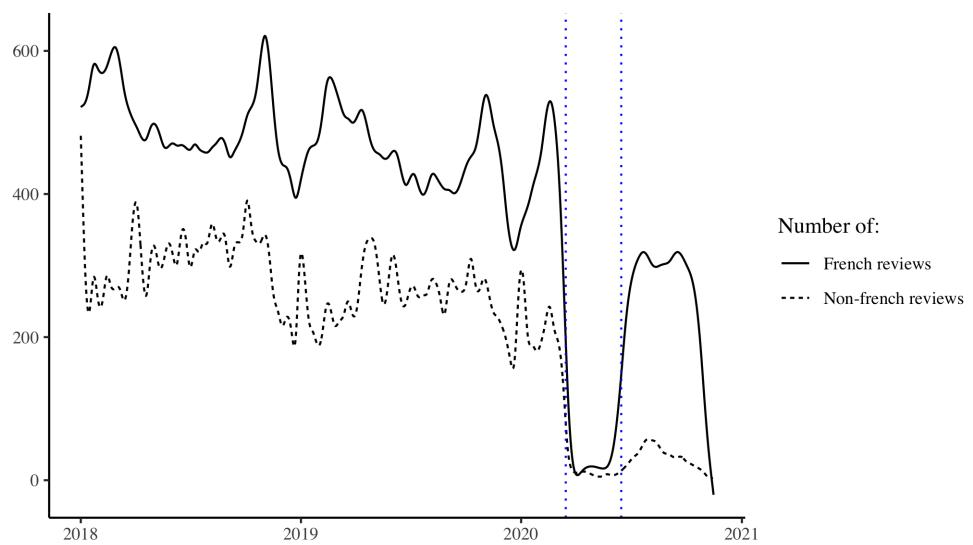
- Allen, T., Fuchs, S., Ganapati, S., Graziano, A., Madera, R., and Montoriol-Garriga, J. (2020). Is tourism good for locals? evidence from barcelona.
- Almagro, M. and Dominguez-Iino, T. (2019). Location sorting and endogenous amenities: Evidence from amsterdam. Technical report, Working Paper.
- Althoff, L., Eckert, F., Ganapati, S., and Walsh, C. (2020). The city paradox: Skilled services and remote work.
- Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–80.
- Capocchi, A., Vallone, C., Pierotti, M., and Amaduzzi, A. (2019). Overtourism: A literature review to assess implications and future perspectives. *Sustainability*, 11(12):3303.
- Carlino, G. A. and Saiz, A. (2019). Beautiful city: Leisure amenities and urban growth. *Journal of Regional Science*, 59(3):369–408.
- Carvão, S., Koens, K., and Postma, A. (2018). Presentation of unwto report ‘overtourism? understanding and managing urban tourism growth beyond perceptions’.
- Couture, V., Dingel, J. I., Green, A., Handbury, J., and Williams, K. R. (2021). Jue insight: Measuring movement and social contact with smartphone data: a real-time application to covid-19. *Journal of Urban Economics*, page 103328.
- Couture, V. and Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, 119:103267.
- Coven, J., Gupta, A., and Yao, I. (2020). Urban flight seeded the covid-19 pandemic across the united states. *Available at SSRN 3711737*.
- De Fraja, G., Matheson, J., and Rockey, J. (2020). Zoomshock: The geography and local labour market consequences of working from home. *Available at SSRN 3752977*.

- Faber, B. and Gaubert, C. (2019). Tourism and economic development: Evidence from mexico's coastline. *American Economic Review*, 109(6):2245–93.
- Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer city. *Journal of economic geography*, 1(1):27–50.
- Gupta, A., Mittal, V., Peeters, J., and Van Nieuwerburgh, S. (2021). Flattening the curve: Pandemic-induced revaluation of urban real estate. Technical report, National Bureau of Economic Research.
- Gupta, A., Van Nieuwerburgh, S., and Kontokosta, C. (2020). Take the q train: Value capture of public infrastructure projects. Technical report, National Bureau of Economic Research.
- Kuang, C. (2017). Does quality matter in local consumption amenities? an empirical investigation with yelp. *Journal of Urban Economics*, 100:1–18.
- Lee, S. (2010). Ability sorting and consumer city. *Journal of urban Economics*, 68(1):20–33.
- Miyauchi, Y., Nakajima, K., and Redding, S. J. (2021). Consumption access and agglomeration: evidence from smartphone data. Technical report, National Bureau of Economic Research.

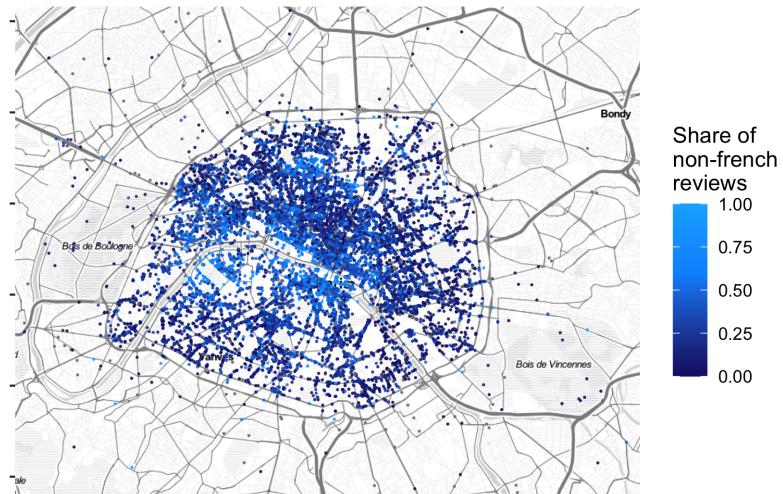
Daily Number of Reviews in Paris (since launch of Tripadvisor)



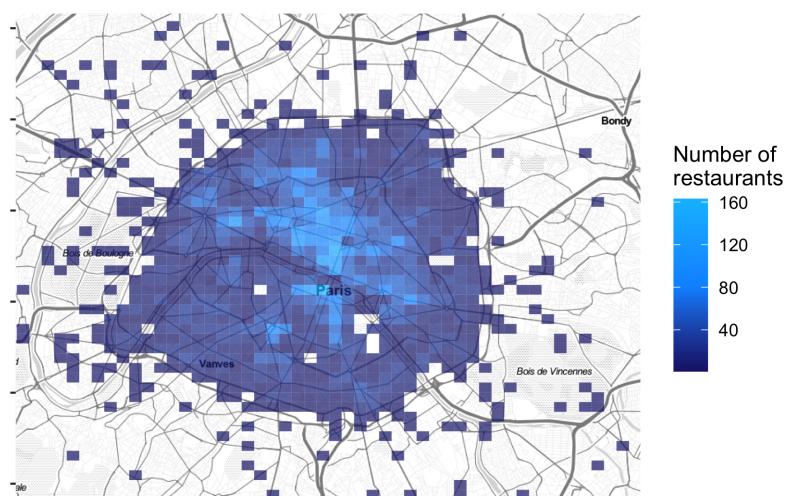
Daily Number of Reviews in Paris



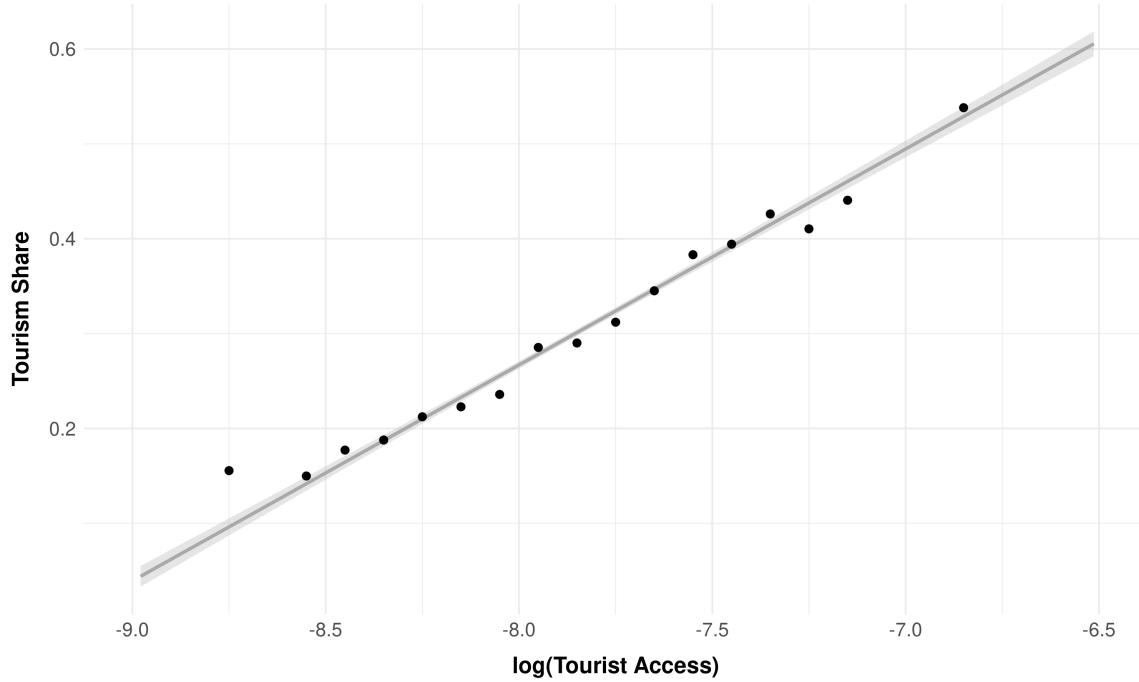
Map of restaurants by share of non-french reviews



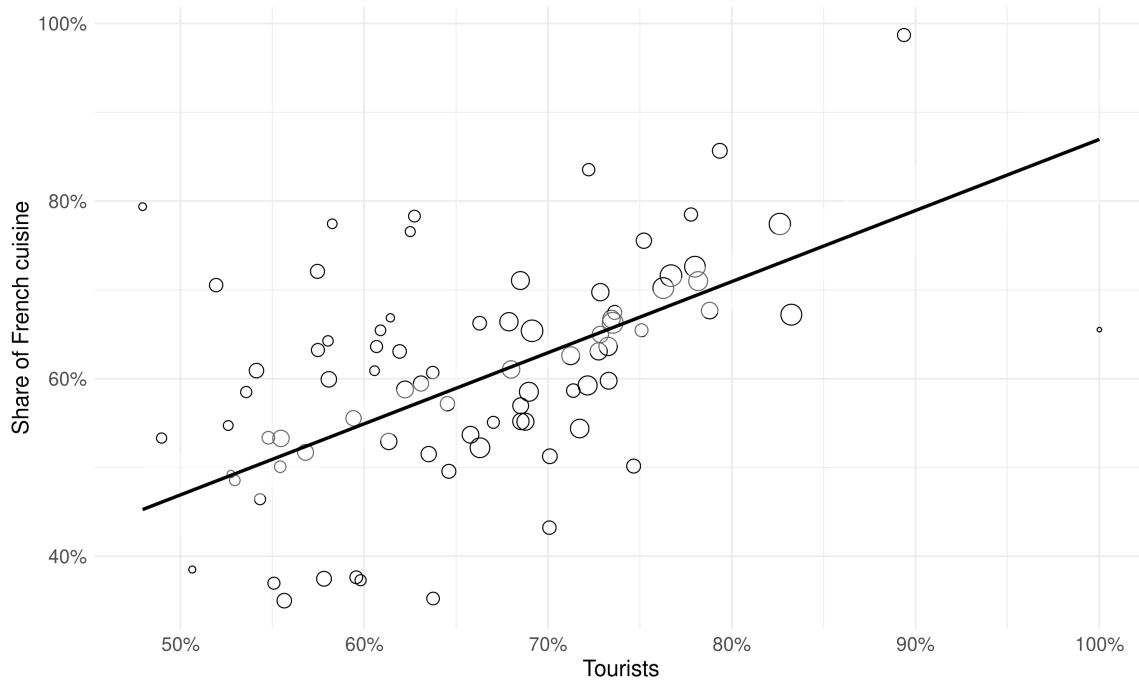
Grid map of restaurants density



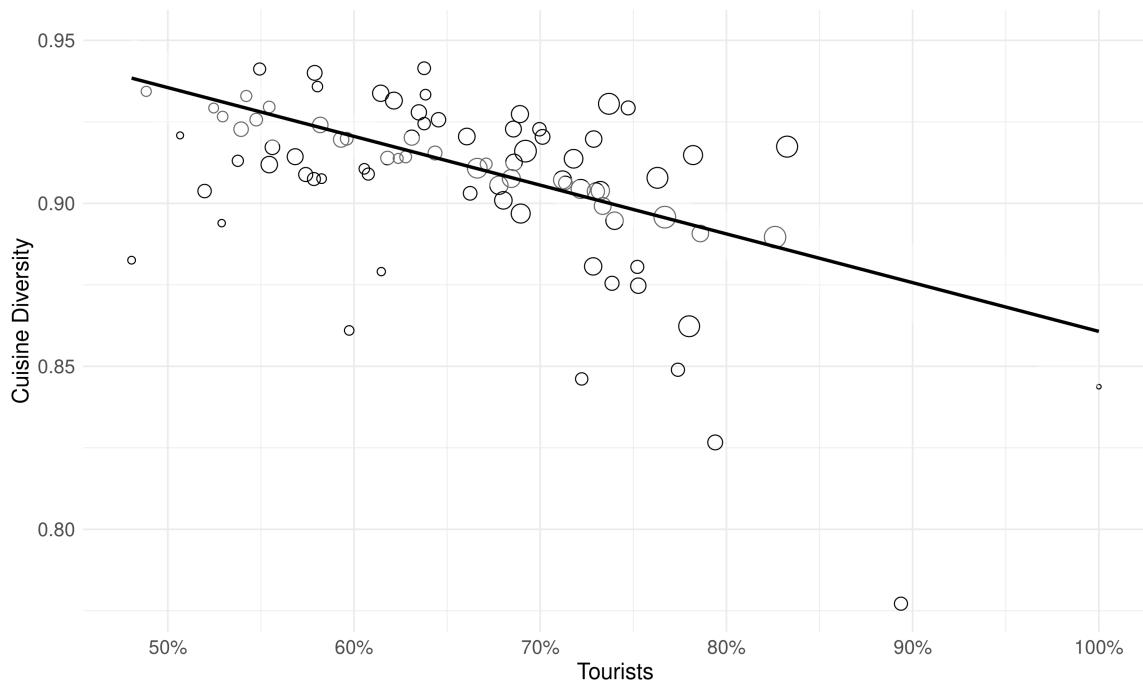
### Tourist Access vs Tourism Proxy



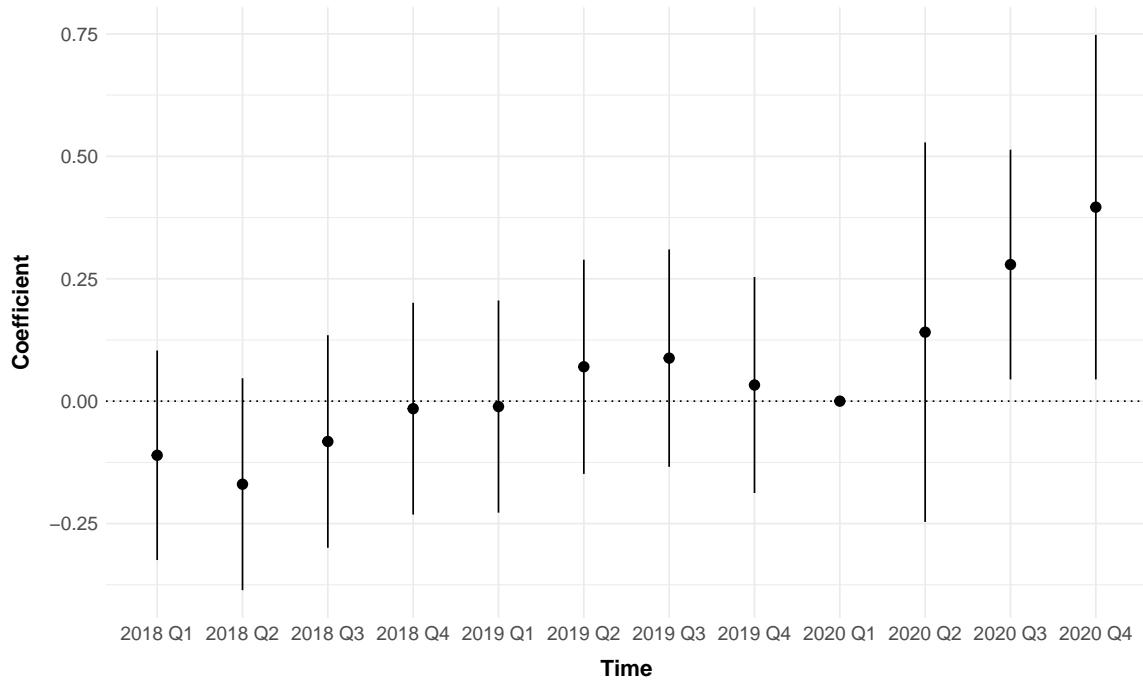
### Share of French Cuisine



### Diversity of Cuisine Types



## Dynamic Effects



## 9. Tables

Table 2.1: Stylized Facts: User Preferences

Dependent Variable:	Rating		
Model:	(1)	(2)	(3)
<i>Variables</i>			
Tourism Share	-0.3932*** (0.0856)	-0.2541*** (0.0710)	-0.3068*** (0.0700)
log(Num of Reviews)	0.0245* (0.0130)	0.0089 (0.0100)	0.0189** (0.0093)
<i>Fixed-effects</i>			
User		Yes	Yes
Quartier			Yes
<i>Fit statistics</i>			
Observations	109,210	109,210	109,210
R <sup>2</sup>	0.00274	0.61455	0.61866
Dependent variable mean	3.8669	3.8669	3.8669

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is an individual review. Dependent variable is a review's rating. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.2: Main Result: Tourism and Restaurant Ratings by Parisians (Restaurant-Level)

	Avg. Rating by Parisian			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Tourism share $\times$ Post-Lockdown	0.3008*** (0.0789)	0.3244*** (0.0952)		
Top 25% Most Touristic $\times$ Post-Lockdown			0.1110*** (0.0368)	0.1037** (0.0410)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month $\times$ Quarter		Yes		Yes
<i>Fit statistics</i>				
Observations	75,876	75,876	75,876	75,876
R <sup>2</sup>	0.35637	0.38035	0.35631	0.38029
Dependent variable mean	3.8599	3.8599	3.8599	3.8599

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is a pair Month  $\times$  Restaurant. Dependent variable is an average rating of restaurants among users with home location in Paris. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.3: Main Result: Tourism and Restaurant Ratings by Parisians (Review-Level)

	Rating			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Tourism Share × Post-Lockdown	0.2781*** (0.0830)	0.1866* (0.0969)	0.2587** (0.1205)	0.3393** (0.1558)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
User		Yes	Yes	
Month × Quarter			Yes	Yes
User × Post-Lockdown				Yes
<i>Fit statistics</i>				
Observations	120,314	120,314	120,314	120,314
R <sup>2</sup>	0.28145	0.73488	0.74564	0.76153
Dependent variable mean	3.8803	3.8803	3.8803	3.8803

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is an individual review. The sample consists of reviews left by users with home location in Paris. Dependent variable is a review's rating. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.4: Tourism and “Dans Ma Rue” Complaints

	# Complaints			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Share Tourism $\times$ Post-Lockdown	-0.6570*** (0.2272)	-0.2581* (0.1364)		
Top 25% Most Touristic $\times$ Post-Lockdown			-0.3527*** (0.1213)	-0.1504** (0.0726)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month $\times$ Quarter		Yes		Yes
<i>Fit statistics</i>				
Observations	366,930	305,332	366,930	305,332
R <sup>2</sup>	0.48157	0.68477	0.48024	0.68481
Dependent variable mean	0.40114	0.48207	0.40114	0.48207

*Notes.* This table reports PPML estimates. The dependent variable is the number of complaints registered on the “Dans ma rue” platform within 100m of a restaurant in a given month. The tourism share is measured as the share of non-French reviews left on a restaurant’s page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Standard-errors clustered at quartier level are in parentheses. Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.5: Textual Outcomes

	Tourists (1)	Low Food Quality (2)	Too Expensive (3)	Too Noisy (4)	Long Wait (5)
<b>Panel A: restaurant-level</b>					
<i>Variables</i>					
Tourism Share × Post-Lockdown	-0.0646*** (0.0112)	-0.0032 (0.0190)	0.0044 (0.0142)	0.0093 (0.0109)	-0.0132 (0.0123)
<i>Fixed-effects</i>					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month x Quarters	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	75,997	75,997	75,997	75,997	75,997
R <sup>2</sup>	0.24881	0.23065	0.19966	0.18782	0.19802
Dependent variable mean	0.02306	0.07168	0.04727	0.02365	0.02561
<b>Panel B: review-level</b>					
<i>Variables</i>					
Tourism Share × Post-Lockdown	-0.0891*** (0.0222)	-0.0032 (0.0311)	-0.0334 (0.0278)	0.0145 (0.0265)	-0.0332 (0.0223)
<i>Fixed-effects</i>					
User-Post-Lockdown	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Quarters	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	111,756	111,756	111,756	111,756	111,756
R <sup>2</sup>	0.56827	0.60988	0.53738	0.47727	0.53808
Dependent variable mean	0.02274	0.07506	0.05095	0.02816	0.02702

*Notes.* This table reports OLS estimates. In all columns of Panel A the unit of analysis is a pair restaurant × month. In all columns of Panel B the unit of analysis is an individual review. Dependent variable is constructed from reviews' texts with the help of dictionaries described in Appendix. In panel A dependent variable is a share of reviews related to the corresponding topic (by restaurant-month). In panel B depended variable is a dummy that switch on when a review is related to a topic. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.6: Social Proximity

	Avg. Rating by Parisian	(1)	(2)	(3)	(4)
<i>Variables</i>					
Tourism Share × Post-Lockdown	0.3073** (0.1206)				
Tourism Share × Post-Lockdown × High SCI	0.1623 (0.1506)				
Tourism Share × Post-Lockdown × Low SCI	0.3379*** (0.1209)				
Top 25% Most Touristic × Post-Lockdown		0.0865 (0.0571)			
Top 25% Most Touristic × Post-Lockdown × High SCI			0.0384 (0.0674)		
Top 25% Most Touristic × Post-Lockdown × Low SCI				0.1209* (0.0637)	
<i>Fixed-effects</i>					
Restaurant		Yes	Yes	Yes	Yes
Month x Quarter		Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	62,050	62,050	62,050	62,050	62,050
R <sup>2</sup>	0.36701	0.36705	0.36696	0.36698	
Dependent variable mean	3.8055	3.8055	3.8055	3.8055	

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is a pair Month × Restaurant. Dependent variable is an average rating of restaurants among users with home location in Paris. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Measure of network proximity between countries of origin are constructed using Facebook data. Restaurants with different proximity score were divided into two groups: above and below median proximity, High and Low SCI respectively. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.7: Textual Outcomes and Social Proximity

	Tourists (1)	Low Food Quality (2)	Too Expensive (3)	Too Noisy (4)	Long Wait (5)
<i>Variables</i>					
Tourism Share	-0.0491*** (0.0096) (0.0153)	0.0197 (0.0177)	0.0295 (0.0334)	0.0043 (0.0241)	-0.0162 (0.0130)
× Post-Lockdown					
× High SCI					
Tourism Share	-0.0816*** (0.0160)	-0.0221 (0.0247)	0.0077 (0.0183)	0.0171 (0.0120)	-0.0135 (0.0135)
× Post-Lockdown					
× Low SCI					
<i>Fixed-effects</i>					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Quarter	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	62,079	62,079	62,079	62,079	62,079
R <sup>2</sup>	0.24497	0.22017	0.18684	0.18442	0.18753
Dependent variable mean	0.02580	0.07424	0.04878	0.02452	0.02618

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is a pair Month × Restaurant. Dependent variable is constructed from reviews' texts with the help of dictionaries described in Appendix. It is a share of reviews related to the one of corresponding topics (by restaurant-month). The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Measure of network proximity between countries of origin are constructed using Facebook data. Restaurants with different proximity score were divided into two groups: above and below median proximity, High and Low SCI respectively. Standard-errors clustered at the quarters level are in parentheses.

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.8: Spillovers

Dependent Variable:	Avg. Rating by Parisian			
Model:	(1)	(2)	(3)	(4)
<i>Variables</i>				
Tourism Share × Post-Lockdown	0.3053*** (0.0836)	0.2790*** (0.1007)	0.3095*** (0.1020)	0.2775*** (0.1036)
Touristic Area (<100m) × Post-Lockdown		-0.1396 (0.1512)		0.0018 (0.1551)
Touristic Area (100m-300m) × Post-Lockdown		0.4084* (0.2432)		0.4558* (0.2657)
Touristic Area (300m-500m) × Post-Lockdown		0.0834 (0.2977)		0.1179 (0.3427)
Touristic Area (500m-1000m) × Post-Lockdown		-0.3662 (0.2911)		0.0816 (0.4458)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
Month × Quarter			Yes	Yes
<i>Fit statistics</i>				
Observations	63,410	63,410	63,410	63,410
R <sup>2</sup>	0.34439	0.34445	0.37327	0.37333
Dependent variable mean	3.8157	3.8157	3.8157	3.8157

*Notes.* This table reports OLS estimates. In all columns the unit of analysis is a pair Month × Restaurant. Dependent variable is an average rating of restaurants among users with home location in Paris. The tourism share is measured as the share of non-French reviews left on a restaurant's page until 2020. Post-lockdown is a dummy, which is switched on in June, 2020 – after the first COVID-19 lockdown. Standard-errors clustered at the quarters level are in parentheses.

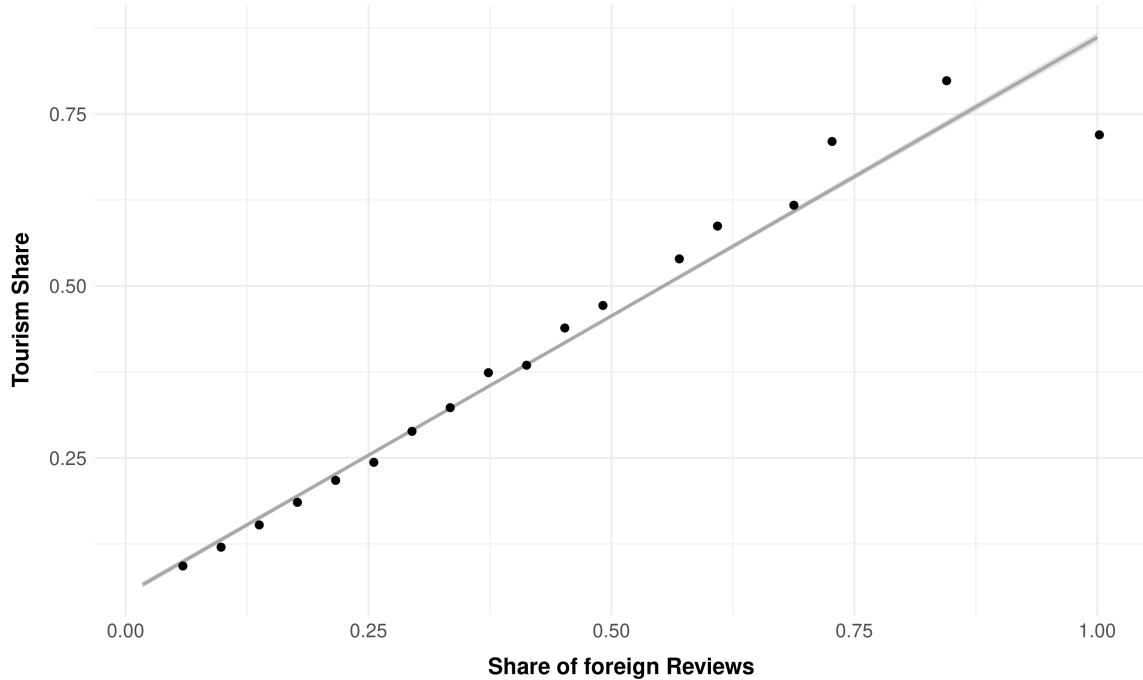
Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## A. Additional Plots

Tripadvisor interface

The screenshot shows the Tripadvisor review submission interface for 'Cafe de Flore' located at 172 boulevard Saint Germain, 75006 Paris, France. At the top, there's a navigation bar with a menu icon, the Tripadvisor logo, and a search icon. Below the header is a thumbnail image of the cafe. The main section contains a message: 'Your first-hand experiences really help other travelers. Thanks!'. Underneath, there's a rating scale with five circles, followed by a button labeled 'Click to rate'. A note says 'Changes will save automatically'. Below the rating is a 'Title of your review' field and a larger 'Summarize your visit or highlight an interesting detail' text area. Further down are fields for 'Your review' and 'Tips for writing a great review', with the latter providing guidance on what to include in the review.

Correlating Different Tourism Proxies



## B. Robustness Checks

### B.1. Alternative Identification: November 2015 Paris attacks

Table 2.B.1: Tourism and Rating: November 2015 Paris attacks

	Rating by Parisians (1)	Rating by Non-Parisians (2)	(3)	(4)
<i>Variables</i>				
Tourism Share × Post-Attack	0.0992** (0.0445)	0.1096** (0.0508)	0.0216 (0.0264)	0.0248 (0.0314)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month × Quarter		Yes		Yes
<i>Fit statistics</i>				
Observations	44,572	44,572	64,387	64,387
R <sup>2</sup>	0.35707	0.37938	0.31664	0.33293
Within R <sup>2</sup>	0.00015	0.00015	$1.36 \times 10^{-5}$	$1.36 \times 10^{-5}$

*One-way (Restaurant) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## B.2. Location-Based Tourism Measure

Table 2.B.2: Location-Based Measure: Tourism and Restaurant Ratings by Parisians: Restaurant-Level Analysis

	Avg. Rating by Parisian			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Tourism Share (location-based) × Post-Lockdown	0.4356*** (0.0925)	0.3984*** (0.0985)		
Top 25% Most Touristic (location-based) × Post-Lockdown			0.1569*** (0.0409)	0.1438*** (0.0442)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month x Quarter		Yes		Yes
<i>Fit statistics</i>				
Observations	75,822	75,822	75,822	75,822
R <sup>2</sup>	0.35615	0.38011	0.35608	0.38007
Dependent variable mean	3.8595	3.8595	3.8595	3.8595

*Clustered (quarter level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.B.3: Location-Based Measure: Tourism and Restaurant Ratings by Parisians:  
Review-Level Analysis

	Rating			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Share Tourism (location-based) × Post-Lockdown	0.4290*** (0.0983)	0.3172*** (0.1156)	0.3592*** (0.1288)	0.3868*** (0.1430)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
User		Yes	Yes	
Month × Quarters			Yes	Yes
User × Post-Lockdown				Yes
<i>Fit statistics</i>				
Observations	120,252	120,252	120,252	120,252
R <sup>2</sup>	0.28131	0.73480	0.74557	0.76145
Dependent variable mean	3.8800	3.8800	3.8800	3.8800

*Clustered (quarter-level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Table 2.B.4: Location-Based Measure: Textual Outcomes

	Tourists (1)	Low Food Quality (2)	Too Expensive (3)	Too Noisy (4)	Long Wait (5)
<i>Variables</i>					
Tourism Share (location-based) × Post-Lockdown	-0.0562*** (0.0111)	-0.0213 (0.0186)	0.0013 (0.0155)	-0.0014 (0.0109)	-0.0165 (0.0119)
<i>Fixed-effects</i>					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Quarter	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	75,943	75,943	75,943	75,943	75,943
R <sup>2</sup>	0.24864	0.23044	0.19964	0.18781	0.19802
Dependent variable mean	0.02308	0.07171	0.04730	0.02367	0.02563

*Clustered (quarter-level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

### B.3. Aggregation of Language-Based Tourism Measure by Different Periods

Table 2.B.5: Tourism and Ratings: Language-Based Tourism Aggregated by Different Periods

		Avg. Rating by Parisian				
		(1)	(2)	(3)	(4)	(5)
<i>Variables</i>						
Tourism share (before 2017) × Post-Lockdown		0.2659** (0.1114)				
Tourism share (before 2018) × Post-Lockdown			0.3171*** (0.1082)			
Tourism share (before 2019) × Post-Lockdown				0.3451*** (0.0987)		
Tourism share (before 2020) × Post-Lockdown					0.3244*** (0.1016)	
Tourism share (before 2021) × Post-Lockdown						0.3290*** (0.1095)
<i>Fixed-effects</i>						
Restaurant		Yes	Yes	Yes	Yes	Yes
Month x Quarter		Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations		57,292	65,515	72,112	75,876	76,350
R <sup>2</sup>		0.37559	0.37228	0.37469	0.38035	0.38273
Dependent variable mean		3.7902	3.8156	3.8433	3.8599	3.8626
<i>Clustered (quarter-level) standard-errors in parentheses</i>						
<i>Signif. Codes:</i> ***: 0.01, **: 0.05, *: 0.1						

#### B.4. Clustering

Table 2.B.6: Tourism and Ratings: Different Clustering

	Avg. Rating by Parisian			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
Tourism Share × Post-Lockdown	0.3244*** (0.1016)	0.3244*** (0.0979)	0.3257*** (0.0952)	0.3257*** (0.0952)
<i>Fixed-effects</i>				
Restaurant	Yes	Yes	Yes	Yes
Month × Quarter	Yes	Yes	Yes	Yes
<i>Clustering</i>				
	Quarter	Grid cell	Restaurant	No
<i>Fit statistics</i>				
Observations	75,876	75,884	75,961	75,961
R <sup>2</sup>	0.38035	0.38046	0.38098	0.38098
Dependent variable mean	3.8599	3.8598	3.8592	3.8592

*Clustered (quarter-level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## C. Validation of Tourism Measures

Table 2.C.1: Tourist Access

	Tourism Share			
	(1)	(2)	(3)	(4)
<i>Variables</i>				
log(Tourist Access)	0.2443*** (0.0171)	0.2170*** (0.0369)	0.2450*** (0.0215)	0.1409*** (0.0326)
Weighted			Yes	Yes
<i>Fixed-effects</i>				
Quartier		Yes		Yes
<i>Fit statistics</i>				
Observations	10,179	10,179	10,179	10,179
R <sup>2</sup>	0.22746	0.31021	0.26590	0.39319
Dependent variable mean	0.31451	0.31451	0.31451	0.31451

*Clustered (quarter-level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## D. Text Analysis

Table 2.D.1: Dictionary for Text Analysis

<b>Low Food quality</b>			
pas bon	sans goût	aucun saveur	réchauff
pas très bon	aucun goût	fade	cuisine bof
mauvaise cuisson	goût bizzare	industriel	avarié
pas assez cuit	trop cuit	supermarch	tombé malade
pas cuit	sans saveur	mauvaise qualité	vomir
indigestion	intoxication	pas frais	surgel
insipide	dégueulass	degueulass	micro-ond
pas fait maison			
<b>Too Expensive</b>			
prix élevés	cher	prix sont élevés	prix sont très élevés
<b>Too Noisy</b>			
bruyant	beaucoup de bruit		
<b>Long Wait</b>			
long	lent		
<b>Tourism</b>			
touris			

*Notes.* This table reports phrases that were used in our text analysis. Terms are not always the full forms of the words, which helps to take into account the syntax. We also do not include to this table potential distortions of the same phrases, which were also used in our analysis (missing accent marks, common misspellings).

Table 2.D.2: Summary Statistics for Textual Variables

Variable	N	Mean	St. Dev.
Tourism	1,154,860	0.025	0.157
Low Food Quality	1,154,860	0.066	0.248
Too Expensive	1,154,860	0.050	0.218
Too Noisy	1,154,860	0.028	0.165
Long Wait	1,154,860	0.024	0.153

Table 2.D.3: Ratings and Textual Variables

	Rating					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Tourists	-0.3413*** (0.0370)					-0.2868*** (0.0363)
Low Food Quality		-1.163*** (0.0208)				-1.138*** (0.0207)
Too Expensive			-0.4439*** (0.0228)			-0.3939*** (0.0214)
Too Noisy				-0.2186*** (0.0275)		-0.1930*** (0.0255)
Long Wait					-0.4257*** (0.0280)	-0.3845*** (0.0255)
<i>Fixed-effects</i>						
User	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes	Yes
Date	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	112,905	112,905	112,905	112,905	112,905	112,905
R <sup>2</sup>	0.74586	0.76787	0.74789	0.74560	0.74653	0.77195
Dependent variable mean	3.8863	3.8863	3.8863	3.8863	3.8863	3.8863

*Clustered (quarter-level) standard-errors in parentheses*

*Signif. Codes:* \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

# Chapter 3

## Going Viral in a Pandemic: Social Media and Allyship in the Black Lives Matter Movement

### Abstract

How can modern social movements broaden their base? Prompted by the viral video footage of George Floyd's murder, the Black Lives Matter (BLM) movement gained unprecedented momentum and scope in the spring of 2020. Using Super Spreader Events as a source of plausibly exogenous variation at the county-level, we find that pandemic exposure led to an increase in the likelihood of observing online and offline BLM protests. This effect is most pronounced in whiter, more affluent and suburban counties. We develop a novel index of social media penetration at the county level to show that this effect is driven by higher social media take-up among non-traditional users. Specifically, we find that a one standard deviation increase in pandemic exposure led to a doubling of new Twitter accounts in counties with no BLM protest history. Our results suggest that the pandemic acted as a demand shock to social media among non-traditional users, mobilizing new segments of society to join the movement for the first time. We find supporting evidence for this mechanism using individual-level survey data and rule out competing channels, such as pandemic induced salience of racial inequality, lower opportunity cost of protesting or higher overall agitation and propensity to protest.

# 1. Introduction

*There is a far more representative cross-section of America out on the streets [...]*

*That didn't exist back in the 1960s. That broad coalition.*

- Barack Obama, June 3rd 2020

The effectiveness of social movements depends on their ability to mobilize allies, build coalitions and inspire reform through collective action Olson (1989); Ostrom (1990); Della Porta and Diani (2015, 2020). Traditionally, mobilization was carried out at the local level via face-to-face interactions. Today, activism is organized in the virtual space. For instance, the Civil Rights Movement in the 1960s depended heavily on local chapters as decision making, mobilization, coordination and persuasion tools Morris (1986). One of its successors - the Black Lives Matter (BLM) movement - was born on Twitter in 2013 and relies primarily on social media to communicate with the broader public and mobilize protesters.<sup>1</sup>

The #BlackLivesMatter hashtag has become one of the most frequently used hashtags on Twitter, peaking at 8.8 million tweets per day in May 2020 (PEW, 2020). Videos on Twitter about the murder of George Floyd by the police officer Derek Chauvin were watched over 1.4 billion times within two weeks.<sup>2</sup> The ensuing protest in May of 2020 were labeled the “largest” and the “broadest” social movement in the history of the United States.<sup>3</sup>

What led to the broadening of the movement’s coalition during the pandemic? We approach this question in two parts. First, we establish a causal link between exposure to COVID-19 and protest participation at the county level, using Super Spreader Events as a source of exogenous variation. We show that exposure to COVID-19 is associated with an increase in protest behavior but only among those counties that have never protested for a BLM-related cause before.

Second, we develop a novel index of social media penetration at the county level to show that this effect is driven by higher social media take-up during the pandemic but before the protest trigger. While we cannot fully rule out that other mechanisms were at play, we show evidence that alternative explanations such as *i*) a pandemic-induced rise in the salience of racial inequality, *ii*) lower opportunity costs of protesting, *iii*) higher overall propensity to protest and *iv*) a scattering rather than a broadening protest are not driving our results.

Previous work has shown that social media can solve the collective action and coordination problem for individuals already sympathetic to a political cause Enikolopov et al. (2020);

---

<sup>1</sup>As McKersie (2021) notes: ”Even though an organization like BLM does not have a constituent base like the CCCO, through which affiliated congregations and neighborhood organizations issued calls for participants, current BLM organizations more than compensate by utilizing the power of social media to mobilize participants for protests.”

<sup>2</sup>See Listing of Twitter Videos with George Floyd and BLM hashtag

<sup>3</sup>See New York Times and Washington Post

Manacorda and Tesei (2020). In contrast, we focus on the role of social media as a tool that can *broaden* alliances and mobilize new fractions of society. In addition, previous papers exploit supply side constraints (informal networks or infrastructure) in the early stages of internet or social media roll-out going back to the early 2000s Guriev et al. (2019); Müller and Schwarz (2020); Enikolopov et al. (2020); Manacorda and Tesei (2020). However, initial constraints become less relevant over time and do not account for more recent determinants of social media penetration. To the best of our knowledge, we are the first to show that COVID-19 acted as a demand shock for social media among "non-traditional" users and that this is an important driver behind the broadening of the BLM movement during the pandemic.

Our identification is based on a small window between the end of March and mid April of 2020 during which COVID-19 was prevalent enough but lock-down stringency lax enough to allow for so-called Super Spreader Events (SSE) to occur. These events are characterized by the presence of one highly infectious individual (a super-spreader) and took place mainly at birthday parties, nursing homes or prisons. We exploit cross-sectional variation in the number of SSEs within a 50 kilometer radius from the county border but not within the county 6 weeks prior to the murder of George Floyd to construct our instrument for exposure to COVID-19 at the county level. We include state fixed effects and a vast set of county level controls, most notably the number of historical BLM events between 2014 and 2019, as well as socio-demographic variables and proxies for political leaning and social capital.

We find robust evidence that exposure to COVID-19 increased BLM protest. We estimate that a one standard deviation increase in the number of COVID-19 related deaths in a county at the time of George Floyd's murder (approximately 25 deaths per 100K inhabitants), increases the likelihood of a BLM event occurring in the three weeks following the murder by 5%. Our baseline result is entirely driven by counties with no prior BLM protests and the effect doubles in size and is more precisely estimated for this sub-sample.

We summarize all robustness checks on our instrument and main results in section 7 and present them in more detail in Appendix Appendix A and Appendix B. We preview here that we perform several exercises to probe the plausibility of the exclusion restriction. Most importantly, we *i*) show in a placebo test that SSEs do not predict past BLM events, and using LASSO *ii*) we weight SSEs by their inverse probability of occurrence and *iii*) include a control variable that captures the pre-pandemic protest propensity.<sup>4</sup> Our results hold for various iterations of our SSE instrument (varying distance, time lag, and cases associated with SSEs). Moreover, we check the robustness of our main results with respect to changes in sample composition, spatial correlation, and definition of the treatment and outcome

---

<sup>4</sup>We describe the LASSO selected model in detail in Appendix section B.3.

variables.

In addition, we propose three alternative identification strategies and show that our results replicate. First, using large scale mobile phone mobility data by *SafeGraph*, we instrument pandemic exposure with tourist flows to one of the largest SSEs in the US - Florida spring break in March 2020. Second, we employ a difference in differences approach, for which we scrape information on all similar BLM protest triggers since 2014 to estimate the differential response to a protest trigger before and after the pandemic. Third, we use a LASSO-based matching approach, comparing counties with similar pre-pandemic protest probabilities.

In a next step, we investigate various sources of heterogeneity and show that - in line with the idea of a broadening movement - our baseline results are driven by whiter, more affluent and sub-urban counties. We also look at alternative outcomes and find that exposure to COVID-19 increases the frequency of BLM protest without diminishing its scope (total number of participants or average number of participants per event). Moreover, we also find evidence that exposure to COVID-19 increases online protest, measured as the number of BLM-related tweets and the number of followers of the official BLM twitter account. Lastly, we geo-localize street art related to George Floyd from the *Urban Anti-Racist Street Art Mapping* project and find no effect of exposure to COVID-19 on pro-BLM street art. We interpret this outcome as form of BLM protest with high barriers to entry (unlike offline and online protest) as it relies on existing networks and cultural capital.

In the second part of the paper, we investigate whether the uptake in social media can account for the pandemic-induced broadening of the BLM movement. We start by repeating the above analysis, this time using a novel index of social media penetration as our main outcome variable. The index is measured *before* the protest trigger but after the outbreak of the pandemic in the United States (i.e. the first detected case on January 20, 2020 prior to George Floyd's murder on May 25th). We use the first principle component of multiple variables: *i*) the (log) cumulative number of new twitter accounts, which we obtain by scraping and geo-coding information on the creation date of new twitter accounts at the county level from approximately 45 million tweets, *ii*) the (log) number of new followers of the official BLM account *iii*) Google searches for the term "Twitter", hypothesizing that new users will Google the term first to create an account and *iv*) Google mobility data at the county level, assuming that increased residential stays (time spent at home) as well as lower social, work and leisure mobility is associated with more time spent online.<sup>5</sup>

---

<sup>5</sup>We use a normalized index of search activity for the term 'twitter' provided by Google Trends. Search activity indices are provided as integers from zero to 100 with an unreported privacy threshold. Each observation is the number of searches of the given term divided by the total searches from the geography and time range, which is then normalized between regions such that the region with the largest measure is

We find that the pandemic has a positive and significant effect on our social media index and that this is entirely driven by the sub-sample of counties that have never protested before. For instance, we show that a one standard deviation increase in pandemic exposure led to a doubling of twitter accounts among counties with no prior BLM event, without affecting counties that traditionally protest.

In a next step, we zoom in on the role of twitter in mobilizing BLM protesters. First, we interact baseline twitter penetration (before the pandemic) with exposure to COVID-19. We address the concern that our results could capture underlying factors that drive both Twitter penetration and protest participation, replicating the SXSW instrument for baseline Twitter penetration used by Müller and Schwarz (2020). We show that counties with higher baseline twitter penetration react more to pandemic exposure. This is in line with two mutually non-exclusive interpretations. First, counties with higher baseline twitter penetration may react more to the social media demand shock, as the marginal users has a bigger incentive to join social media when the existing network is large. Second, the pandemic may also serve as a demand shock at the intensive margin with existing users spending more time on social media. Additionally, we interact pandemic exposure with contemporaneous twitter penetration and find that the effect of COVID-19 on protest is entirely driven by counties with higher twitter take-up during the pandemic.

To probe the social media mechanism further, we use individual-level survey data. Interpreting these results with caution, we find that individuals living in a county with higher COVID-19 deaths are more likely to receive news about George Floyd through social media than through other channels.<sup>6</sup> We also find that COVID-19 exposure is associated with more sympathy for the movement and higher salience of racial injustice among respondents (controlling for race, gender, education, income, and political leaning) without changing attitudes towards other progressive issues, such as "illegal" immigration.

In the last part of our paper, we look at competing mechanisms. Naturally, the pandemic has affected a number of important dimensions that are not limited to higher social media take-up. First, we consider the possibility that our results are driven by a scattering rather than a broadening of BLM protest. More specifically, we verify that the effect is not driven by a substitution away from some locations to others. Second, the pandemic may have increased the overall salience of racial inequality *before* the murder of George Floyd. We test this by interacting COVID-19 with a proxy for disproportional death burden on Blacks and the number of BLM-related search terms on Google before the protest trigger. Third,

---

set to 100. The Google Trends data is defined on a designated market area (DMA) level.

<sup>6</sup>The data set does not contain information on the location of the respondent but only whether they live in a low, medium or high COVID-19 county. Therefore, we cannot employ our instrument for exposure to COVID-19.

we investigate whether the pandemic has decreased the opportunity cost of protesting. We interact COVID-19 with the unemployment rate at the county level and stringency at the state level. If individuals choose to protest in lieu of going to work or engage in social activities, we should see a larger effect in counties with higher unemployment rates or stricter stringency measures. Third, we look at the effect of COVID-19 on other protests. If the pandemic increased overall agitation and propensity to protest, then we would expect this to also hold for other causes beyond BLM. We show that these channels are unlikely to drive our results.

We contribute to the nascent literature on the effect of the internet on political outcomes Falck et al. (2014); Lelkes et al. (2017); Boxell et al. (2017); Campante et al. (2018); Guriev et al. (2019) and the effect of social media on xenophobia, polarization, political preferences, social capital and protests more specifically Acemoglu et al. (2018); Enikolopov et al. (2018); Bursztyn et al. (2019); Enikolopov et al. (2020); Manacorda and Tesei (2020); Müller and Schwarz (2020); Zhuravskaya et al. (2020); Müller and Schwarz (2021); Fujiwara et al. (2021); Campante et al. (2021). To the best of our knowledge, we are the first to investigate the role of social media in *broadening* political coalitions through persuasion, rather than mobilizing individuals that are already sympathetic to the movement's grievances.

Typically, these papers consider (the lack of) protest mobilization as a collective action problem, where access to information reduces coordination costs and therefore increases participation. For instance, Cantoni et al. (2019) and Bursztyn et al. (2021) show in an experimental setting in Hong-Kong that information about other people's turnout encourages individual protest participation and that this has longer-run effects on the propensity to protest if a sufficiently large fraction of the network is mobilized. They conclude that one-time mobilization shocks can have persistent effects on the dynamics of social movements.

Most similar to our study, Enikolopov et al. (2020) show that social media helps to solve the collective action problem in a one-shot setting, where the expansion of a social media platform coincides with a contested election in Russia. Similarly, Manacorda and Tesei (2020) exploit the expansion of mobile phone reception in Africa to show that access to information and communication technologies will only increase protest if economic grievances are high and opportunity costs are low (e.g., during economic downturns). In contrast to these papers, we are able to identify for which groups exposure to social media is particularly effective and how it can persuade individuals at the margin. In addition, we overcome important challenges in identifying the causal effect of social media in saturated markets.

Our analysis also contributes to a large literature that analyzes the determinants of social movements and protests, ranging from macro level drivers, such as local institutions or socio-economic conditions Lipsky (1968); Eisinger (1973); McCarthy and Zald (1977);

Besley and Persson (2011); Dube and Vargas (2013); Berman et al. (2017), to micro level drivers, including individual decision making processes Ellis and Fender (2011); Guriev and Treisman (2015); Sangnier and Zylberberg (2017) and different aspects of individual and social psychology, as well as protest as a collective action problem Guriev and Treisman (2015); Sangnier and Zylberberg (2017); Passarelli and Tabellini (2017); Cantoni et al. (2019); Enikolopov et al. (2020); Manacorda and Tesei (2020); González and Prem (2020); Hager et al. (2020); Bursztyn et al. (2021).

The remainder of the paper is organized as follows. In section 2, we provide some background on the BLM movement, present some motivating evidence and describe our main data sources. We present our empirical strategy in section 3 before moving to our main results in section 4. Section 5 provides various pieces of evidence for the social media mechanism. Section 6 addresses competing mechanisms. Section 7 provides a summary of all robustness checks performed and section 8 concludes.

## 2. Background and Data

### 2.1. *BLM History and Motivating Evidence*

The Black Lives Matter (BLM) movement emerged on social media after the acquittal of George Zimmerman in the deadly shooting of a Black teenager named Trayvon Martin. The movement was founded by three Black activists, Alicia Garza, Patrisse Cullors, and Opal Tometi in July of 2013 with the aim to end systemic racism, abolish white supremacy and state-sanctioned violence Black Lives Matter (2020), and more generally, to “fundamentally shape whites’ attitudes toward Blacks” Mazumder (2019).

Over the following months, an ever-increasing but small number of activists coalesced under the hashtag #BlackLivesMatter on Twitter and Facebook. In August of 2014, after a court decision to not indict the responsible police officer in the fatal shooting of Michael Brown in Ferguson, #BLM became one of the most widely used hashtags on Twitter (the hashtag was used 1.7 million times in the three weeks following the court decision, compared to 5000 tweets in all of 2013, see Freelon et al. (2016); Anderson and Hitlin (2016)), confirming its status as a mainstream social media phenomenon. The shooting of Michael Brown was followed by a large and protracted protest in the city of Ferguson. The consequences of this shooting rippled throughout American society, generating counter-movements under the hashtag #AllLivesMatter and #BlueLivesMatter and mobilizing protesters (for and against the cause) far beyond the city’s borders.

BLM played a crucial role in transforming localized activism into a coordinated move-

ment across various locations within and outside of the United States. The founders state that "[...] when it was time for us to leave, inspired by our friends in Ferguson, organizers from 18 different cities went back home and developed Black Lives Matter chapters in their communities and towns — broadening the political will and movement building reach catalyzed by the #BlackLivesMatter project" Black Lives Matter (2020). The *Black Lives Matter Global Network Infrastructure* was designed to provide decentralized actors with resources and guidelines to organize protests, receive information about the movement, and coordinate through social media.<sup>7</sup>

In the following years, the BLM movement expanded geographically and demographically, attracting an unprecedented number of participants after the murder of George Floyd in Minneapolis on May 25th 2020. Protesters took to the streets when a video of the murder of George Floyd went viral on social media, showing how police officer Derek Chauvin suffocated George Floyd using a choke-hold. The video spurred unrest in Minneapolis but the protests quickly expanded to other parts of the United States, including communities that had never engaged in BLM protests before. The number of BLM protests quadrupled in May and June of 2020, compared to previous peaks in 2016 (see Figure 3.1).

The surge in BLM protests in the spring of 2020 is all the more remarkable as the COVID-19 pandemic was well underway. At the time of George Floyd's murder almost 100,000 COVID-19-related deaths had been recorded in the United States and the country was reeling under the first wave of the pandemic (see Figure 3.2). Tough lockdown and social distancing measures were imposed in many counties to prevent the spread of the virus. Average lockdown stringency peaked in May Hale et al. (2020) and the Center for Disease Control and Prevention urged the public to "remain out of congregate settings, avoid mass gatherings, and maintain distance from others when possible" CDC (2020).

A key motivating observation for our study is the exceptionally high level of participation in BLM protests after the murder of George Floyd (see Figure 3.1). While the outbreak of the pandemic and the peak in BLM protests coincided, the surge in protests may still have been driven by counties that were less exposed to the pandemic. If we split the sample into above and below median COVID-19-related deaths at the county level and plot the BLM protests in 2020 in the top panel of Figure 3.3, we also find a geographical link between exposure to COVID-19 and BLM protests. In the bottom panel of Figure 3.3, we plot the evolution of tweets that mention the hashtags #BLM or #BlackLivesMatter. Using an algorithm that assigns tweets to geographic locations, we are able to assign these tweets to counties that experience above and below median COVID-19-related deaths. We find that locations that were more affected by COVID-19 increase their online protest activity. These descriptive

---

<sup>7</sup><https://blacklivesmatter.com/herstory/>

plots suggest that - despite the fear of contagion and the stringency of social distancing measures - there is both a temporal and a geographical relationship between COVID-19 intensity and occurrence of BLM protests.

Lastly, we find that - in line with public perception - the BLM movement has broadened in scope. We divide the counties into those that always protest for BLM and those that protested for the first time after George Floyd was murdered.<sup>8</sup> Figure 3.4 plots in black counties that had at least one BLM protest pre-pandemic and also protested after George Floyd's death. Counties that recorded their first BLM protest only after George Floyd's murder are shown in green. Our data reveals that the geographic spread of first time protesters does not follow the typical coastal geographic clusters, but rather spread across all of the United States. Interestingly, counties with no BLM events prior to George Floyd's murder make up half of the counties protesting in the weeks following Floyd's murder.

There are three takeaways from this evidence. First, the BLM movement has gained unprecedented scope during the pandemic. Second, there is a geographic link between COVID-19 exposure and online and offline BLM protests. Third, a meaningful proportion of protesters in 2020 come from counties that have never protested for a BLM-related cause before. We use these observations to guide our empirical analysis.

## 2.2. Main Data Sources

In this section, we present the primary data sources on the COVID-19 pandemic, BLM and other protests, Twitter data and other county-level socio-demographic and political information. Summary statistics are presented in Table 3.1 and a breakdown of summary statistics by sub-samples (counties with and without prior BLM events) is presented in Appendix Table 3.C.1. We describe the additional data sources in more detail in Appendix Appendix D and provide an overview of the main sources in Appendix Table 3.D.1.

**COVID-19.** Data on COVID-19 related deaths and cases in the USA at the county level comes from the New York Times. This data set provides the cumulative count of cases and deaths every day for each county in the USA, starting from January 21, 2020 when the country's first COVID-19 case was reported. A key limitation of COVID-19 cases data is that it depends on the testing facility and availability of the test kits in the region. We therefore mainly rely on COVID-19 related deaths as a measure of exposure to the pandemic. We also obtain data on daily COVID-19 hospitalizations and deaths by race and ethnicity at the state-level from the Center for Disease Control and Prevention.

---

<sup>8</sup>We use data from *Elephrame* on BLM events between 2014 and 2020 and describe this data set in more detail in the next section and in Appendix Appendix D.

**Super spreader events.** We collect data on COVID-19 super spreader events from a project started by independent investigators and researchers from London School of Hygiene and Tropical Medicine Leclerc et al. (2020). Data are put together based on scientific journals and news reports on super spreader events, which are defined as "clusters" or "outbreaks" of COVID-19 infections with a minimum of 2 infections outside of the home. For the whole period (January 2020 to January 2021), we identify a total of 1074 super spreader events in the USA. Most commonly, events occur in nursing homes, prisons, factories, and retribution (correction facility) or medical centers. Figure 3.5 shows the distribution of these events by their type and Table 3.C.2 provides descriptive statistics about each type of event. We describe the nature of these events in more detail in section 3 and lay out the limitations of the SSE data set and how we address those in Appendix Appendix D.

**Black Lives Matter.** This data comes from the crowd-sourced platform Elephrame. It provides information on the place and date of each BLM protest and estimated number of participants, as well as a link to a news article covering the protest. We extracted records of all protests from June 2014 to September 2020 and geo-coded their location. We also collected and geo-located cross-sectional information on street art with BLM and George Floyd-related content from the Urban Art Mapping George Floyd and Anti-Racist Street Art database. We add information on non BLM-related protests from the US Crisis Monitor, a joint project between ACLED and the Bridging Divides Initiative (BDI) at Princeton University, that collects real-time data on different types of political violence and protests in the US from Spring 2020 to present day.

**Twitter.** We collect three types of Twitter data at different points in time (before the pandemic, during the pandemic but before the murder of Floyd and in the three weeks after the murder of Floyd). First, from the Twitter API we collect the universe of tweets with BLM related hashtags. This includes the hashtags #BlackLivesMatter, #BlackLifeMatters, #BLM, #AllLivesMatter, and #BlueLivesMatter.<sup>9</sup> Second, we collect data to proxy the broader use of Twitter by taking a random sample of tweets that use the most common 100 words in the English language. Third, we scrape information on all followers of the official Black Lives Matter Twitter account (as of March 2022). With the help of a geo-location algorithm, we can assign about 5 to 20% of Twitter users (depending on the sample) to counties. We show in Appendix Table 3.D.3 that, reassuringly, the characteristics of counties for which we have geo-located tweets are remarkably similar to the full sample of counties. Using this data we are able to proxy *i*) online protest for and against BLM with the number

---

<sup>9</sup>We present a selection of tweet examples from our collected sample in Appendix Table 3.D.4

of tweets containing the relevant hashtags *ii*) the number of new Twitter accounts, using the creation date of the Twitter accounts and *iii*) information on baseline Twitter penetration. Finally, to reproduce the instrument for Twitter usage used by Müller and Schwarz (2020) we collect the list of followers of the account of the SXSW festival, which provided an initial boost to Twitter usage. Appendix Appendix D provides more detail on the collection and construction of the Twitter data used in this analysis.

**Google.** We use two main sources from Google. First, data on mobility to understand the mechanism of observing protests during pandemics. This data collects information on the time a person spent on certain mobility tasks like the time spent in parks, being at home, doing groceries, in the transit stations and finally at their workplace (as identified by Google). This information is then aggregated at the county level to measure the aggregate daily mobility. Second, data on Google search terms from the Google Trends API at the Designated Market Area and day level. We use this information to proxy interest in Twitter, George Floyd and the Black Lives Matter Movement at different points in time. In Appendix Appendix D, we describe the Google data and related search terms in more detail.

**Survey Data.** We use data from the American Trends Panel survey conducted by the Pew Research Center to estimate the link between COVID-19 death rates and change in use of social media and public opinion on racial disparities and the BLM movement. We analyse data from wave 68 that took place between June 4th and June 10th, 2020. This data set does not include information on the county of the respondent but only the exposure to COVID-19 (categorized as low, medium or high) in their county of residence at the time of the interview.

**Additional county-level controls.** We include unemployment data available on a monthly basis at the county level from the Local Area Unemployment Statistics of the US Bureau of Labor Statistics and the total population, population by ethnicity, income statistics (such as Black poverty rate and median household income (all in 2018), as well as past Republican vote share (in 2012 and 2016) from the American Community Survey. We use a dummy for rural counties which is constructed from the Office of Management and Budget's February 2013 delineation of metropolitan and micropolitan statistical areas.<sup>10</sup> The measure of social

---

<sup>10</sup>2013 NCHS Urban-Rural Classification Scheme for Counties, Vintage 2012 postcensal estimates of the resident U.S. population. NCHS Urbanization levels are designed to be convenient for studying the difference in health across urban and rural areas. This classification has 6 categories: large “center” metropolitan area (*inner cities*), large “fringe” metropolitan area (*suburbs*), median metropolitan area, small metropolitan area, micropolitan area and non-core (nonmetropolitan counties that are not in a micropolitan area).

capital that we use aggregates the information on the number of local organizations.<sup>11</sup> In addition, we include an index of county resilience towards a pandemic provided by the US Census bureau, which incorporates health and infrastructure indicator and is described in more detail in Appendix Appendix D.

### 2.3. Descriptive statistics

Table 3.1 presents summary statistics on the main variables of interest for the full sample. As outlined above, we use information that is available at different points in time. We present 5 panels that split the variables according to when they are measured: *i*) three weeks after George Floyd’s murder, *ii*) the day of the murder, *iii*) before the murder but after the pandemic started in January 2020, *iv*) later outcomes and *v*) baseline county characteristics before the outbreak of the pandemic. Our main outcome variables are measured in the three weeks following the murder of George Floyd, from May 25th to June 14th of 2020. COVID-19 related deaths and cases, our main treatment variables, are measured at the day of the murder. We measure proxies for online activity and use of social media (new Twitter account, Google searches for Twitter and BLM, mobility patterns etc.) before the murder of Floyd. Some variables are not time-stamped and are only available cross-sectionally at the time of scraping (followers of the main BLM Twitter account and street art were scraped in February 2022). Control variables are drawn from various sources at the closest available year. For instance, variables from the American Community Survey are measured in 2018, vote shares are measured in 2012 and 2016. Appendix Table 3.D.1 reports the exact time frames of all variables used in our analysis.

The average likelihood of observing a BLM-related protest at the county level between May 25th and June 14th lies at about 10%. There are on average 0.25 events per county in the three weeks following George Floyd’s murder and the average number of participants is approximately 270 with a maximum of over 320K participants.<sup>12</sup> If an event occurs, the average number of participants per event is about 540. In the three weeks following George Floyd’s murder we can identify about 820 tweets per county using BLM-related hashtags and about 4 to 5 new users per county (those created after the pandemic started but before the murder of Floyd) who start tweeting about BLM.

The per county average number of cumulative COVID-19 related deaths is 24 (or 0.113 per 1000 population) by May 25th 2020. Absolute cumulative cases are approximately 460

---

<sup>11</sup>This includes: (a) civic organizations; (b) bowling centers; (c) golf clubs; (d) fitness centers; (e) sports organizations; (f) religious organizations; (g) political organizations; (h) labor organizations; (i) business organizations; and (j) professional organizations.

<sup>12</sup>The average sets the number of participants in places with no BLM protests as zero.

per county (or 2.8 per 1000). The maximum number of deaths in a county at the time was 3,300, compared to 31,000 deaths in March 2022. While COVID-19 cases and deaths were comparatively low, the salience of the pandemic was particularly high. In fact, lockdown stringency in the United States peaked in late April 2020. We also report the Black Death Burden (BDB) and find that Blacks were disproportionately affected by the pandemic. The average BDB index is 1.3 indicating that Blacks died at a rate 30% higher than their share of the population would predict. The average county experienced about three Super Spreader Events in its immediate surroundings between January 2020 and April 2020.

In addition, we report detailed summary statistics for the different sub-samples in Table 3.C.1. We report the full sample in the left-hand columns and present a breakdown of the summary statistics by sub-sample in the middle and right-hand side of the table. We distinguish between counties with no BLM events before the pandemic and those with prior BLM events. The vast majority of counties where there was no history of protest for a BLM-related cause continue to not protest after the murder of George Floyd (2,635 counties, which is approximately 85% of all counties). However, we observe that among the sample of "no BLM event before" 133 counties start to protest for the first time during the pandemic. We also report summary statistics on the traditional protesters, i.e. counties that have had a prior BLM protest. Among those 339 traditional protesters, 123 counties stop protesting after the murder of George Floyd and 176 counties continue to protest. As expected, the average probability of observing a protest in response to the murder of George Floyd is 10 times higher among traditional protesters compared to other counties. Remarkably, however, the first-time protesters make up nearly 50 percent of all counties that protested during the pandemic. Counties that traditionally protest have a higher Black population share and higher median household income and are more urban and Democratic leaning than the counties that had never protested before.

### 3. Empirical Strategy

#### 3.1. Baseline Estimating Equation

To study the effect of exposure to COVID-19 on BLM protests, we estimate

$$BLM_c = \beta_0 + \beta_1 Covid_{cs} + \mathbf{X}_c \boldsymbol{\beta}_{\mathbf{X}} + \delta_s + \epsilon_{cs} \quad (3.1)$$

where  $BLM_c$  is a dummy variable for the presence of a BLM protest in county  $c$  during the three weeks following the murder of George Floyd.<sup>13</sup>

---

<sup>13</sup>We restrict the sample for our main outcome of interest to the three weeks after the death of George

We are interested in the coefficient  $\beta_1$ , which captures the effect of one additional COVID-19 related case per 1000 inhabitants in county  $c$  of state  $s$  at the time of George Floyd's murder on May 25th 2020. In addition to state fixed effects  $\delta_s$ , the vector  $\mathbf{X}_c$  includes an array of county level controls (we describe all these variables in detail in Table 3.1). Specifically, we include variables that are associated with participation in the BLM movement, such as a dummy for urban counties and Black population share and the poverty rate among Blacks. Most importantly, we also include two major determinants of BLM protests following the murder of George Floyd, namely the number of BLM events before the murder (starting 2014) and the use of deadly force by police (i.e. number of Black people who died during an encounter with the police, excluding suicides, for two time periods: from summer 2014 to 2019 and in 2020 up to May 25th). We also control for underlying political and attitudinal factors and socioeconomic drivers of protest and social media use, such as the vote share for Republicans in the 2012 and 2016 presidential elections, median household income, unemployment rate, community resilience, and two proxies for social capital (number of civil organizations and number of religious organizations). We cluster standard errors at the state level.

### *3.2. IV Estimation: Super Spreader Events*

A key empirical challenge in ascertaining the causal impact of exposure to COVID-19 on BLM protests is that both occurrences could be driven by unobserved factors. For instance, tight-knit and socially active communities may both increase the spread of the virus and protest more for a BLM-related cause. Alternatively, counties that are in favor of lax social distancing rules (and thus more aligned with the president's views at the time) are less likely to engage in BLM protests. Additionally, we may be concerned that BLM protests themselves could lead to COVID-19 infections. While we can assuage the latter concern by measuring COVID-19 exposure at baseline (e.g. before the murder of George Floyd and the onset of BLM protests), we address the former concern with an instrumental variable approach.

We exploit plausibly exogenous variation in the occurrence of Super Spreader Events (SSEs) to causally identify the effect of COVID-19 on BLM protests at the county level. Specifically, we construct the IV as the sum of all SSEs that occur within 50 km of the county border but not within the county until 6 weeks before the murder of George Floyd.

---

Floyd, that is the period from May 25th to June 14th for several reasons: we can capture a large share of the protest behavior (66 percent of BLM protests following George Floyd's murder can be observed in this three week window) while limiting the potential for confounding factors to arise. Our results hold when we extend this window to six or eight weeks, or reduce it to two weeks (see Table 3.A.4)

The first stage is written as:

$$Covid_c = \zeta_0 + \zeta_1 Z_{cs} + \mathbf{X}_{cs} \zeta \mathbf{x} + \gamma_c + \eta_{cs}, \quad (3.2)$$

$$Z_c = \sum_{m=1}^{t-6} SSE_{csm}^{neighbor} \quad (3.3)$$

The key identifying assumption of this instrument is that - given the set of controls and state fixed effects - SSEs only affect BLM protests through an increase in exposure to COVID-19. We exploit three features of our IV to argue for the validity of the exclusion restriction: *i*) epidemiological features of super spreader events, specifically small events with one highly infectious person present *ii*) the temporal feature, e.g. the short window of opportunity for SSEs to arise , and *iii*) exposure to SSEs *outside* the county. In section 7.1, we also provide a number of empirical tests to verify the plausibility of the exclusion restriction and probe the robustness of our instrument.

**Event types.** Super Spreader Events are defined as the presence of a highly infectious person (a super spreader) in a context where they can infect a large number of people. Super-spreaders are individuals who are an order of magnitude more contagious than others. This phenomenon, well-known in epidemiology, is instrumental in infectious disease spread (e.g. Galvani and May (2005)) and of particular importance for COVID-19, where 70–80% of transmissions can be traced back to just 10–20% of cases Adam et al. (2020); Endo et al. (2020); Miller et al. (2020). It is important to note that these events do not have to be large gatherings or mass events. The majority of the approximately 1000 SSEs in our data<sup>14</sup> take place in prisons, nursing homes, and at birthday parties. SSEs are characterised by the presence of a highly infectious individual. The size of the event is only relevant insofar as it increases the likelihood of a super-spreading individual being present. Therefore, not all mass gatherings are SSEs and not all SSEs are mass gatherings. This is relevant for the exclusion restriction as far as it alleviates concerns about SSEs being a proxy for a county's propensity to organize large public events, including BLM events. In fact, the overwhelming majority of SSEs is recorded – as expected – in the medical care sector (see Figure 3.5).

**Window of opportunity.** Next, we illustrate in Figure 3.6 that the overwhelming majority of SSEs (solid blue line) occurred between the second week of March and the last week of April. This was an opportune period for SSEs for two main reasons. First, infections were sufficiently high to introduce a significant number of super-spreader individuals.

---

<sup>14</sup>Data recorded by scientists from the London School of Hygiene and Tropical Medicine

Second, lock-down measures were not yet stringent enough (in addition to the lack of public awareness) to restrict group gatherings and encourage mask-wearing. The red dotted line of Figure 3.6 shows that the increase in the number of new COVID-19 cases coincided with the increase in SSEs. The green dashed line illustrates that state-issued stringency measures (as measured by the stringency index from the Oxford COVID-19 Government Response Tracker) peaked around the time that SSEs leveled off. We argue that during this time window, the occurrence of SSEs was mainly driven by the presence of a highly infectious person, rather than heterogeneity in risk preferences or other underlying factors that could drive both SSEs and BLM protests. We only include SSEs until April 13th 2020 - 6 weeks prior to George Floyd’s murder, to account for the fact that SSEs further into the pandemic may be more endogenous. We illustrate in Figure 3.7 that this was well into the pandemic (measured as the cumulative number of COVID-19 related deaths) but sufficiently distanced from the surge in BLM protests and its trigger.

**Geographic proximity.** Lastly, we improve on the plausibility of the exclusion restriction by exploiting SSEs *outside* the county and not within the county. Specifically, we use the number of SSEs within a 50km (or approximately 30 mile) radius from the county border in which we measure exposure to COVID-19 and BLM protests. We illustrate the construction of our instrument in Figure 3.8 using the example of Arizona. To create this instrument, we rely on the geo-location information of the SSEs and county borders. We indicate as red dots the SSEs used for our IV in this illustrating case. We first draw a circle from the location of each super spreader event and then use the SSEs whose circle intersects with the county boundary to instrument COVID-19 deaths. We argue that SSEs in geographic proximity but not in the county itself are even less likely to affect BLM events in the county other than through COVID-19 exposure.

In Figure 3.9 we show the geographical distribution of our instrument across US counties. In the top panel, we map at the county level the cumulative number of SSEs 6 weeks prior to Floyd’s murder. In the bottom panel, we illustrate the identifying variation of our instrument, e.g. the number of SSEs in 50 km proximity to the county border up to April 13th. We present the first stage results in Table 3.C.3. Results show, as expected, that a higher number of cumulative SSEs in a 50km radius of neighbouring counties is related to a higher number of COVID-19 deaths per thousand population. On average, an increase of one additional SSE increases the number of COVID-19 deaths per thousand population by between 0.8 and 1.3 points, depending on the specification. For all specifications the F-statistic is well above the standard threshold.

Overall, the features of our instrument (epidemiological feature, small window of oppor-

tunity, geographic distance) lend confidence to a causal interpretation of our IV estimation. We dedicate subsection section 7.1 and appendix Appendix D to carefully addressing concerns about the validity and robustness of our instrument. Let us preview some of the most important checks here. First, we show that SSEs do not predict past BLM events. Second, we incorporate various weighting schemes and additional control variables to improve the plausibility of the exclusion restriction. Third, we run a number of robustness checks, including varying the distance and window of opportunity for SSEs, excluding SSEs in prisons, and controlling for SSEs in the same county.

## 4. COVID-19 and BLM

### 4.1. Main Results

We present our main results in Table 3.2, showing the OLS and IV results for the full sample (Panel A), the sample of counties without BLM events prior to George Floyd’s murder (Panel B) and the sample of traditional protesters, e.g. those with at least one BLM event before (Panel C). Reduced form regressions are presented in Appendix Table 3.A.1.

Column 1 of Table 3.2 reports the effect of COVID-19 deaths on the probability of observing a BLM protest without state fixed effects or controls. We find consistently strong and positive effects of COVID exposure on protest behavior. In columns 2 to 6, we progressively add state fixed effects, demographic controls (share of Black population and degree of urbanization), economic controls (median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience), and political controls (Republican vote share in 2012 and 2016, social capital, i.e. the number of different types of civic organizations, the number of past BLM events between 2014 and 2019, and deadly force used by police between 2000 and 2019).

Our preferred specification is presented in column 7 and includes state fixed effects and the full set of controls. We find that one additional death per 10 000 population increases the likelihood of at least one BLM event occurring in the three weeks following the death of George Floyd by between 2 and 6 percentage points (p.p.) depending on the specification. An increase of one standard deviation in the number of deaths per thousand increases the likelihood of at least one BLM event occurring by between 5 and 14 p.p.

As shown in Figure 3.4, we observe that more than half the counties that take to the streets in response to George Floyd’s murder have never protested for a BLM-related cause before. In Panels B and C of Table 3.2, we turn to the sub-samples of counties with and without protest history. Focusing on column 7 of Panel B, we find that the effect doubles

in size and is more precisely estimated than the full sample. Specifically, we find that a one standard deviation increase in the number of deaths (25 per 100 000), increases the probability of protesting by 10%. On average, a marginal increase of around 1.2 points in the number of COVID-19 deaths per thousand population in all counties that did not protest before the murder of George Floyd would double the number of counties hosting a first demonstration. In Panel C, we show that traditional protesters are not responding to the exposure to COVID-19, confirming that our baseline result is entirely driven by counties protesting for the first time.

Throughout all of our estimations (including the robustness checks presented in subsection section 7.2) the IV estimates exhibit larger coefficients compared to the OLS. In the absence of exogenous variation in changes to the COVID-19 infectious environment, the OLS underestimates the role of COVID-19 as a trigger for BLM protests. The bias in the OLS could stem from unobserved within state county-level determinants that drive both BLM protests and lower levels of COVID-19 exposure.<sup>15</sup> This could be due to underlying attitudes that disapprove of the Trump administration (beyond those that are captured in the past Republican vote shares and the inclusion of state fixed effects). For instance, more progressive counties, such as Travis county (capital Austin Texas) could be more favorable towards the BLM movement and at the same time more cautious vis a vis the pandemic outbreak and adhere to stricter social distancing rules than Montgomery, Texas. Using mobile phone mobility data, we find that counties that protested for BLM after the murder of George Floyd also decrease their workplace and leisure mobility, while increasing residential stay. This is in line with Dave et al. (2020) who show that BLM protesters adhere more to social distancing measures.

Again, we preview here that our results are robust to changes in the construction of the instrument, treatment and outcome variables, to changes in the sample composition, spatial clustering, and additional controls. We describe all of these checks in section 7.2 and provide greater detail in Appendix Appendix A. In addition, we use three alternative identification strategies to corroborate the results, including the use of an alternative instrument; an instrumented difference-in-difference model and a LASSO propensity matching model. These are summarized in section 7.3 and described in detail in Appendix Appendix B

## 4.2. Heterogeneity

What are the characteristics of counties that start to protest in response to the murder of George Floyd? In Table 3.4, we interact exposure to COVID-19 with baseline characteristics

---

<sup>15</sup>Since the treatment (exposure to COVID) is measured before the protest trigger, reverse causality is not the driver of the difference in magnitude.

for the full sample of counties and report the coefficient of the interacting variable in the bottom row. We analyze heterogeneity over the full sample to identify which baseline county characteristics determine protest in response to George Floyd's murder. We instrument both COVID and the interaction term with our SSE variable, and report the F statistics at the bottom of the table.

In column 1, we show the baseline effect for reference. In columns 2 and 3, we consider heterogeneity by race as recorded in the American Community Survey in 2018.<sup>16</sup> The coefficient of the interacting variable indicates that - as expected - counties with a higher non-Black and non-white population share are less likely to protest overall. This is in line with our prior that those who are most affected by the movement's grievances are typically protesting. However, counties with a higher non-Black population share (including whites, Hispanics, Asians and "others") are more likely to respond to exposure to COVID-19, confirming the idea of a broadening BLM coalition. Interestingly, if we look at the effect of counties with higher non-white population shares (this includes other minorities beyond Blacks), we do not see the same response, indicating that whites are driving the results in column 2.

In column 4, we move to the economic prosperity of the county, as proxied by the median household income - again measured in 2018 from the American Community Survey. Richer counties are more likely to protest overall and these counties protest even more in response to the pandemic. This is in line with two mutually non-exclusive interpretations. First, the literature on protest and conflict highlights that individuals need basic resources to be able to engage in protest in the first place Bates et al. (2002); Bazzi and Blattman (2014); Besley and Persson (2011). Only more affluent households may be able to protest when the resources of other households are depleted due to the pandemic. Second, it is possible that - similar to the non-Black counties in the previous columns - richer counties become aware of racial inequalities through the murder of George Floyd and start to protest in response.

As expected, counties with higher vote shares for Donald Trump in the 2016 elections (vote share Republican reported in column 5) are less likely to participate in BLM protests overall. However, the coefficient of the interaction term is negative, not significant and very noisy, indicating that the political leaning is less relevant for the likelihood of a BLM event occurring in response to higher exposure to COVID-19. Conditional on state fixed effects this may not be surprising, as they capture a large share of the variation in political leaning.

In columns 6 to 9, we consider different classifications for a county's degree of urbanization as defined by the 2013 NCHS Urban-Rural Classification Scheme for Counties. Typically, BLM protests occur in large metropolitan areas, like New York or Los Angeles and less frequently in smaller cities, suburban or rural areas. In column 6, we look at the effect

---

<sup>16</sup>Self reported racial identification with the categories: white, Black, Asian, Hispanic and "other".

of the pandemic on counties that are not part of a large city. This encompasses fairly big sub-urban areas like Bergen County, New Jersey (adjacent to Bronx County in New York) to small rural areas like Mariposa County, California. Similarly, we also consider only suburban counties in column 7. Both of these county types experience an increase in BLM protests in response to the pandemic. Unsurprisingly, small towns and rural areas are less responsive to COVID-19 exposure.

Overall, these results confirm our prior that the pandemic broaden the kind of counties mobilizing for BLM. These recently joined counties are characterized by having a higher share of non-Black and affluent populations and for having a higher probability of being located in suburbs and smaller cities.

We repeat the analysis, now focusing on the sub-sample of counties with no prior BLM protests. While the previous exercise sheds light on heterogeneity in the characteristics of counties that respond to exposure to COVID-19, this analysis excludes traditional protesters and investigates which of the counties join the movement in response to the pandemic, and which counties remain inactive (rather than continue to protest). We present these results in Table 3.C.4 and find similar patterns. While the racial composition of the county points in the same direction (but is more noisy), the effect of income and degree of urbanization become larger and more precisely estimated.

#### *4.3. Alternative Outcomes*

Our main variable of interest, so far, was the likelihood of observing any BLM protest in the three weeks following the murder of George Floyd. In Table 3.3, we consider the frequency and scope of BLM protests and include other forms of political expression, including online protest and street art.

We report the baseline result for the sub-sample of counties with no prior BLM events in column 1. In columns 2 to 4 we look at the structure of these protests, investigating the number of BLM events in the three week window, as well as the total number of protesters and the average number of protesters per event.

In columns 3 and 4, we look at the total number of participants and the average number of participants, again including counties with no BLM events as zeros. We find negative but non significant and very noisy estimates for the effect of COVID-19 on both measures for the scope of BLM protests. We conclude that the pandemic increase the likelihood and frequency of BLM protest without significantly impacting its scope.

Next, we investigate the impact on online protest. In column 5, we report as an outcome the total number of geo-localized tweets in a county in the three weeks following George

Floyd's murder. These are based on the universe of tweets that use the hashtags #BlackLivesMatter #BlackLifeMatters or #BLM. We find a large effect of pandemic exposure on the number of BLM tweets. Our coefficient is eight times the average number of BLM related tweets in the full sample. In addition, we scrape information on all followers of the official BLM account and geo-localize each Twitter user. We find that places that were more exposed to the pandemic started following the BLM account in greater numbers. This has potential implications for the medium-run mobilizing potential of the movement. The official twitter account serves as a primary coordination, communication and mobilization tool for BLM Black Lives Matter (2020). Therefore, the expansion of the follower base may help activate these groups, when similar protest triggers arise in the future.

While protests on the streets and online may have a low barrier to entry, there are other forms of political expression that require more cultural or political capital. For instance, street art (and art more generally) has become a major form of advocacy in anti-racist movements Cappelli et al. (2020); Mathieu (2018) but is not as accessible and is harder to replicate among counties that are new in hosting BLM events. We geo-locate street art containing references to Black Lives Matter and George Floyd from the *Urban Art Mapping George Floyd and Anti-Racist Street Art* database. In line with our priors, newly mobilized counties can mobilize in the arena of online and offline protest but cannot quickly replicate forms of protest that are more deeply rooted in the BLM movement.

## 5. Social Media and BLM

### 5.1. COVID-19 and the Use of Social Media

*Average monetizable DAU [daily active users] grew 24% year over year... The increase in mDAU was driven by ... an increased engagement due to the COVID-19 pandemic.*

Twitter letter to shareholders of April 30th 2020

The literature on the effect of social media on protest and other political outcomes exploits supply side constraints to the access to social media, typically leveraging a version of a staggered roll-out design Enikolopov et al. (2020); Manacorda and Tesei (2020); Müller and Schwarz (2020). These approaches go back to the early 2000s and become less relevant as social media becomes widely accessible. In this paper, we hypothesize that the pandemic shifted a substantial proportion of communication and social interactions to the digital spaced. More specifically, we argue that the pandemic acted as a demand shock to social media, particularly Twitter. In this section, we will show that the pandemic-induced uptake in social media happened disproportionately in areas with no BLM history. We argue that these

“non-traditional” users were then exposed to an unexpected and highly viral protest trigger - the murder of George Floyd - which in turn mobilized them to take to the streets for the first time during the pandemic.

To further motivate this prior, we show some descriptive figures in Appendix Figure 3.C.1 and Figure 3.C.2. We see that in the period prior to the protest trigger, the mean stringency of social distancing and lockdown measures (as proxied by the Oxford Government Response Tracker) increased substantially. Measures mostly included recommendations to socially distance (interestingly, mask wearing recommendations - a sub-category in this index - only started many weeks later). In Figure 3.C.2, we use Google mobility data and show that residential stay increased, whereas other types of mobility (particularly, work, transit, and retail) decreased substantially. This already points to a probable decrease in social activities and an increase in online activities between March and May. Moreover, many online services reported substantial increases in the number of users during the first months of the pandemic. For instance, Netflix attributed 16 million new subscribers to lockdown measures<sup>17</sup> and TikTok experienced growth of 180 percent during the pandemic<sup>18</sup>.

To test this hypothesis more systematically, we create a novel index of social media penetration that comprises the first principle component of four main variables (plus the log of two of them)<sup>19</sup>: *i*) the (log) cumulative number of new Twitter accounts, which we obtain by scraping and geo-coding information on the creation date of new Twitter accounts at the county level from approximately 45 million tweets; *ii*) the (log) number of new followers of the official BLM Twitter account, which we obtain by scraping the BLM account followers, identifying their creation date and localizing them; *iii*) the normalized index of search activity for term ‘Twitter’ provided by Google Trends, hypothesizing that new users will Google the term and then create an account<sup>20</sup>; and *iv*) Google mobility data at the county level, assuming that increased residential stay (time spent at home) as well as lower social, work and leisure mobility is associated with more time spent online.

All of these variables are measured between January 2020 and May 24th 2020, i.e. after

---

<sup>17</sup><https://www.bbc.com/news/business-52376022>

<sup>18</sup>TikTok Usage

<sup>19</sup>We include both the absolute number of accounts and the log number of accounts (new Twitter accounts and new BLM followers) for two reasons. On the one hand, we do not have a prior as to whether the absolute number Twitter users or share of Twitter users is important for the occurrence of a BLM event. It is possible that irrespective of county size or Twitter penetration at the county level, there is a threshold level of individuals that need to be mobilized for a BLM event to occur. The average number of protesters at a BLM event in counties with no prior BLM events is about 350 individuals. On the other hand, in the absence of a good measure for relative importance of Twitter (by population, baseline Twitter usage, overall social media users) we want to give less weight to counties with higher Twitter penetration. Including both in the principle component will allow us to account for distributional features of Twitter penetration. The principle component will only capture the residual correlation between the two variables.

<sup>20</sup>The Google Trends data is defined on a designated market area (DMA) level

the outbreak of the pandemic but before the murder of George Floyd. We limit the observation period, such that the BLM events themselves do not impact online activity but we are still able to observe the pandemic-induced increase in online activity. We show the features of our index in Table 3.C.7, presenting the correlation between the different sub-components in Panel a), the eigenvalues of the principle components in Panel b) and the factor loadings in Panel c).

In Table 3.5, we show the results for the full sample (Panel A), counties with no BLM events before George Floyd's murder (Panel B) and counties with prior BLM events (Panel C). Again, we use the instrumented exposure to cumulative COVID-19 deaths per 1000 population until May 24th as a main explanatory variable. In column 1, we confirm that the pandemic has led to an increase in online activity as measured by our index for social media penetration. Importantly, the effect is 10 times as large and more precisely estimated for the subset of counties with no prior BLM protest history.

We then zoom into the specific sub-components of the index and find in column 2 that increased exposure to the pandemic had no effect on the raw number of new Twitter accounts created until May 24 (just before George Floyd's murder) for the full sample, or the sample of traditional protesters, but is large and significantly positive for the sub-sample of counties with no prior BLM events. When we consider the log of new Twitter accounts in column 3, we find an even stronger effect for the sub-sample of counties with no BLM before George Floyd's murder.

Focusing on Twitter search terms on Google as an additional proxy for the use of Twitter in column 4, we find that - again - search terms only significantly increased among counties with no prior BLM events. Then we show residential stay (column 5), using Google mobility data at the county level in the month leading up to George Floyd's murder and find that for all samples there has been an increase in residential stay - and more so among counties with no prior BLM events. Lastly, we find a positive but noisy effect of COVID-19 on the number of new BLM followers and no effect on the log number of new followers. This is possibly due to a noisy measure of BLM followers as we scrape this information in February 2022, when many accounts may have been deleted or have unfollowed the BLM account.

Taken together, these results show, consistent with our prior, that the pandemic has increased online activity and particularly the use of Twitter - but only among those counties that never protested for a BLM-related cause before. It is important to note again that we measure this online activity cumulatively at the day of George Floyd's murder, capturing the pandemic-induced increase in social media use and excluding the effect of George Floyd's murder on social media use directly. The pandemic acted as a demand shock to social media in areas with lower prior BLM salience.

## 5.2. Twitter and BLM protests

In the previous subsection, we have established that the pandemic is associated with higher online activities. Importantly, this is driven by the sub-sample of counties with no prior BLM protest, which are also those that start to protest in response to the pandemic. In this section, we establish a more direct link between online activity, particularly Twitter usage, and protest behavior.

However, it is possible that among the sub-sample of counties with no prior BLM protest, those counties that experienced an increase in social media uptake are not the same as those where protests occurred. Therefore, we interact different measures of Twitter penetration (we detail the construction of this variable in Appendix D) with (instrumented) exposure to COVID-19 to see whether within the sub-sample of counties with no prior BLM protest. We caveat now that baseline Twitter penetration may be related to unobserved factors that co-determine BLM protests. Additionally, new Twitter accounts are a bad control as they are co-determined by exposure to COVID-19. We will address this point in the subsequent analysis but focus, for now, on the following heterogeneity. We estimate a second stage regression of the form:

$$\begin{aligned} BLM_{cs} = & \beta_0 + \widehat{\beta_1 Covid_c} + \beta_2 Twitter_c \\ & + \widehat{\beta_3 Covid_c \times Twitter_c} \\ & + \mathbf{X}_c \beta_{\mathbf{X}} + \delta_s + \epsilon_{cs} \end{aligned} \tag{3.4}$$

where  $Twitter_c$  is either (i) the number of users posting about BLM registered in 2020 before May 24 in county  $c$  of state  $s$ , or (ii) the number of users from the county observed in a sample of tweets collected on December 2019. The logarithm of this number (plus one, to avoid missing values) is interacted with COVID 19 deaths per 1000 population.<sup>21</sup> We instrument COVID-19 deaths and their interaction with users by SSEs and their interaction with  $Twitter_c$ .

We present results in Table 3.6 for the sample of counties with no prior BLM protests. In column 1, we show the interaction effect between instrumented COVID-19 and baseline twitter penetration, measured as the log number of users in December 2019. We find that the effect of COVID-19 entirely runs through counties with higher levels of baseline users. The baseline effect of both COVID-19 and baseline Twitter penetration are insignificant. In column 2, we repeat the same exercise, this time interacting instrumented COVID-19 with

---

<sup>21</sup>We use the logarithm instead of the actual number of tweets to overcome potential problems with outliers

the log number of new accounts created during the pandemic. Remarkably, we find a positive and significant coefficient of almost identical magnitude. We take this as first indicative evidence that baseline penetration combined with COVID-19 exposure is a major predictor of new users in the pandemic. This is in line with the literature on the path dependence in technology adoption Arrow (2000); Arthur (1989); Liebowitz and Margolis (1999); Müller and Schwarz (2020). The marginal utility of joining a social network increases with the size of the existing network. Therefore, it may be unsurprising that the pandemic induced increase in the use of social media operates through the sub-sample of counties with sufficiently large baseline network size.

In Table 3.7 we repeat the analysis for the subsample of counties that had already hosted a BLM event before the murder of George Floyd. Results show no differential effect of COVID-19 on protest neither in counties with higher baseline Twitter penetration (column 1), nor in counties with more new Twitter accounts created during the pandemic (column 2) for this subsample.

The different results of this exercise for the subsample with and without previous BLM protest suggest that exposure to the George Floyd murder and the following reaction though social media is important in the fractions of the population that are not yet conscious of the problems faced by Black people and of systemic racism more generally. As shown in previous sections, counties without previous BLM events are generally whiter, richer and less urban. It is not surprising that people living in whiter, richer and less urban areas have been less exposed (directly or indirectly) to the problem of racial inequality. Indeed, Black people do not need external input to learn about racial inequality, and people who live in counties that already hosted a BLM event are more likely to have already been exposed to narratives highlighting the problem. This exposure could have happened through different channels, and notably through BLM protest themselves as protests can serve as information shocks Lohmann (1994).

As cautioned some paragraphs above, these results cannot be interpreted causally: while we have an instrument for COVID-19, the number of pre-existing and new Twitter users is endogenous and potentially correlated with the error term. Even with the fixed effects and various controls, Twitter usage at baseline could be driving BLM protest differentially for counties with higher COVID-19 exposure.

To address this concern, we instrument pre-pandemic Twitter penetration in December 2019. Specifically, we reproduce the SXSW instrument for Twitter usage described by Müller and Schwarz (2019). SXSW (South by Southwest) is an annual festival in Austin, Texas. During the March 2007 edition Twitter was heavily promoted, leading to a rapid increase in the social network's popularity. To reproduce this instrument, we collect the location of

all followers of the @SXSW account of the South by Southwest festival and the date they joined Twitter.

The dataset we end up with is not entirely identical: some users created on or before March 2007 might have started or stopped following SXSW later. They might also have changed their location between the time Müller and Schwarz collected their dataset and when we collected ours (2019 versus November 2021). Finally, our geolocation method might be different.<sup>22</sup>

Following Müller and Schwarz (2020), we compute for each county the number of followers whose account was created in March 2007 and the number of users whose account was created before this date. With our data collection and user localization strategy, this leads to users being located in 172 counties, only 67 of which did not have BLM events before (Müller and Schwarz find 155 affected counties). To increase the number of treated counties, and thus the power of our identification, we also consider users in neighboring counties created during this period: assuming that Twitter presence diffuses geographically in part (again following the Müller and Schwarz approach), these counties should also have a higher number of Twitter users. We find 817 such counties, 618 of which did not have a BLM protest before.

We estimate the log number of observed Twitter users in December 2019 using the number of users that joined during SXSW controlled by the number of SXSW followers that joined before,<sup>23</sup> with the following regression:

$$\begin{aligned} Users_c = & \xi_0 + \xi_1 SXSW\ Users_{sc} + \xi_2 Pre\ SXSW\ Users_{sc} \\ & + \mathbf{X}_c \xi \mathbf{x} + \gamma_s + \eta_{cs} \end{aligned} \tag{3.5}$$

where SXSW Users<sub>sc</sub> is the log number of SXSW followers who created their account in March 2007 in the county and neighboring counties, and Pre SXSW Users<sub>sc</sub> is the log number of SXSW followers in the county and neighboring counties that created their account before March 2007.

For the subsample of counties without BLM event before George Floyd's murder, the results of this first stage regression are reported in Appendix Table 3.C.6. The coefficient of SXSW users is positive and highly significant, and the first stage is strong ( $F = 13.02$ ). We re-run the above specification, this time instrumenting pre-existing Twitter users by the SXSW instrument. The results for the second stage are presented in column 3 of Table 3.6 We report

---

<sup>22</sup>We automatically geocode the location given by the user using Nominatim, as described in the Data section. Müller and Schwarz (2019) do not detail their geolocation method. Fujiwara et al. (2021) indicates that 58% of users that joined between 2006 and 2008 are geocoded; we attribute 52% of users to US counties (excluding imprecise locations and locations outside the US).

<sup>23</sup>This variable controls for the interest in SXSW festival and also acts as a proxy control for the general interest in Twitter in the county.

the per-coefficient F statistic of weak identification following Sanderson and Windmeijer (2016).

While this result supports our hypothesis, it needs to be interpreted with caution. First, though we focus on new users, we do not observe the extensive margin of Twitter usage: our collection method only allows us to observe users that actually post on Twitter or retweet existing posts, but not users that only read and like tweets. In this way the results that we capture underestimate the effect exposure to social media had on BLM protests.

Second, we cannot disentangle the effects of having higher share of users at the baseline from the effect of additional people joining Twitter due to the pandemic. These two measures may be related. On one hand, they are related by network effects. COVID-19 created a shock increasing the demand for online activities, including social media. Potential social media users faced a choice between different options of online activity to adopt. The likelihood of adopting Twitter is higher for people that know a number of friends using Twitter, both because they can use it to communicate with their friends, but also because it is more likely that these friends share interesting tweets through other channels. On the other hand, there is also a saturation effect: a higher Twitter penetration pre-pandemic in a county reduces the potential number of people that can join Twitter or use it more. This is not likely to be the case here, as the extra Twitter usage derives from a demand shock, where users have more time to spend on Twitter, instead of an offer shock. Moreover, Twitter reports only 77 million users in January 2022 in the United States, while for instance Facebook reports 180 millions which makes an absolute saturation (i.e. saturation driven not by users' maximum level of willingness to join or spend time on social media but by the absolute availability of time or of new possible users) unlikely.

Despite the limitations discussed above, we can interpret these results as suggestive evidence that social media (either at baseline or its increase in usage during the pandemic) played a crucial role in mobilizing for BLM in counties whose population is less likely to have been exposed to narratives that denounce the presence of racial inequality and discrimination (i.e. whiter, richer, less urban counties that have not hosted any BLM event before).

### 5.3. *News Consumption and Attitudes towards BLM*

In this subsection we examine the social media mechanisms more closely by exploiting individual-level survey data. We ask whether exposure to COVID-19 at the individual level caused a shift in news consumption away from traditional media and towards social media. We then investigate whether this shift is accompanied by a change in attitudes towards Blacks and the Black Lives Matter movement more generally.

It is important to note that a causal interpretation of these results is not possible, as we do not have precise information on the location of the respondent; we only have information on the severity of exposure to COVID-19 in their county of residence at the time of the interview in June 2020. However, the rich set of individual-level controls and placebo checks assuage concerns about omitted variable bias.

We use survey data from the Pew Research Center to conduct individual-level multivariate regressions on different outcomes, controlling for respondent characteristics: race, whether or not they live in a metropolitan area, gender, age, education, income and whether or not they lean towards the Democratic party. Table 3.8 shows the results. Columns 1 - 3 show the intensity and form of news consumption in the context of George Floyd's murder. Higher levels of COVID-19 are positively and significantly associated with more news consumption about George Floyd and more social media news consumption about George Floyd. In column 3, we show that individuals in counties with higher COVID-19 exposure also consume relatively more news about George Floyd on social media, confirming a change in the information set - or at least their source.

Then, we analyze whether this change in mode of news consumption is accompanied by a change in attitudes. In column 4, we find that individuals are more likely to report that higher hospitalization rates of Blacks during the pandemic are caused by circumstances beyond their control, rather than personal choices or lifestyle. Respondents are also more likely to agree with the statement that the BLM protests arise because of structural racism and not as an excuse for criminal behavior. To rule out that exposure to COVID-19 in the earlier stages of the pandemic is just a proxy for more progressive leaning counties, we use an additional question that deals with an unrelated progressive issue: legal status for undocumented immigrants. Individuals living in counties with higher exposure to COVID-19 are not more likely to prefer more rights for undocumented immigrants, alleviating some of the concern about unobserved heterogeneity.

## 6. Competing Mechanisms

In this section, we consider alternative (non-exclusive) mechanisms for the pandemic-induced increase in BLM protests, considering *i*) a scattering rather than a broadening of protest *ii*) pandemic-induced salience of racial inequality *iii*) lower opportunity costs of protesting and *iv*) increased overall agitation and propensity to protest.

## *6.1. Broadening versus Scattering of Protest*

In this section we discuss the possibility that spatial spillovers from BLM protest (say, from the cities to the suburbs) are driving our results. Specifically, we investigate whether the observed broadening of the coalition is in fact just a substitution of protesters in time and space. In fact, it is possible that we observe new counties protesting for reasons unrelated to the idea of an increase in allyship for the BLM movement. First, the pandemic may have changed the scope and structure of BLM protests (smaller but more numerous). Second, neighboring counties may inspire future protest in close proximity.<sup>24</sup> Third, the pandemic and its restrictions on mobility may have led to a geographic spread of the protest movement, substituting large protests in cities with smaller protests in suburbs. We address the concern that the pandemic may have simply led to a substitution of protest locations and frequencies, rather than a true broadening of sympathizers.

**Number of participants and protests.** If the observed increase in the number of counties hosting a BLM event for the first time after George Floyd’s murder is simply driven by a substitution of protest across space (e.g. re-location of protesters themselves or creation of multiple smaller protest events), we should observe that the number of protests increases while the number of participants should decrease. We show in columns 2 to 4 of Table 3.3 that neither is the case. We take this as first evidence that the pandemic does not change the structure of these protests.

Moreover, we consider the possibility that individuals who protest might, in response to the pandemic, decide to protest closer to home and not protest in the city center of the neighboring county. For instance, protesters could be affected by restrictions and closures of public transport, preventing them from going to a demonstration further away. They might also consider that a smaller, more local demonstration is safer, as they would come into contact with fewer people, limiting the risk of spreading coronavirus between communities.

**Traditional protesters as neighbors.** While we should pick up some of this in the number of participants and protests in the previous analysis, we test this more systematically by constructing a dummy variable equal to one if one of the county’s neighbors is a ”traditional protester” (e.g. had a BLM related protest before May 25th 2020). We use this variable in two ways. First we include it as an additional control (column 1 of Table 3.9) and second we interact this dummy variable with COVID-19 deaths per 1000 population (column 2 of Table 3.9). Results show that having a traditional protester as a neighbor does not increase

---

<sup>24</sup>If SSEs and BLM protests themselves have spill-over effects, we may falsely attribute an increase in protest to the pandemic.

the probability of protesting overall within the sample of counties that had never protested before. More importantly, the interaction term between exposure to COVID-19 and having a traditional protester as a neighbor in column 2 is not significant, and if anything reduces the likelihood of protesting in response to the pandemic. This seems to indicate that the displacement effect is not a driver of our results.

**Recent protesters as neighbors.** Lastly, it is possible that protests in one county could inspire protests in neighboring counties over time. While this would not go against the idea of a broadening BLM coalition, it indicates that protests during the pandemic inspire subsequent protests in neighboring counties. We therefore construct an indicator similar to the "traditional protester as neighbor" but apply this to the period after George Floyd's murder. More specifically, we construct a dummy variable that indicates whether the county has a neighboring county that protested *before* they start to protest. This allows us - even in our cross-sectional setup - to account for spillovers in time. However, this approach suffers from an important caveat: protests in neighboring counties during the pandemic could be endogenous and therefore a bad control. We consider these effects in columns 3 and 4 of Table 3.9 with these caveats in mind.

If spillovers exist, we would expect that having a neighboring county that recently protested increases the likelihood of observing a protest yourself. We include this variable as a control in columns 3 of Table 3.9 and find no change in our results. In column 4 we interact this variable with COVID-19 deaths per 1000 population and find that the effect of COVID-19 on the likelihood of protest is not higher among counties who's neighbours protested before. This suggests that these temporal spillovers across neighboring counties are not driving our main results.

Lastly, we analyze the geographic diffusion of protest. The viral video footage of police officer Derek Chauvin murdering George Floyd inspired large scale protest in the city, starting the day after the murder on May 26th 2020. President Trump infamously tweeted that "when the looting starts the shooting starts", referring to the escalation of protests in Minneapolis on May 27th. Minneapolis quickly became one of the main focal points in the Black Lives Matter movement. In columns 5 and 6 of Table 3.9, we investigate whether proximity to the earliest and largest protest hub affected protest behavior. We use the distance and squared distance to Minneapolis and find no significant impact of proximity to Minneapolis. If anything, counties further away may respond slightly more to COVID-19 exposure, with the caveat that the first stage of the interaction term becomes weak in column 6.

Overall, we take these results as evidence that the observed spread of the BLM protest is a true broadening of the BLM movement and not driven by the spread of existing protesters.

In addition, we also find no evidence for learning or imitation through time and space. We argue that this is also consistent with the use of social media mechanism because exposure to the protest trigger on social media is much less dependent on learning over time or through geographic proximity.

## 6.2. *Salience of Racial Inequality*

The second alternative mechanism we test is a rise in the salience of racial inequality due to the pandemic itself and not through exposure to BLM-related content online. For instance, an *a priori* indiscriminate virus should affect whites and Blacks equally but if there are racial disparities in death rates, then people may be more inclined to believe that there are systemic disadvantages afflicting the Black community. We test this mechanism in two ways. First, we hypothesize that if this mechanism is at place, counties facing a higher proportion of Black deaths due to COVID-19 (respect to the total proportion of COVID-19 deaths) would be more likely to protest after the trigger of George Floyd's death. Column 1 of Table 3.10 shows the estimate of the interaction term between COVID-19 death per 1000 population and the Black death burden<sup>25</sup>. Results show that the effect on COVID-19 on protest is not higher in counties with relatively more death burden of Blacks.

Additionally, we test whether the results are driven by an increase in the awareness and sympathy towards BLM-related issues during the pandemic but *before* the murder of George Floyd. We hypothesize that if people are empathizing with problems faced by the Black community because of the pandemic itself, we would observe an increase in interest towards BLM already *before* the murder of George Floyd. If this is the mechanism driving our results, counties that have gained awareness about BLM-related issues *before* the murder of George Floyd would be the ones that protest the most after the murder of George Floyd. We test this in column 2 of Table 3.10, where we interact the relative popularity of BLM search terms on Google in the month leading up to George Floyd's murder with the number of COVID-19 deaths per 1000 population. We do not find that an increased interest in racial injustice before the protest trigger (measured with BLM Google searches) increased the probability of a demonstration.

Overall, we do not find that an increase in sympathy or interest towards BLM-related issues before the murder can explain the effect of COVID-19 on BLM protest following George Floyd's death.

---

<sup>25</sup>Black death burden is computed as the ratio of the Black COVID-19 deaths per 1000 Black population over the total COVID-19 deaths per 1000 population

### 6.3. Opportunity Cost of Protesting

Next, we test whether the results can be explained by a decrease in the opportunity cost of protesting. It is possible that new people joined the movement because they had a lower opportunity cost of protesting during the pandemic. We consider two possible channels.

First a decrease in the overall opportunity cost of protesting can be due to a decrease in employment and economic opportunities due to the pandemic. According to Bureau of Labor Statistics (2020): “in June 2020, 40.4 million people reported that they had been unable to work at some point in the last 4 weeks because their employer closed or lost business due to the coronavirus pandemic —that is, they did not work at all or worked fewer hours” which “represented 16 percent of the civilian non institutional population”. We proxy the decrease of economic opportunity cost using the unemployment rate before the murder of George Floyd. Column 3 of Table 3.10 shows the interaction between unemployment and COVID-19 deaths per 1000 population. Results show that the effect of COVID-19 on protest is not higher in counties with higher unemployment rate.

Second, we consider the decrease of the *social* opportunity costs as a possible channel. An alternative use of the time spent protesting, could *a priori* be spent in social and leisure activities like going to a restaurant or to the cinema. Lockdown and social distancing measures made those alternatives uses of time not available, decreasing the *social* opportunity cost of protesting. We proxy the decrease of *social* opportunity cost with the stringency of social distancing measures at the state level. Columns 4 of Table 3.10 shows the interaction between the stringency of social distancing measures and COVID-19 deaths per 1000 population. Results show that effect of COVID-19 on protest is not higher in counties having stricter lock-down and social distancing measures.

### 6.4. Agitation and Propensity to Protest

Lastly, we investigate whether COVID-19 has increased agitation in the public space generally. It is possible that the increase we find in protest is due to an increased general agitation and discontent and has nothing to do with BLM itself. We therefore look at the effect on other protests, using the ACLED US Crisis Monitor protest data. We exclude BLM-related protests from this data set and expand the observation period to 3 months after George Floyd’s murder to make sure we do not capture a substitution effect between BLM protests and other protests immediately after the BLM protest trigger. We report the results in column 5 of Table 3.10; we do not find an effect of COVID-19 on other protests. This remains true (column 6) even if we consider only COVID-19 related protests (which are largely comprised of anti-mask protests). Additionally, we verify whether the pandemic also

mobilized the counter-movement to BLM. Two of the most popular hashtags in opposition to BLM were `#AllLivesMatter` and `#BlueLivesMatter`. We show in columns 7 and 8 that the pandemic did not lead to a counter-mobilization on Twitter.

## 7. Robustness

In this section, we describe the large set of robustness checks we conduct. We first consider and test various possible threats to the validity of our instrument and the identification assumption. We then move to a brief description of the battery of robustness check we conduct to further validate the main results of this paper. We expand the discussion of the different checks for the instrument and the main results in Appendix A. Finally, we present the three different alternative identification strategies we conduct that we explain more in detail in Appendix B.

### 7.1. Instrument validity

We provide various checks to probe the validity of the identification assumption in Table 3.A.4. Specifically, we investigate whether - despite the features of our instrument described above - SSEs capture some underlying factors that co-determine BLM protests. We always present results for the full sample and the sub-sample of counties that never experienced a BLM protest before. Firstly and importantly, we show that SSEs in neighboring counties do not predict the likelihood of past BLM events between 2014 and 2019. If our instrument was related to some unobserved heterogeneity that drives BLM events, we should observe a direct effect of SSEs on past BLM events. Reassuringly, this is not the case.

In addition, we consider the following possibility: the likelihood of being treated by our instrument is not the same across all counties. For instance, counties neighboring large cities may have a higher probability of having an SSE in close proximity.

This heterogeneity in the probability of being treated could be related to certain county characteristics that relate to their intrinsic probability of participating in a BLM protest. We address this issue by weighting each observation (i.e. each county) by their inverse probability of being treated, using LASSO.<sup>26</sup> In doing so, we give more weight to counties that had a low a-priori likelihood of being treated by the instrument. As shown in Appendix Table 3.B.3, this weighting procedure does not change our results, further alleviating concerns about a violation of the exclusion restriction.

---

<sup>26</sup>We describe this approach in more detail in Appendix section B.3

Lastly, we expand on the idea of controlling for overall BLM protest probability, beyond the important but simple (discrete) measure of past BLM protests. Using LASSO, we select the subset of relevant county-level variables that determine past BLM events and create a propensity score for protesting, based on the selection of these variables.<sup>27</sup> This gives us a continuous measure of protest probability that also covers counties that did not end up protesting for a BLM-related cause in the past, despite having all the features typically associated with protesters. We include this variable as an additional control in column 3 of Table 3.A.5 and confirm that our results remain robust to the inclusion of this variable. Finally, we group counties in sets of 10, 100 and 1000 with similar propensity to protest and add a group fixed effect (Column 4 to 6 of Table 3.A.5).

We probe the robustness of our instrument in Appendix Table 3.A.2 and Table 3.A.3 (Appendix A provides a more detailed description of these exercises). We report the first stage coefficient of our preferred specification where the instrument is the cumulative number of SSEs in neighbouring counties within a 50km radius up to 6 weeks prior to the murder of George Floyd. We include the full set of fixed effects and controls as specified in our baseline estimation. In the top panel, we show results for the full sample; in the bottom panel we focus on the sub-sample of counties with no prior BLM protests. We show both the coefficient for SSEs on COVID-19 ("first stage coefficient") and the second stage results (IV: COVID). In this section, we focus on the first stage robustness but preview that our second stage is largely robust to these changes.

In column 1 of Table 3.A.2, we show that one additional SSE increases the number of COVID-19 deaths by 0.93 per 100 000 population for the full sample. The first stage F statistics lie well above the conventional threshold (Kleibergen-Paap F of 36) and find a slightly smaller coefficient and a weaker first stage (Kleibergen-Paap F of 27) for the sub-sample of counties that have never protested before. In columns 2 to 4, we consider the baseline time lag of 6 weeks, i.e. SSEs until April 13th 2020, but vary the distance to the border between 25km and 200km. Our results hold but as expected, the coefficient decreases and the first stage becomes weaker if we move too far from the county border. Next, we use the number of cases associated with SSEs and our results largely hold. Then, we keep the 50km distance but vary the time lag of SSEs until the protest trigger, reducing it to five weeks and expanding it to seven and eight weeks in columns 6 to 8 and our results hold as well.

In Appendix Table 3.A.3, we continue our robustness checks. Again, we report our baseline in column 1. In column 2, we exclude SSEs in prisons as they may impact the public perception of exposure to the pandemic differently and may also be related to factors

---

<sup>27</sup>We describe this approach in more detail in Appendix section B.3

that drive BLM protests. Next, in column 3, we also include the number of SSEs in-county to account for correlation between neighboring and own SSEs. Then we consider the specific distance to the geo-located SSE. We include both the simple linear distance and squared distance to the SSE in columns 4 and 5. Then, we also consider the extent of the overlap of the 50km radius and the county's territory in column 6. Our results remain robust to changes in the definition of the instrument.

## *7.2. Robustness of main results*

In the previous section, we have provided an array of checks on the plausibility of the exclusion restriction and robustness of our instrument to changes in definition (in the first stage and reduced form). We describe these in more detail in Appendix A. In the top row of each panel of Appendix Table 3.A.2 and Table 3.A.3, we show the second stage results and - reassuringly - find consistent results throughout. The coefficient of COVID-19 on the likelihood of BLM protests among counties with no prior BLM history remains positive, significant and similar in magnitude.

We now move on to the robustness of our results to changes in sample composition, spatial correlation, and definition of the treatment and outcome variables. First, in columns 3 and 4 of Table 3.A.4, we exclude counties and whole states on the coasts and our results hold. We do this for two reasons: first, counties and states next to the ocean will mechanically have fewer neighboring counties with SSEs. Second, when thinking about a "broadening" of the BLM coalition, we want to verify that this does not just apply to states with pre-existing progressive leanings. In columns 5 to 7, we shorten the time horizon to 2 weeks and to 6 and 8 weeks after the murder of George Floyd. In column 8, we use COVID-19 related cases, instead of deaths. The last column includes, as an additional control, the number of COVID-19 related deaths in the past seven days. This is designed to account for heterogeneity in the trajectory of the COVID-19 pandemic when cumulative deaths over the whole period are similar. All of these checks yield consistent results. We provide further robustness checks in Table 3.A.5. In column 2, we run an IV Probit regression instead of a 2SLS. In column 3, we include as an additional control the pre-pandemic protest probability, which we derive from the LASSO matching strategy which we outline in more detail in Appendix Appendix A. In columns 4 to 6, we include fixed effects to compare counties with similar pre-pandemic protest probabilities, in 3 groups (with 1000 counties each), 30 groups (with 100 counties each) and 300 groups (with 1000 counties each). In columns 7 and 8, we replace the state clustering with spatial clustering, allowing correlation in a 50 km radius for column 7, and between neighbors for column 8. Column 9 omits clustering altogether. Reassuringly, our

results are not sensitive to these changes.

### 7.3. Alternative Identification Strategies

We complement our preferred estimation strategy in three ways: i) we design an alternative instrument ii) we exploit the panel dimension of our data set to estimate an instrumented difference-in-differences model and iii) we perform a LASSO matching approach comparing counties with a similar pre-pandemic protest probability. We give a brief summary of the approaches here and describe the strategies in more detail in Appendix Appendix B. All of these approaches confirm the baseline results.

#### *Alternative Instrument: Florida Spring Break*

Instead of collecting information on multiple independent SSEs as in the previous section, we now focus on one single, large-scale event known to have contributed substantially to the spread of COVID-19, the Florida Spring Break in March of 2020 Mangrum and Niekamp (2020). We use *SafeGraph* mobile phone data with over 45 million data entries to identify spring break tourists and their home counties and calculate the share of devices that were present at one of the main spring break beaches in March of 2020 relative to all devices of the origin county. As expected, the first stage for this instrument (reported in Table 3.B.1) is below the conventional threshold. When we include the full set of controls the F-Stats become weak but the results qualitatively hold.

#### *Difference-in-Differences: Notable Deaths Sample*

We expand our data set and include BLM events at the county-week level starting in 2014. We scrape information on all police-related deaths of Blacks since July 2014 that were covered in a major national newspaper like the Washington Post, that were covered on TV by CNN and/or have a dedicated Wikipedia page. We include county and state-week fixed effects to account for all time-invariant county level heterogeneity and common time-varying characteristics at the state level. We interact these "Notable Deaths" (time variation) with the instrumented exposure to COVID-19 (county variation). In this instrumented difference-in-differences approach, we exploit differences in protest behavior following a "notable" death in the presence and absence of COVID-19. We show the results in Table 3.B.2 and we find a sufficiently strong first stage and a strongly significant effect consistent with our baseline results.

### *LASSO Matching: Propensity to Protest*

We additionally exploit the previously constructed dataset of notable deaths and BLM events to construct a measure of the propensity of a county to protest after a notable death. The controls used in the model are selected using LASSO logit regression. We use this propensity measure to construct a matching of counties with and without COVID-19 deaths and with a similar propensity to protest. The results (presented in Table 3.B.3) are highly significant and consistent with our baseline results.

## 8. Conclusion

Protests are an important tool for bringing about social change and holding politicians and institutions accountable. Particularly in the context of minority rights, social movements have to rely on allies to put pressure on decision makers and translate demands into legislation, social and institutional change. However, the way to build and broaden allyship in modern social movements is still poorly understood.

In this paper, we shed light on the role of social media in generating mobilization in counties whose characteristics are closer to the median voter and where a larger part of the population is not directly impacted by the movement's grievances. We first document that around half of the protests following George Floyd's murder occur in counties that are hosting a BLM event for the first time. We next show that exposure to the pandemic increased protest behavior and that this effect is driven by those counties hosting a protest for the first time. We then turn to the study of the role of social media in explaining this effect. We first present evidence showing that the pandemic lead to an increase in the time spent on online activities and in the use of social media in all counties, and more so in counties hosting their BLM first event after George Floyd's murder. Then, we show that counties where social media was more widely used at the beginning of the pandemic and counties where a higher number of new Twitter users were created during the pandemic show a higher effect of COVID-19 on their protest behaviour. This differential effect is only present in counties with no prior BLM-related protest activity, which suggests that exposure to social media content related to a protest trigger can increase mobilization in parts of the population that were not yet conscious of the problems faced by the aggrieved minority.

Our research highlights the importance of social movements' online presence. Exogenous changes in the use of social media may increase political mobilization, notably among people not directly impacted by the movement's grievances but close enough to sympathize. However, our research also ties into the potential drivers of an increasing political polarization in

the United States. If this effect is symmetric across the ideological spectrum, we may expect similar forms of political mobilization in response to other protest triggers, as the attack on the Capitol on January 6, 2021 illustrates.

# References

- Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.
- Abouk, R. and Heydari, B. (2020). The immediate effect of covid-19 policies on social-distancing behavior in the united states. *Public Health Reports*, page 0033354920976575.
- Acemoglu, D., Hassan, T. A., and Tahoun, A. (2018). The power of the street: Evidence from egypt's arab spring. *The Review of Financial Studies*, 31(1):1–42.
- Adam, D., Wu, P., Wong, J., Lau, E., Tsang, T., Cauchemez, S., Leung, G., and Cowling, B. (2020). Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (sars-cov-2) infections in hong kong.
- Anderson, M. and Hitlin, P. (2016). The hashtag# blacklivesmatter emerges: Social activism on twitter. *Pew Research Center*, 15.
- Arrow, K. J. (2000). Increasing returns: historiographic issues and path dependence. *The European Journal of the History of Economic Thought*, 7(2):171–180.
- Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, 99(394):116–131.
- Askitas, N., Tatsiramos, K., and Verheyden, B. (2020). Lockdown strategies, mobility patterns and covid-19. *arXiv preprint arXiv:2006.00531*.
- Bates, R., Greif, A., and Singh, S. (2002). Organizing violence. *Journal of Conflict Resolution*, 46(5):599–628.
- Bazzi, S. and Blattman, C. (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4):1–38.
- Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This Mine Is Mine! How Minerals Fuel Conflicts in Africa. *American Economic Review*, 107(6):1564–1610.

- Besley, T. and Persson, T. (2011). The logic of political violence. *The quarterly journal of economics*, 126(3):1411–1445.
- Black Lives Matter (2020). Home - black lives matter. <https://blacklivesmatter.com/>. Accessed: 2020-10-31.
- Boxell, L., Gentzkow, M., and Shapiro, J. M. (2017). Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617.
- Bureau of Labor Statistics (2020). Supplemental data measuring the effects of the coronavirus (covid-19) pandemic on the labor market. <https://www.bls.gov/cps/effects-of-the-coronavirus-covid-19-pandemic.htm>. Accessed 2022-04-19.
- Bursztyn, L., Cantoni, D., Yang, D. Y., Yuchtman, N., and Zhang, Y. J. (2021). Persistent political engagement: Social interactions and the dynamics of protest movements. *American Economic Review: Insights*, 3(2):233–50.
- Bursztyn, L., Egorov, G., Enikolopov, R., and Petrova, M. (2019). Social media and xenophobia: evidence from russia. Technical report, National Bureau of Economic Research.
- Campante, F., Durante, R., and Sobbrio, F. (2018). Politics 2.0: The multifaceted effect of broadband internet on political participation. *Journal of the European Economic Association*, 16(4):1094–1136.
- Campante, F. R., Durante, R., and Tesei, A. (2021). Media and social capital. Working Paper 29230, National Bureau of Economic Research.
- Cantoni, D., Yang, D. Y., Yuchtman, N., and Zhang, Y. J. (2019). Protests as Strategic Games: Experimental Evidence from Hong Kong’s Anti-Authoritarian Movement. *Quarterly Journal of Economics*, 134(2).
- Cappelli, M. L. et al. (2020). Black lives matter: The emotional and racial dynamics of the george floyd protest graffiti. *Advances in Applied Sociology*, 9(10):323.
- CDC (2020). Risk Assessment and Management. Technical report, Centers for Disease Control and Prevention.
- Colella, F., Lalivé, R., Sakalli, S. O., and Thoenig, M. (2019). Inference with Arbitrary Clustering. IZA Discussion Papers 12584, Institute of Labor Economics (IZA).

- Dave, D. M., Friedson, A. I., Matsuzawa, K., Sabia, J. J., and Safford, S. (2020). Black Lives Matter Protests, Social Distancing, and COVID-19. Technical report, National Bureau of Economic Research.
- Dave, D. M., McNichols, D., and Sabia, J. J. (2021). Political violence, risk aversion, and non-localized disease spread: Evidence from the us capitol riot. Technical report, National Bureau of Economic Research.
- De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.
- Della Porta, D. and Diani, M. (2015). *The Oxford handbook of social movements*. Oxford University Press.
- Della Porta, D. and Diani, M. (2020). *Social movements: An introduction*. John Wiley & Sons.
- Dube, O. and Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from colombia. *The review of economic studies*, 80(4):1384–1421.
- Eisinger, P. K. (1973). The conditions of protest behavior in american cities. *The American Political Science Review*, 67(1):11–28.
- Ellis, C. and Fender, J. (2011). Information cascades and revolutionary regime transitions. *Economic Journal*, 121(553):763–792.
- Endo, A., Abbott, S., Kucharski, A. J., Funk, S., et al. (2020). Estimating the overdispersion in covid-19 transmission using outbreak sizes outside china. *Wellcome Open Research*, 5(67):67.
- Enikolopov, R., Makarin, A., and Petrova, M. (2020). Social media and protest participation: Evidence from russia. *Econometrica*, 88(4):1479–1514.
- Enikolopov, R., Petrova, M., and Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics*, 10(1):150–74.
- Falck, O., Gold, R., and Hebligh, S. (2014). E-lections: Voting behavior and the internet. *American Economic Review*, 104(7):2238–65.
- Freelon, D., McIlwain, C., and Clark, M. (2016). Beyond the hashtags: #Ferguson, #Blacklivesmatter, and the online struggle for offline justice. Technical report, Center for Media & Social Impact.

- Friedson, A. I., McNichols, D., Sabia, J. J., and Dave, D. (2020). Did California's shelter-in-place order work? early coronavirus-related public health effects. Technical report, National Bureau of Economic Research.
- Fujiwara, T., Müller, K., and Schwarz, C. (2021). The effect of social media on elections: Evidence from the united states. Technical report, National Bureau of Economic Research.
- Galvani, A. P. and May, R. M. (2005). Dimensions of superspreading. *Nature*, 438(7066):293–295.
- González, F. and Prem, M. (2020). Police repression and protest behavior: Evidence from student protests in chile. *Available at SSRN 3705486*.
- Guriev, S., Melnikov, N., and Zhuravskaya, E. (2019). 3g internet and confidence in government. *Available at SSRN 3456747*.
- Guriev, S. and Treisman, D. (2015). How modern dictators survive: an informational theory of the new authoritarianism. Technical report, National Bureau of Economic Research.
- Hager, A., Hensel, L., Hermle, J., and Roth, C. (2020). Political activists as free-riders: Evidence from a natural field experiment.
- Hale, T., Webster, S., Petherick, A., Phillips, T., and Kira, B. (2020). Oxford covid-19 government response tracker. *Blavatnik School of Government*, 25.
- Heckman, J. J. (1987). *The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence*. University of Chicago Center for Mathematical studies in Business and Economics.
- Ince, J., Rojas, F., and Davis, C. A. (2017). The social media response to black lives matter: how twitter users interact with black lives matter through hashtag use. *Ethnic and racial studies*, 40(11):1814–1830.
- Jinjarak, Y., Ahmed, R., Nair-Desai, S., Xin, W., and Aizenman, J. (2020). Accounting for global covid-19 diffusion patterns, january-april 2020. Technical report, National Bureau of Economic Research.
- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of econometrics*, 95(2):391–413.

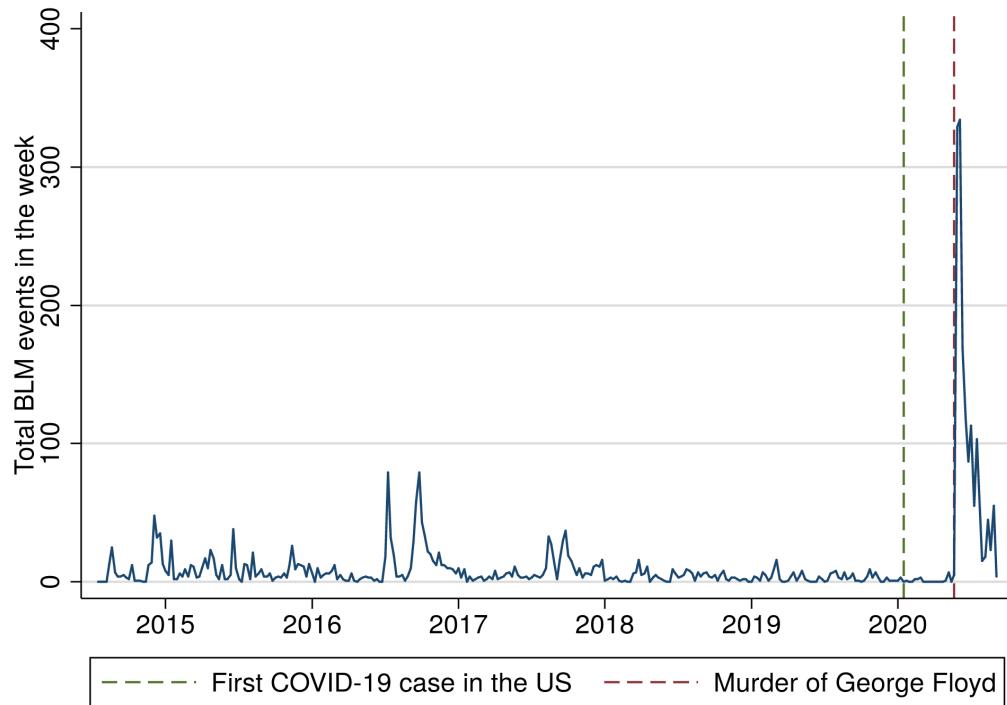
- Lasry, A., Kidder, D., Hast, M., Poovey, J., Sunshine, G., Winglee, K., Zviedrite, N., Ahmed, F., Ethier, K. A., Program, C. P. H. L., et al. (2020). Timing of community mitigation and changes in reported covid-19 and community mobility—four us metropolitan areas, february 26–april 1, 2020. *Morbidity and Mortality Weekly Report*, 69(15):451–457.
- Leclerc, Q. J., Fuller, N. M., Knight, L. E., Funk, S., Knight, G. M., cmmid covid 19 Working Group, et al. (2020). What settings have been linked to sars-cov-2 transmission clusters? *Wellcome open research*, 5.
- Lelkes, Y., Sood, G., and Iyengar, S. (2017). The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science*, 61(1):5–20.
- Liebowitz, S. J. and Margolis, S. E. (1999). Path dependence. *Encyclopedia of law and economics*.
- Lipsky, M. (1968). Protest as a political resource. *The American Political Science Review*, 62(4):1144–1158.
- Lohmann, S. (1994). Information aggregation through costly political action. *American Economic Review*, 84(3):518–530.
- Manacorda, M. and Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567.
- Mangrum, D. and Niekamp, P. (2020). Jue insight: College student travel contributed to local covid-19 spread. *Journal of Urban Economics*, page 103311.
- Marcus, M. and Sant'Anna, P. H. (2021). The role of parallel trends in event study settings: An application to environmental economics. *Journal of the Association of Environmental and Resource Economists*, 8(2):235–275.
- Mathieu, L. (2018). Art and social movements. *The Wiley Blackwell Companion to Social Movements*, pages 354–368.
- Mazumder, S. (2019). Black Lives Matter for Whites' Racial Prejudice:Assessing the Role of Social Movements in Shaping Racial Attitudes in the United States. Technical report, Harvard University.
- McCarthy, J. D. and Zald, M. N. (1977). Resource mobilization and social movements: A partial theory. *American journal of sociology*, 82(6):1212–1241.

- McKersie, R. B. (2021). The 1960s civil rights movement and black lives matter: Social protest from a negotiation perspective.
- Miller, D., Martin, M. A., Harel, N., Kustin, T., Tirosh, O., Meir, M., Sorek, N., Gefen-Halevi, S., Amit, S., Vorontsov, O., et al. (2020). Full genome viral sequences inform patterns of sars-cov-2 spread into and within israel. *medRxiv*.
- Morris, A. D. (1986). *The origins of the civil rights movement*. Simon and Schuster.
- Müller, K. and Schwarz, C. (2019). From hashtag to hate crime: Twitter and anti-minority sentiment. *Available at SSRN 3149103*.
- Müller, K. and Schwarz, C. (2020). From hashtag to hate crime: Twitter and anti-minority sentiment. *Available at SSRN 3149103*.
- Müller, K. and Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4):2131–2167.
- Olson, M. (1989). Collective action. In *The invisible hand*, pages 61–69. Springer.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge university press.
- Passarelli, F. and Tabellini, G. (2017). Emotions and political unrest. *Journal of Political Economy*, 125(3):903–946.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.
- Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *arXiv preprint arXiv:2201.01194*.
- Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221. Endogeneity Problems in Econometrics.
- Sangnier, M. and Zylberberg, Y. (2017). Protests and trust in the state: Evidence from African countries. *Journal of Public Economics*, 152:55–67.
- Takhteyev, Y., Gruzd, A., and Wellman, B. (2012). Geography of twitter networks. *Social networks*, 34(1):73–81.

- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.
- Zhuravskaya, E., Petrova, M., and Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, 12:415–438.

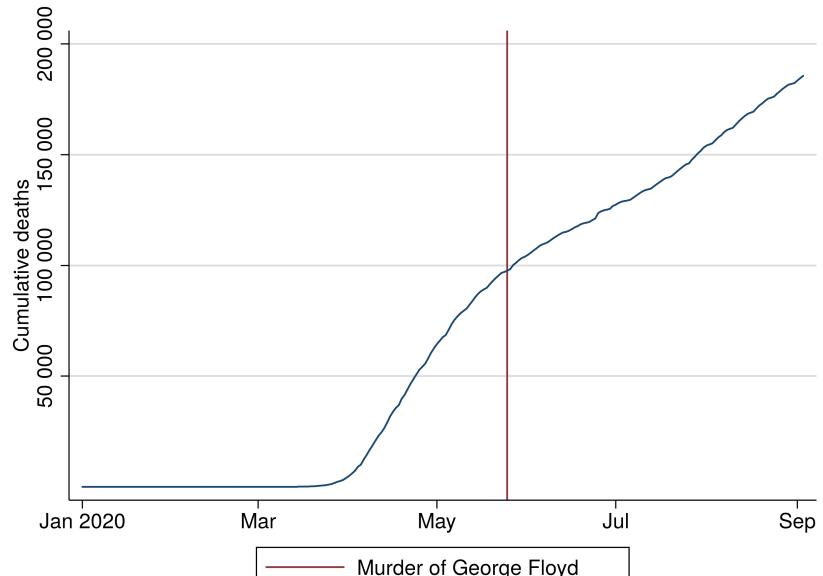
## 9. Figures and Tables

BLM events over time

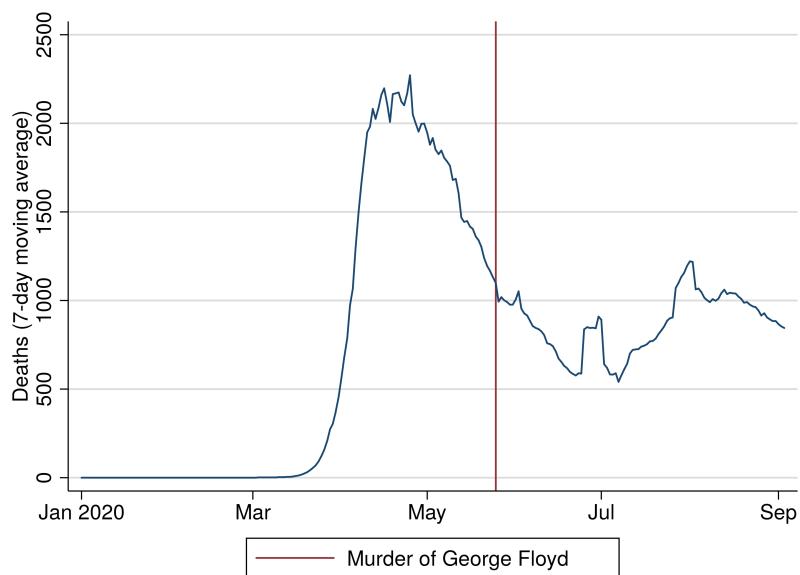


Note: Number of BLM events per week in the US from June 2014 to September 2020. The green vertical line denotes the week of the first confirmed COVID-19 case in the US (January 21, 2020), and the red vertical line denotes the week of the murder of George Floyd (May 25, 2020).

## COVID-19 deaths and timing of GF's murder



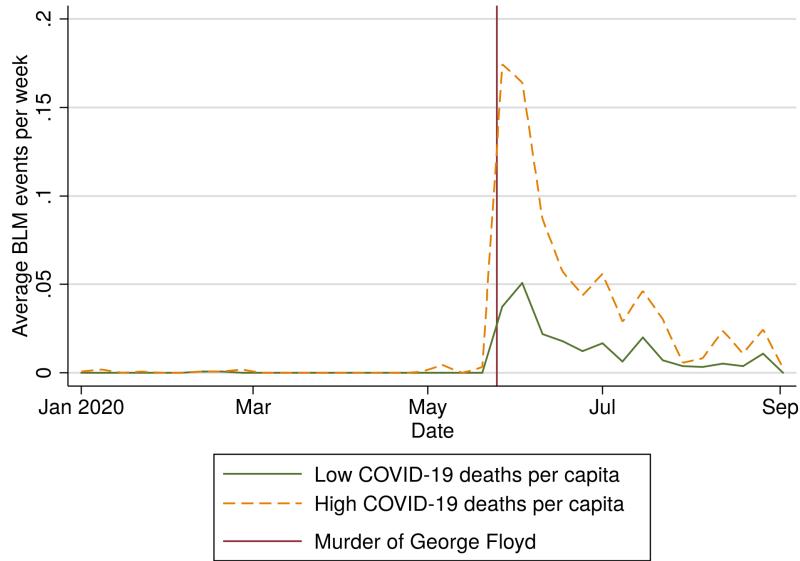
(a) Cumulative deaths



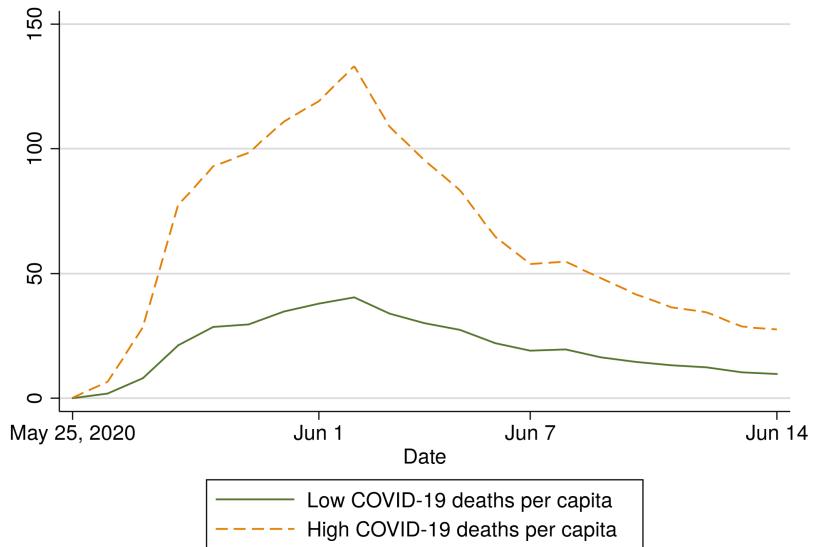
(b) New deaths

Note: Number of cumulative COVID-19 deaths and daily new COVID-19 deaths in the US between January and September 2020. New COVID-19 deaths are presented as a 7-day moving average. The red vertical line denotes the day of the murder of George Floyd (May 25, 2020).

BLM events and tweets in counties with above and below median COVID-19 deaths per capita



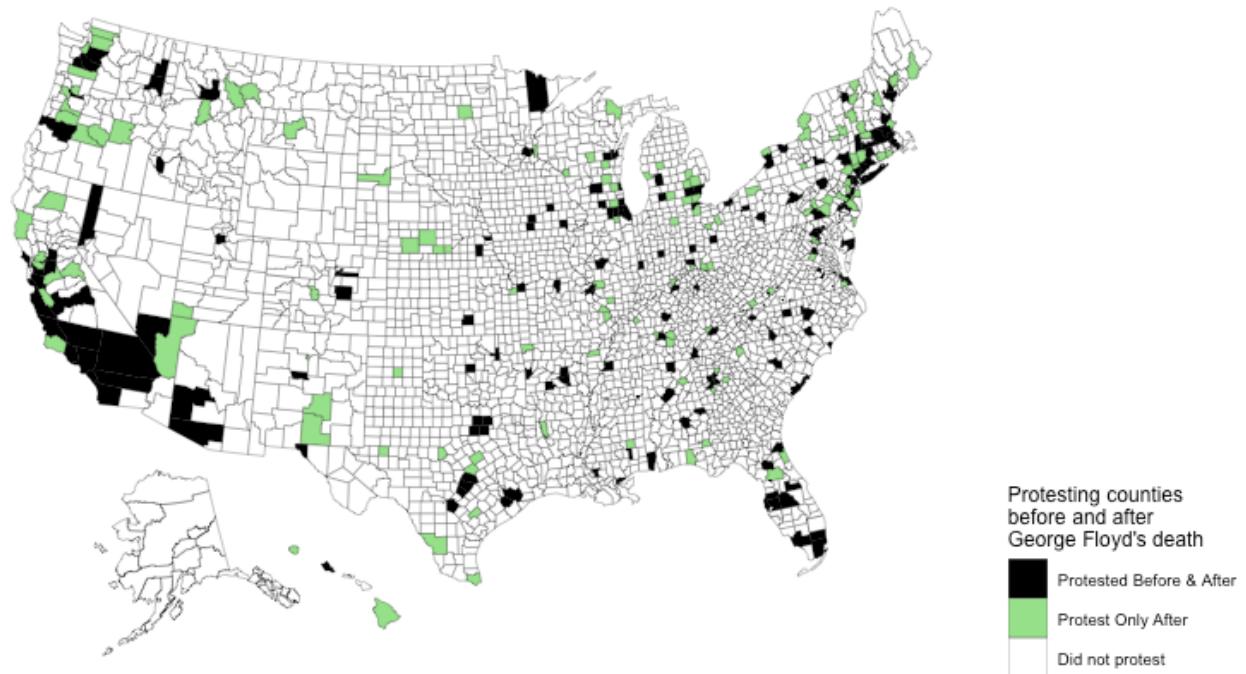
(a) Average BLM protests per week



(b) Average tweets mentioning BLM per day

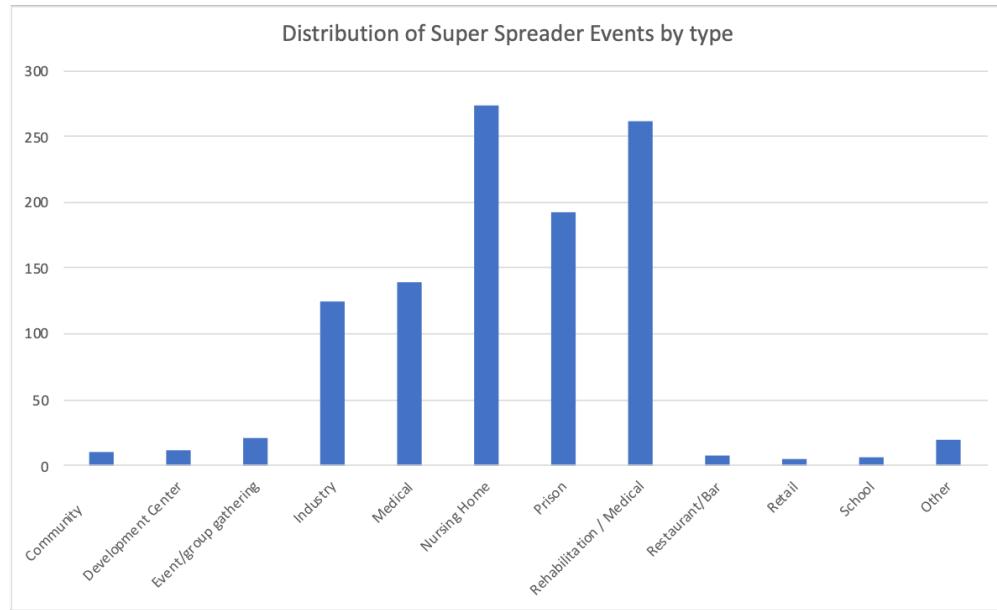
Note: Evolution of two variables over time in counties with below and above median COVID-19 deaths per capita. Subgraph (a) presents the average number of BLM protests per week between January and September 2020. The red vertical line represents the day of the murder of George Floyd (May 25, 2020). Subgraph (b) presents the average number of daily tweets mentioning “BLM” or “Black Lives Matter” from May 25 to June 14, 2020.

Spatial distribution of US counties based on their BLM protest activities before and after George Floyd's murder

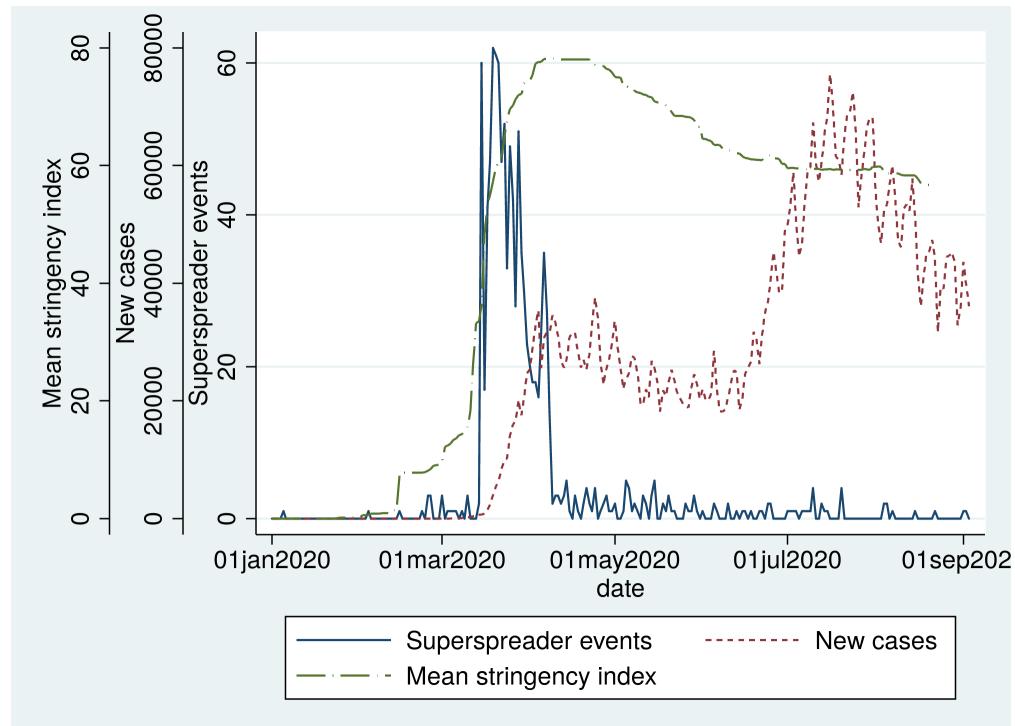


Note: Own visualization based on data from *Elephrame*. This map represents whether US counties that protested in the three weeks following the murder of George Floyd (May 25 to June 14, 2020) already held a BLM protest before the murder of George Floyd. Counties in black protested both before and after the murder of George Floyd. Counties in green are counties whose first BLM protest was after George Floyd's murder. Counties in white did not protest after the murder.

## Distribution of super-spreader events in the US by their type

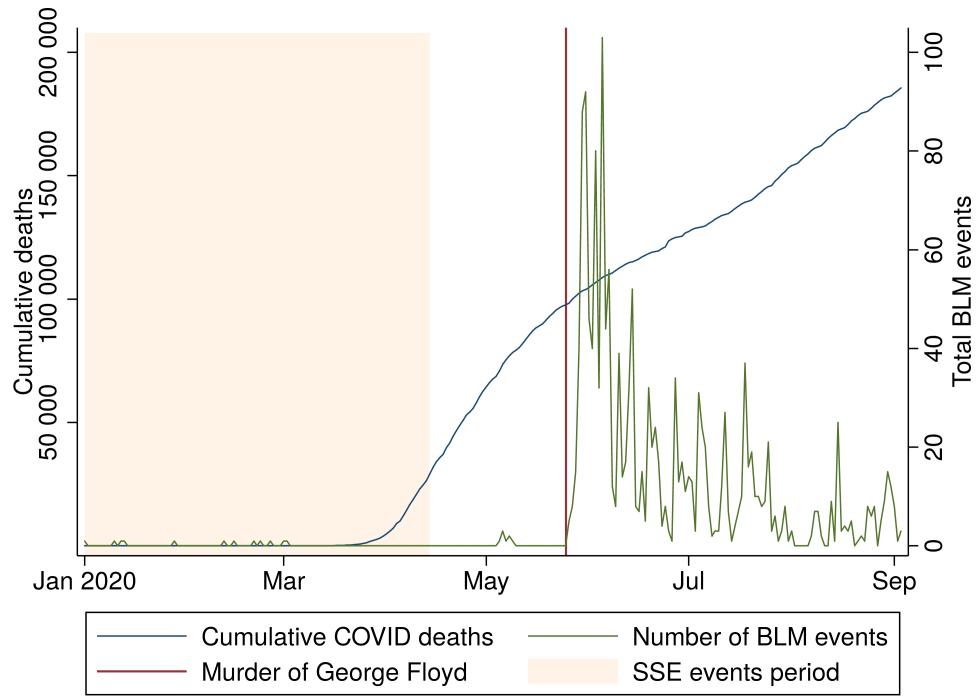


## Window of opportunity for SSEs



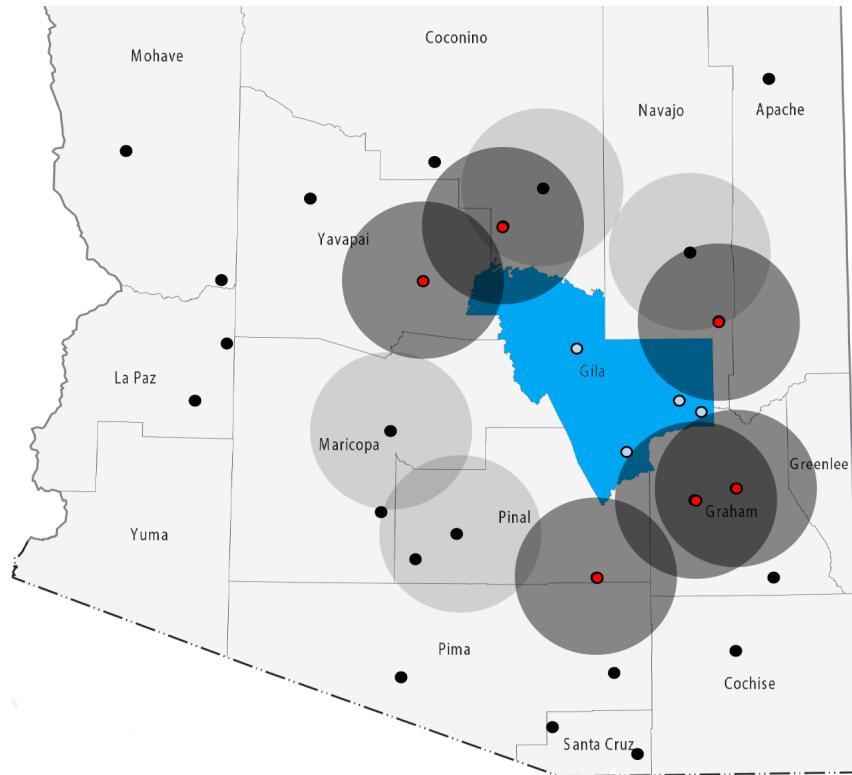
Note: Solid (blue) line represents the number of daily total SSEs over time (January 2020 to September 2020). Dashed (green) line shows the daily average stringency index across all US states, as measured by the Oxford COVID response tracker. Dotted (red) line shows the number of daily new COVID-19 cases as recorded by the New York Times.

## Timing of SSEs relative to Floyd's murder, protest and COVID-19 deaths



Note: Cumulative COVID-19 deaths and BLM events per day from January to September 2020. The red vertical line denotes the week of the murder of George Floyd (May 25, 2020), and the orange shaded area is the period we consider for super-spreader events.

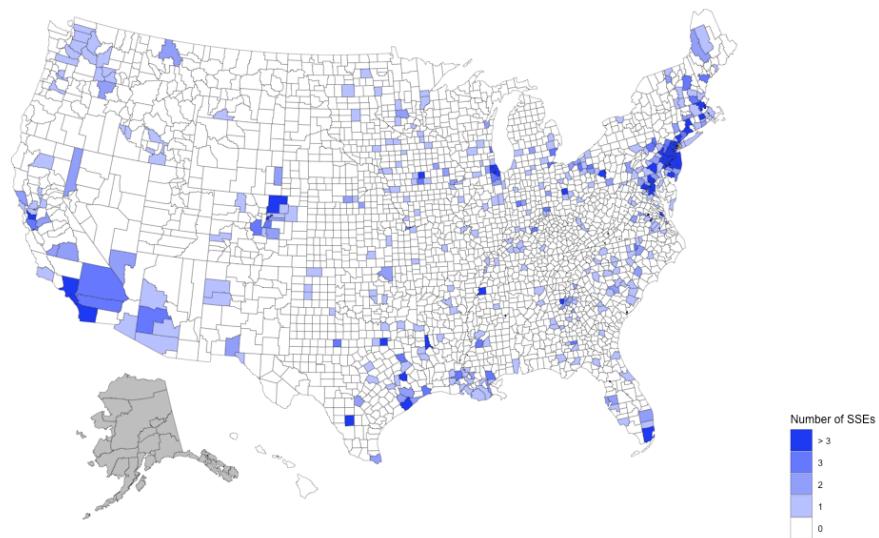
## Construction of the super-spreading events instrument (example)



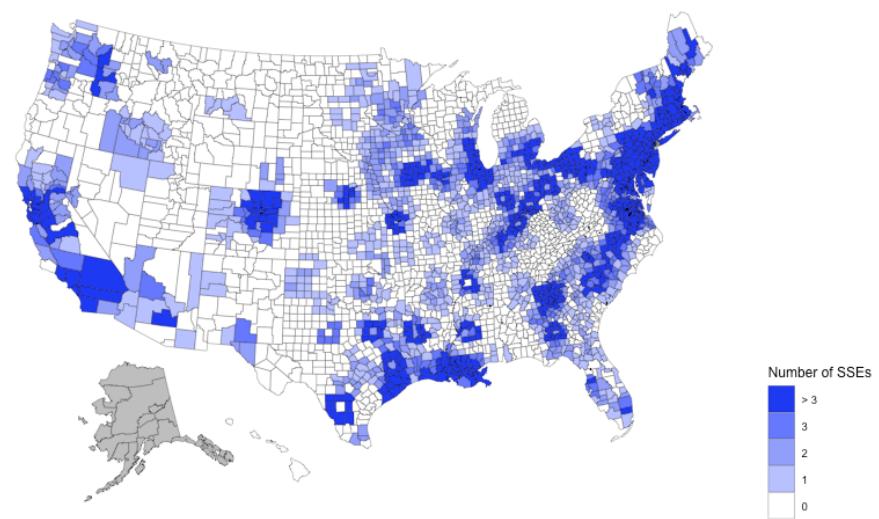
Note: Example of the

construction of the instrument. Red point are the super-spreader events assigned to the blue county. Gray shaded area represents the 50km radius around each super-spreader event. Black points represent super-spreader event that are not assigned to the blue county because are too far away from the border. White points represents super-spreader events that are inside the county and therefore not assigned to the county (to increase exogeneity).

## Geographic distribution of super-spreader events (SSEs)



(a) SSEs at county level 6 weeks prior to Floyd's murder



(b) Variation of SSE instrument  
*Number of SSEs in 50 km radius outside of own county*

Table 3.1: Summary statistics

	N	Mean	SD	Min	Max
<b>From 25th of May to 14th of June 2020:</b>					
Presence of BLM events	3106	0.099	0.298	0.000	1.000
Number of BLM events	3106	0.250	1.348	0.000	36.000
Participants in BLM events	3106	270.759	5968.521	0.000	323687.500
Participants per event	307	539.141	878.429	0.000	8991.319
Tweets mentioning BLM	3106	819.502	7187.496	0.000	243596.000
New users tweeting about BLM	3106	4.586	53.812	0.000	2442.000
Followers of @BlkLivesMatter created during the pandemic	3106	1.540	11.207	0.000	453.000
Tweets mentioning #AllLivesMatter	3106	134.741	833.066	0.000	28943.000
Tweets mentioning #BlueLivesMatter	3106	17.753	113.478	0.000	4117.000
Neighbor protested first	3106	0.348	0.477	0.000	1.000
Other Protests	3108	0.081	0.386	0.000	7.000
COVID-19 Protests	3108	0.030	0.204	0.000	4.000
<b>On the 25th of May 2020:</b>					
COVID deaths (total)	3106	24.461	141.132	0.000	3304.000
COVID cases (total)	3106	459.678	2438.202	0.000	72010.000
COVID deaths (per 1000)	3106	0.113	0.248	0.000	2.935
COVID cases (per 1000)	3106	2.791	5.664	0.000	145.513
Super-spreader events, 6+ weeks ago, neighboring	3106	3.070	9.790	0.000	143.000
Black death burden	3106	1.346	0.963	0.000	4.104
Lockdown stringency index	3106	68.445	8.508	47.220	89.810
<b>Before the 25th of May 2020:</b>					
Google searches for Twitter	3056	61.265	11.222	17.000	100.000
Residential stay	1348	10.633	3.387	3.600	26.286
<b>Later outcomes:</b>					
Followers of @BlkLivesMatter	3106	63.198	495.174	0.000	20058.000
Street art count	3106	0.703	26.735	0.000	1467.000
<b>County characteristics:</b>					
Black police-related deaths (2014-2019)	3106	0.677	3.207	0.000	84.000
Black police-related deaths (2020)	3106	0.047	0.301	0.000	6.000
Unemployment rate (year average)	3106	4.691	1.550	0.708	19.650
Black population share	3106	0.100	0.147	0.000	0.875
Non-white population share	3106	0.144	0.162	0.000	0.928
Large cities	3106	0.020	0.140	0.000	1.000
Suburban areas	3106	0.118	0.323	0.000	1.000
Smaller towns	3106	0.234	0.423	0.000	1.000
Rural areas	3106	0.628	0.483	0.000	1.000
BLM events (2014-2019)	3106	0.617	4.183	0.000	117.000
Black poverty rate	3106	0.281	0.225	0.000	1.000
Population share with 3+ risk factors	3106	25.899	5.019	10.685	48.448
Vote share for republicans (2016)	3106	0.633	0.156	0.083	0.960
Vote share for republicans (2012)	3106	0.596	0.148	0.060	0.959
Median household income (2016)	3106	48795.991	13277.575	20170.891	129150.343
Social capital	3106	1.384	0.705	0.000	6.887
Distance to Minneapolis	3106	1216.679	555.825	11.998	6474.706
Notable Deaths	3106	0.010	0.116	0.000	3.000
Log(SXSW followers created before March 2017)	3106	0.114	0.258	0.000	1.474
Log(SXSW followers created during March 2017)	3106	0.193	0.350	0.000	1.658

Note: Summary of main variables used in our analysis. The sample consists of 3,108 US counties. We report the number of observations, the mean, the standard deviation as well as the minimum and maximum value of each of the variables.

Table 3.2: Main Result - COVID exposure and BLM protest

	Presence of BLM events				
	(1)	(2)	(3)	(4)	(5)
<b>Panel A: All counties</b>					
<b>IV:</b> COVID (deaths/1000)	0.647*** (0.0930)	0.730*** (0.187)	0.589*** (0.167)	0.296** (0.117)	0.215* (0.121)
<b>OLS:</b> COVID (deaths/1000)	0.203** (0.0831)	0.158** (0.0638)	0.0758* (0.0435)	0.0382 (0.0289)	0.0323 (0.0264)
Observations	3,108	3,107	3,107	3,106	3,106
F first stage	95.03	31.92	27.44	38.38	36.05
Mean dep. var.	0.0994	0.0991	0.0991	0.0988	0.0988
<b>Panel B: Counties with no BLM protest before</b>					
<b>IV:</b> COVID (deaths/1000)	0.555*** (0.0745)	0.675*** (0.160)	0.790*** (0.177)	0.467*** (0.170)	0.404** (0.187)
<b>OLS:</b> COVID (deaths/1000)	0.0661 (0.0445)	0.0503 (0.0319)	0.0562* (0.0310)	0.0407 (0.0247)	0.0385* (0.0221)
Observations	2,768	2,767	2,767	2,767	2,767
F first stage	115.1	44.53	29.25	27.95	27.04
Mean dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477
<b>Panel C: Counties with BLM protest before</b>					
<b>IV:</b> COVID (deaths/1000)	0.277*** (0.0597)	0.502** (0.229)	0.386* (0.206)	0.116 (0.289)	0.0104 (0.266)
<b>OLS:</b> COVID (deaths/1000)	0.252*** (0.0494)	0.435*** (0.0963)	0.224*** (0.0740)	0.0733 (0.102)	0.0682 (0.102)
Observations	340	334	334	333	333
F first stage	105.3	37.56	32.01	29.27	28.09
Mean dep. var.	0.521	0.515	0.515	0.514	0.514
State fixed effects		Y	Y	Y	Y
Demographic controls			Y	Y	Y
Economic controls				Y	Y
Political controls					Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of at least one Black Lives Matter event during the three weeks following the murder of George Floyd. Panel A presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument and OLS results for all US counties. Panel B presents these results for the sub-sample of counties with no BLM protest before the murder of George Floyd. Panel C presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. Each column include sequentially different sets of additional controls. Demographic controls: share of Black population, urban (category [1-6]). Economic controls: median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience. Political controls: Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations), number of past BLM events between 2014 and 2019, deadly force used by police against Black people. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.3: Alternative outcomes

	Presence of BLM events	Number of BLM events	Total participants	Participants per event	Tweets	Followers	Street art
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Sample: Counties with no BLM protest before</b>							
IV: COVID	0.404** (0.187)	0.621** (0.247)	-81.00 (261.4)	-106.2 (265.9)	2,586* (1,449)	174.8** (79.07)	0.0837 (0.141)
Observations	2,767	2,767	2,767	2,767	2,767	2,767	2,767
F first stage	27.04	27.04	27.04	27.04	27.04	27.04	27.04
Mean dep. var.	0.0477	0.0636	21.03	16.94	322.6	16.18	0.00940
All controls	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths on different outcomes. Columns 1 to 4 use protest information from *Elephrame*. Column 1 presents our main regression. Column 2 uses the number of BLM events in the three weeks following the murder of Floyd. Column 3 reports results for the number of participants and column 4 divides the number of participants by the number of events in the county (imputing zero to counties with no BLM event). Column 5 reports the number of geo-located tweets that use at least one of the following hashtags #BlackLivesMatter #BlackLifeMatters #BLM collected through the twitter API in the three weeks following the murder. Column 6, reports the number of geo-located accounts that follow the official BLM account @BlkLivesMatter. Column 7 uses the geo-location of street art containing references to Black Lives Matter and George Floyd from the *Urban Art Mapping George Floyd and Anti-Racist Street Art database*. We report Kleibergen-Paap rkWald F statistic for COVID and the respective interaction terms. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.4: Heterogeneity by baseline county characteristic

	Presence of BLM events								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Sample: All counties</b>									
COVID (deaths/1000)	0.215*	-0.908*	0.276	-0.176	0.256	-0.314	-0.0301	0.295**	0.243**
(0.121)	(0.487)	(0.204)	(0.301)	(0.175)	(0.188)	(0.147)	(0.112)	(0.120)	
... × Non-Black population share	1.301**								
(0.548)									
... × Non-white population share		-0.163							
(0.411)		0.004117*							
(0.00243)									
... × Median household income			-0.102						
(0.393)									
... × Vote Republican 2016				0.608***					
(0.155)									
... × Not large cities					0.321***				
(0.112)									
... × Suburban areas						0.0391			
(0.137)									
... × Smaller towns							-0.155		
(0.159)									
... × Rural areas									
Interacting variable	-0.191	-0.111**	0.00224**	-1.056***	-0.566***	-0.0572**	0.0703***	-0.0652***	
(0.185)	(0.0537)	(0.000926)	(0.191)	(0.0865)	(0.0251)	(0.0220)	(0.0223)		
Observations	3,106	3,106	3,106	3,106	3,106	3,106	3,106	3,106	
F COVID	36.05	20.93	22.58	22.62	24.95	42.35	37.12	20.25	17.36
F interaction	17.19	46.17	18.79	13.81	110.8	48.84	42.69	9.229	
All controls	Y	Y	Y	Y	Y	Y	Y	Y	
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of at least one Black Lives Matter event during the three weeks following the murder of George Floyd. Table presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument for the full sample (all counties with and without prior BLM protest). All controls and state fixed effects are included. Demographic controls: share of Black population, urban (category [1-6]). Economic controls: median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience. Political controls: Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations), number of past BLM events between 2014 and 2019, deadly force used by police. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.5: COVID-19 exposure and social media use

	PC 1	New Twitter accounts	(log) New Twitter accounts	Google searches for Twitter	Residential stay	New followers @BlkLivesMatter	(log) New followers @BlkLivesMatter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Panel A: All counties</b>							
<b>IV: COVID (deaths/1000)</b>	0.911* (0.523)	-0.709 (20.17)	0.690* (0.376)	12.94* (6.453)	3.155*** (0.592)	4.466* (2.273)	0.405 (0.289)
Observations	1,332	3,106	3,106	3,056	1,348	3,106	3,106
F first stage	27.65	36.05	36.05	35.71	27.52	36.05	36.05
Mean dep. var.	-1.27e-09	4.586	0.586	61.27	10.63	1.540	0.342
<b>Panel B: Counties with no BLM protest before</b>							
<b>IV: COVID (deaths/1000)</b>	1.395** (0.527)	17.88** (7.871)	1.317*** (0.339)	18.28** (8.838)	3.885*** (0.931)	3.816* (2.091)	0.522 (0.373)
Observations	1,014	2,767	2,767	2,730	1,022	2,767	2,767
F first stage	20.21	27.04	27.04	26.03	20.16	27.04	27.04
Mean dep. var.	-0.474	1.808	0.420	60.52	10.01	0.441	0.200
<b>Panel C: Counties with BLM protest before</b>							
<b>IV: COVID (deaths/1000)</b>	0.127 (0.476)	-37.13 (62.07)	-0.374 (0.395)	5.724 (6.164)	2.437*** (0.886)	10.20* (5.871)	-0.386 (0.272)
Observations	312	333	333	320	320	333	333
F first stage	25.27	28.09	28.09	26.65	26.13	28.09	28.09
Mean dep. var.	1.518	27.47	1.931	67.62	12.62	10.56	1.507
All controls	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on use of social media. Column 1 shows the first principal component of the four outcomes of interest: new Twitter accounts tweeting about BLM (and its log), Google searches for Twitter, residential stay and new Twitter accounts following the official BLM Twitter account (and its log). Table 3.C.7 details the construction of the principal component. Column 2 (resp column 3) shows estimates for new Twitter accounts (log of new accounts) created after the beginning of the pandemic but before Floyd's murder that tweet about BLM in the three weeks following Floyd's death. Column 4 (resp column 5) shows results for Google searches for Twitter (residential stay) between April 13 to May 24. Column 6 (resp column 7) shows estimates for new Twitter accounts (log of new accounts) created after the beginning of the pandemic but before Floyd's murder that followed the official BLM Twitter account. Panel A presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument and OLS results for all US counties. Panel B presents these results for the sub-sample of counties with no BLM protest before the murder of George Floyd. Panel C presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.6: Effect of Twitter on BLM protest

Sample: <i>Counties with no BLM protest before</i>	<i>Uninstrumented users</i>		<i>Instrumented users</i>
	(1)	(2)	(3)
COVID (deaths/1000)	-0.599 (0.409)	-0.0444 (0.277)	-0.578 (0.568)
× Log(Preexisting users)	0.245*** (0.0880)		0.232* (0.118)
× Log(New users)		0.205** (0.0834)	
Log(Preexisting users)	0.0128 (0.00854)		0.0406 (0.0453)
Log(New users)		0.0193* (0.0102)	
Mean of dep. var	0.0477	0.0477	0.0477
F COVID	11.35	15.28	8.530
F users			19.31
F interaction	47.35	60.91	18.87
Observations	2,767	2,767	2,767
Instruments	SSE	SSE	SSE & SXSW
All controls	Y	Y	Y
Pre-SXSW users			Y
State fixed effects	Y	Y	Y

Note: Column 1 and 2 show the effect of uninstrumented pre-existing or new users interacted with COVID deaths (instrumented by SSE) on the presence of BLM events in a county. Column 3 shows an IV estimate of the model of column 1, with pre-existing users instrumented by SXSW users. The first stage regression is reported on Table 3.C.6. We present results for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and all standard controls. First stage F statistic for weak identification per second-stage coefficient (F COVID, F users, F interaction) following Sanderson and Windmeijer (2016). Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.7: Effect of Twitter on BLM protest

Sample: <i>Counties with BLM protests before</i>	<i>Uninstrumented users</i>		<i>Instrumented users</i>
	(1)	(2)	(3)
COVID (deaths/1000)	-0.338 (0.649)	0.368 (0.341)	-0.000256 (5.753)
× Log(Preexisting users)	0.0738 (0.107)		-0.0351 (1.090)
× Log(New users)		-0.133 (0.100)	
Log(Preexisting users)	0.158*** (0.0404)		-0.462 (1.373)
Log(New users)		0.0744** (0.0306)	
Mean of dep. var	0.514	0.514	0.514
F COVID	22.99	43.74	0.859
F users			0.309
F interaction	30.31	56.90	0.750
Observations	333	333	333
Instruments	SSE	SSE	SSE & SXSW
All controls	Y	Y	Y
Pre-SXSW users			Y
State fixed effects	Y	Y	Y

Note: Column 1 and 2 show the effect of uninstrumented pre-existing or new users interacted with COVID deaths (instrumented by SSE) on the presence of BLM events in a county. Column 3 shows an IV estimate of the model of column 1, with pre-existing users instrumented by SXSW users. The first stage regression is reported on Table 3.C.6. We present results for the sub-sample of counties with BLM protests before the murder of George Floyd. All specifications include state fixed effects and all standard controls. First stage F statistic for weak identification per second-stage coefficient (F COVID, F users, F interaction) following Sanderson and Windmeijer (2016). Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.8: Survey data: COVID-19, news consumption and attitudes towards BLM, Blacks and COVID-19

	News consumption			Attitudes towards Blacks, BLM & COVID-19			Placebo
Follow news about GF	about GF on social media	Ratio overall GF news	COVID hospitaliz. not their fault	Higher Black structural racism	Protest because criminal behaviour	Protest because racism	Illegal immigration (7)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
COVID-19 deaths per capita (category)	0.0480*** (0.00964)	0.0343** (0.0152)	0.0225* (0.0134)	0.0115* (0.00645)	0.0259*** (0.00907)	-0.0254** (0.0109)	-0.00641 (0.00540)
Black	Y	Y	Y	Y	Y	Y	Y
Metropolitan area	Y	Y	Y	Y	Y	Y	Y
Female	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y	Y
Education	Y	Y	Y	Y	Y	Y	Y
Income	Y	Y	Y	Y	Y	Y	Y
Democrat	Y	Y	Y	Y	Y	Y	Y
Observations	9,201	9,121	9,111	9,212	9,190	9,212	9,183

Note: Relation between living in a county with different levels of COVID-19 deaths per capita on different outcomes related to news consumption and attitudes towards Blacks, BLM and COVID-19. Columns 1 to 3 present the estimates for outcomes related to news consumption. In particular, column 1, 2 and 3 show respectively: the interest in George Floyd related news, the amount of GF related news received through social media and the ratio of the variable of column 2 over the variable of column 1. Columns 4 to 6 show the results for the outcomes related to attitudes towards BLM and racism awareness. Column 4 corresponds to the likelihood of answering that the higher COVID-19 mortality rate faced by Blacks is due to their disadvantaged circumstances instead of to their personal life style choices. Columns 5 and 6 correspond to the likelihood of answering that the protest following George Floyd's death is related with structural racism or to criminal behaviour respectively. Finally, column 7 shows a placebo result. The exact framing of the questions is as follows: column 1: "*How closely have you been following news about the demonstrations around the country to protest the death of George Floyd, a black man who died while in police custody?*"; column 2: *How much, if any, news and information about the demonstrations to protest the death of George Floyd have you been getting on social media (such as Facebook, Twitter, or Instagram)?*; column 4: *Do you think the reasons why black people in our country have been hospitalized with COVID-19 at higher rates than other racial or ethnic groups have more to do with... Circumstances beyond people's control*; column 5: *How much, if at all, do you think each of the following has contributed to the demonstrations to protest the death of George Floyd? Longstanding concerns about the treatment of black people in the country*; column 6: *Some people taking advantage of the situation to engage in criminal behavior*; column 7: *Which comes closer to your view about how to handle undocumented immigrants who are now living in the U.S.? There should be a way for them to stay in the country legally, if certain requirements are met*. All columns include controls for various characteristics of the respondent: race, whether or not they live in a metropolitan area, gender, age, education, income and whether or not they lean towards the democratic party. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.9: Competing Mechanisms: Broadening versus Scattering of Protest

	(1)	(2)	(3)	Presence of BLM events		
Sample: Counties without BLM protests before				(4)	(5)	(6)
IV: COVID (deaths/1000)	0.410** (0.189)	0.306 (0.343)	0.412** (0.191)	0.175 (0.257)	-0.153 (0.806)	3.307 (2.464)
× Neighbor protested historically		0.116 (0.345)				
× Neighbor protested currently			0.236 (0.240)			
× Distance to Minneapolis				0.000371 (0.000532)	-0.00770 (0.00473)	
× Distance to Minneapolis (squared)					3.73e-06* (2.11e-06)	
Neighbor protested historically	-0.0114 (0.0120)	-0.0223 (0.0317)				
Neighbor protested currently			-0.0103 (0.0143)	-0.0289 (0.0253)		
Distance to Minneapolis					0.0000161 (0.0000479)	0.000247 (0.000250)
Distance to Minneapolis (squared)						-0.000000138 (0.000000114)
Observations	2,767	2,767	2,767	2,767	2,767	2,767
F first stage	27.08	13.75	26.60	13.47	29.44	16.61
F interaction		16.57		32.07	21.22	11.51
F interaction sq					7.753	
Mean of dependent variable	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477
All controls from preferred specification	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of BLM protests. Column 1 (column 3) shows estimates for instrumented COVID deaths controlling for a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 2 and 4 present heterogeneous effects for the presence of a neighbouring county that protested before. Columns 2 (column 4) shows the interaction term with a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 5 and 6 presents results for interaction with distance to Minneapolis and distance to Minneapolis squared. All results are shown for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.10: Competeting Mechanisms: Salience, Opportunity Cost, and Agitation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Sample: Counties with no BLM protest before</b>								
COVID (deaths/1000)	0.444** (0.1994)	0.585 (0.342)	0.507 (0.536)	0.0423 (0.980)	0.279 (0.224)	0.087 (0.104)	528.1 (348.2)	62.57 (39.19)
... × Black death burden	1.391 (1.476)							
... × Google BLM search		-0.004 (0.022)						
... × Unemployment			-0.003 (0.054)					
... × Stringency				0.006 (0.264)				
Interacting variable	-0.257 (0.176)	0.0006 (0.001)		0.0047 (0.008)				
Observations	2,767	2,647	2,767	2,768	2,767	2,767	2,767	2,767
F stat COVID	31.95	19.09	24.93	85.33	41.12	41.12	27.04	27.04
F stat Interaction	3.89	25.33	13.61	94.27				
Mean of dependent variable	0.0477	0.0477	0.0477	0.0477	0.0321	0.010	47.45	6.123
All controls	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on presence of several types of protest. Column 1 shows estimates for instrumented COVID deaths. Columns 2 to 4 show heterogeneous effects for Black death burden weeks prior to GF's murder, Google searched for BLM 3 weeks prior to GF's murder, unemployment and stringency 3 weeks after GF's murder. The coefficient for the interacting variable in column 4 is dropped as stringency is measured at the state level and state fixed effects are included. Column 5 (resp column 6) presents results for all other protests besides BLM (protest related to COVID-19; e.g. anti-mask protest) during the 3 weeks following Floyd's murder. Columns 7 and 8 show as an outcome the number of tweets including the pro-police and anti-BLM hashtags #AllLivesMatter and #BlueLivesMatter. All results are shown for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level.

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

## A. Appendix: Robustness Checks

Our robustness checks focus on three dimensions: *i*) robustness to changes in the definition and construction of our instrumental variable *ii*) robustness of our main results to sample composition, spatial correlation and other confounding factors and *iii*) the possibility that our results are driven by a relocation of protesters across time and space rather than a "broadening" of the BLM coalition. We present our results in Table 3.A.2 to Table 3.A.5.

### A.1. Instrument Robustness

We present results on the robustness of the instrument in Table 3.A.2 and Table 3.A.3, showing the IV result and first stage coefficient for both the full sample (Panel A) and the sub-sample of counties without prior BLM events (Panel B). Our baseline results are always reported in column 1 for reference.

**Changing the radius around SSEs.** In the baseline specification, we choose the 50km threshold as a distance of the SSE to the county border, as it is approximately two times the average radius of a county in the US.<sup>28</sup> To make sure that this choice is not driving our results, we change the radius of influence to 25 km, 100 km and 200 km (columns 2, 3 and 4 of Table 3.A.2 respectively). For both samples the coefficient remains significant and becomes slightly larger in magnitude.

**Changing the time window of SSEs.** Similarly, in our preferred specification, we take into account the SSEs that occurred in a specific time window that we call "window of opportunity" where there were enough cases to observe SSEs and the social distancing measures were not applied strictly or widely enough. Specifically, we count the number of SSEs between the beginning of the COVID-19 outbreak until April 13th 2020 (e.g., six weeks before Floyd's murder). In columns 6 to 8 of Table 3.A.2 we expand and narrow this window to make sure our results are not driven by the specific timing of SSEs. In particular, we count SSEs until April 20th, 5 weeks before the murder of Floyd (column 6), on April 6th, 7 weeks before (column 7) and on March 30th, 8 weeks before (column 8). Results are robust to change in the time window.

**Excluding SSEs in prisons.** A non-negligible number of SSEs occurred inside prisons. We exclude SSEs in prisons in a robustness check in column 2 of Table 3.A.3 for two reasons. First, it is likely that by the nature of prisons, the geographical spread of cases stemming

---

<sup>28</sup>For reference, the average radius of a county is 28 km and the average radius of a state is 220 km.

from an SSE in a prison is quite limited and less relevant for the overall population and the protesting population. In this case, we would expect a bigger effect when excluding these SSEs. Second, SSEs in prisons may have an effect on BLM protests other than through overall exposure to COVID, for instance, by raising the salience of the overproportional incarceration of Black people. In this case, we would expect the coefficient to decrease in magnitude when excluding these SSEs. While the salience of racial inequality in prisons may be a possible mechanism, with this exercise we investigate whether our results are indeed solely driven by this subsample of SSEs. We exclude SSEs in prisons in column 2 and find that our results slightly increase in magnitude and precision.

**Controlling for SSEs in the county.** Our first stage compares the effect of having an SSE outside the county within 50 km of the county border and excluding the effect of SSEs that take place within its border. Therefore, in our analysis a county is "not affected" by an SSE if its border is either further than 50 km from the SSE, or the SSE happened within its boundaries. We expect the effect of SSEs to be different between these groups: presumably, counties far away will have no COVID-19 cases from this SSE, while the county where the SSE took place will have a lot of cases and deaths caused by the event. To assuage the concern that correlation of SSEs across counties is driving the variation in SSE exposure, we add as a control the number of SSEs that occurred within the county itself. Estimates are presented in column 3 of Table 3.A.3 and show that the results of the baseline specification are robust to the addition of this control for the counties with no BLM before, and become imprecisely estimated for the sample of all counties (with a p-value of 0.122).

**Weighting SSEs by distance.** In our baseline specification, we count any SSE that occurred in a 50 km radius outside the border of a county as an additional SSE affecting the county. However, an SSE 1 km away from the border is likely to have a different level of influence from a SSE 49 km away. To ensure that this simplification is not driving the results, we refine the level of influence in three different ways. First we weight the SSEs by a linear function decreasing with distance (column 4 of Table 3.A.3), giving less weight to events that are more distant. Second, we repeat the analysis but with a quadratic function (column 5 of Table 3.A.3), weighting distant events less and increasingly so. The results are robust to these distance weighting procedures.

**Weighting SSEs by the inverse probability of occurrence.** The probability of being near a county that has an SSE is not constant over all counties. For instance, counties neighboring cities have likely a higher probability of being treated by our instrument as

their neighbors may be more likely to experience an SSE. This could be a violation of the exclusion restriction because the probability of being treated by our instrument at a certain level is not uniform, and this heterogeneity could be related to certain county characteristics that could in turn be related to the probability of protesting. To address this concern, we weight each observation by the inverse probability of being treated. Using LASSO (a regularized regression procedure that performs variable selection and avoids overfitting, Tibshirani (1996)), we select relevant variables predicting (by a logit model) the probability of having a neighbor with an SSE among a set of county characteristics, including a large set of socio-demographic and economic characteristics extracted from the American Community Survey (such as population, population density, race distribution, age groups, poverty rates, among others), indicators for different levels of urbanization, geographical indications (latitude, longitude, and state dummies), as well as the minimum and maximum of these variables for neighboring counties. We use the LASSO selected model to predict the probability of a county having a neighbor with an SSE, then weight the observations by the inverse of this probability. This means that counties with a higher probability of having a neighbor with an SSE that actually had a neighbor with an SSE are weighted less than counties with a lower probability of being treated that are actually treated. Estimates are presented in column 7 of Table 3.A.2 and show that our results are robust to this weighting procedure.

**Plausibility of exclusion restriction.** If our instrument were to pick up any underlying factors correlated with the overall likelihood of protesting for a BLM-related cause, then this would challenge a causal interpretation of our estimates. To probe the plausibility of the exclusion restriction, we estimate the effect of instrumented COVID-19 on the likelihood of observing past BLM protests. If our instrument were correlated with the county unobservables that also predict the likelihood of observing BLM protests, then we would expect to see a statistically significant relationship between our instrumented COVID-19 and likelihood of observing a BLM protest in the past. In column 2 of Table 3.A.4, we show that exposure to COVID-19 does not predict the presence of BLM events between 2014 and 2019. We take this as additional evidence for the plausibility of our identifying assumption.

## A.2. *Robustness of Main Results*

In this section, we focus on our main results and run robustness checks including changing definitions in treatment and outcome, estimation method, spatial correlation and concerns about the overall propensity to protest. We present these checks in Table 3.A.4 and Table 3.A.5.

**Excluding coastal counties and states.** Coastal states and counties might behave differently, either with regard to our instrument or to the process of COVID-19 contagion. Coastal regions are generally denser, which increases the chance of having an SSE (Figure 3.9 shows the density of SSEs). On the other hand, our instrument behaves differently, as half of the area where SSEs could affect the county is actually ocean. Coastal regions are also more internationally connected, and were the first affected by COVID-19 in the US (the first reported case was in the state of Washington, and the first reported death in California). We show that our main result for the counties with no BLM protest before is robust to excluding coastal counties (column 3 of Table 3.A.4), as well as coastal states (column 4). Estimates for panel A remain with similar magnitude but become imprecisely estimated.

**Time window of protests.** In our baseline specification, we choose the three week window following Floyd's murder since it captures the vast majority of BLM-related protests (see Figure 3.3), while being close enough to the exposure to COVID-19 on May 24th, right before the protest trigger. We show that our main results (Panel B) are robust to reducing this time window to 2 weeks and expanding this time window to 6 and 8 weeks (columns 5 to 7 of Table 3.A.4 respectively). The coefficient of interest in both samples is more precisely estimated the further we expand the time window of protest.

**COVID-19 cases.** In our baseline specification we use the number of COVID-19 deaths per thousand in the county as an explanatory variable for protest. It is possible that COVID-19 deaths may have a different or distinct effect on BLM protest. This could be due to - for instance - different threat perceptions or salience of the pandemic. In column 8 of Table 3.A.4, we show that the results hold when using the number of COVID-19 related cases instead of the number of deaths. As expected, the number of COVID-19 related cases exhibits significantly smaller coefficients but continues to significantly and positively affect protest behavior.

**Probit estimation.** In our baseline specification the effect of COVID-19 is additive. It might be the case that the effect would be multiplicative of some characteristics of the counties. Using a Probit model accounts for this possibility. Non-linear models with many covariates (typically when using fixed effects) suffer from the incidental parameter problem resulting in bias of the estimates Heckman (1987); Lancaster (2000); Wooldridge (2015). To reduce the extent of this problem we omit the state fixed effects, which significantly reduces the number of covariates. We use an OLS in the first stage, but estimate the second stage

with a Probit model. Results are presented in column 2 of Table 3.A.5. The Probit model delivers larger and more precisely estimated coefficients for the sub-sample of counties with no prior BLM event and positive (and largely similar in size) but more imprecisely estimated coefficients (with a p value of 0.11) for the full sample.

**Controlling for propensity to protest.** Our main specification already controls for the number of BLM events that took place in the county in previous years. While this gives some indication of the county's propensity to protest, this is essentially an imprecise measure, since counties having a non-zero probability of protesting might simply not have protested before by random chance. We re-use the propensity to protest that we constructed for our matching-based alternative identification (the construction of this propensity measure is detailed in Appendix section B.3) as a control in our regression. We first use it directly as a control (column 3 of Table 3.A.5). This holds constant the overall probability of observing BLM protests in the past, improving on identification. Our results remain robust and are more precisely estimated.

In addition, we include fixed effects for different levels of the propensity to protest. We group observations by groups of 1000, 100 and 10 units with similar propensity to protest and add fixed effects for each group. Results are shown in columns 4 to 6 of Table 3.A.5. This is essentially a matching-like strategy, where the fixed effects ensure that observations with similar propensity are compared. Results are robust to the inclusion of fixed effects for the panel of interest (panel B) and become non-significant for some specifications of the whole sample.

**Accounting for spatial correlation.** Observations are likely to be spatially correlated for several reasons. For instance, there could be spatially-correlated unobserved factors influencing the decision to protest (such as weather conditions or available TV and radio stations). Clustering by state does not entirely remove these errors because correlation across state borders remains Colella et al. (2019). To overcome this problem, we use Conley standard errors that allow for spatial correlation within a certain distance. Column 7 of Table 3.A.5 shows the estimates when allowing spatial correlation between observations in a 50 km radius. Column 8 of Table 3.A.5 shows the estimates when allowing spatial correlation with all neighboring counties. Reassuringly, our results remain robust.

**Estimation without clustering.** Our preferred specification clusters at the state level and includes state fixed effects Abadie et al. (2017). Column 9 of Table 3.A.5 shows our baseline results when we do not cluster the standard errors.

Table 3.A.1: Reduced form: superspread events on the presence of BLM events.

	(1)	(2)	(3)	(4)	(5)
<b>Panel A: All counties</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00777*** (0.00145)	0.00831*** (0.00138)	0.00620*** (0.00121)	0.00277** (0.00120)	0.00200 (0.00128)
<b>Panel B: Counties with no BLM protest before</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00548*** (0.000966)	0.00595*** (0.00178)	0.00630*** (0.00167)	0.00360*** (0.00159)	0.00303** (0.00163)
Observations	2,768	2,767	2,767	2,767	2,767
Mean dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477
<b>Panel C: Counties with BLM protest before</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00380*** (0.000760)	0.00632*** (0.00213)	0.00468*** (0.00198)	0.00110 (0.00270)	0.000100 (0.00256)
Observations	340	334	334	333	333
Mean dep. var.	0.521	0.515	0.515	0.514	0.514
State fixed effects	Y	Y	Y	Y	Y
Demographic controls	Y	Y	Y	Y	Y
Economic controls			Y	Y	Y
Political controls				Y	Y

Note: Reduced form estimation of the effect of superspread events near the county on the presence of at least one Black Lives Matter event during the three weeks following the murder of George Floyd. Panel A presents the estimation for all US counties. Panel B presents these results for the sub-sample of counties with no BLM protest before the murder of George Floyd. Panel C presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. Each column include sequentially different sets of additional controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3A.2: Instrument robustness - SSE timing and distance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Panel A: All counties</b>							
<b>IV: COVID (deaths/1000)</b>	0.215*	0.278**	0.275	0.373*	0.209*	0.225*	0.240*
	(0.121)	(0.130)	(0.175)	(0.207)	(0.121)	(0.119)	(0.120)
<b>First stage coefficient:</b>	0.00930***	0.0141***	0.00388***	0.00132***	0.00919***	0.00962***	0.0112***
	(0.00155)	(0.00215)	(0.000703)	(0.000379)	(0.00154)	(0.00164)	(0.00208)
Observations	3,106	3,106	3,106	3,106	3,106	3,106	3,106
F first stage	36.05	42.84	30.38	12.08	35.66	34.63	28.91
Mean of dep. var.	0.0988	0.0988	0.0988	0.0988	0.0988	0.0988	0.0988
<b>Panel B: Counties with no BLM protest before</b>							
<b>IV: COVID (deaths/1000)</b>	0.404**	0.503*	0.499**	0.503**	0.383**	0.440**	0.410**
	(0.187)	(0.266)	(0.191)	(0.225)	(0.188)	(0.193)	(0.187)
<b>First stage coefficient:</b>	0.00751***	0.0126***	0.00304***	0.000901***	0.00738***	0.00770***	0.00926***
	(0.00144)	(0.00331)	(0.000309)	(0.000272)	(0.00139)	(0.00154)	(0.00170)
Observations	2,767	2,767	2,767	2,767	2,767	2,767	2,767
F first stage	27.04	14.40	97.13	10.95	28.12	24.87	29.78
Mean of dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477
Distance	50 km	25 km	100 km	200 km	50 km	50 km	50 km
Lag	6 weeks	6 weeks	6 weeks	5 weeks	7 weeks	8 weeks	
All controls	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y

Note: Results on robustness of our instrument. Panel A shows the full sample, Panel B shows the sub-sample of counties with no prior BLM event. Column 1 corresponds to our baseline specification. Columns 2 to 4 vary the distance at which SSE are counted from 25 to 200km. Columns 5 to 7 vary the time lag between the murder of Floyd and the last SSE, going back 5, 7 or 8 weeks. All specifications include the whole set of controls and state fixed effects. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level.  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3A.3: Instrument robustness - SSE definition and weighting

	Presence of BLM event					
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: All counties</b>						
IV: COVID (deaths/1000)	0.215*	0.255** (0.121)	0.235 (0.109)	0.275*** (0.149)	0.307*** (0.108)	0.248* (0.139)
First stage coefficient:	0.00930*** (0.00155)	0.0100*** (0.00177)	0.00842*** (0.00176)	0.0207*** (0.00341)	0.0274*** (0.00452)	0.00965*** (0.00158)
Observations	3,106	3,106	3,106	3,106	3,106	3,105
F first stage	36.05	31.92	22.92	36.72	36.83	37.35
Mean of dep. var.	0.0988	0.0988	0.0988	0.0988	0.0988	0.0952
<b>Panel B: Counties with no BLM protest before</b>						
IV: COVID (deaths/1000)	0.404*** (0.187)	0.482*** (0.180)	0.401* (0.232)	0.515*** (0.217)	0.581*** (0.266)	0.363* (0.195)
First stage coefficient:	0.00751*** (0.00144)	0.00798*** (0.00159)	0.00653*** (0.00130)	0.0178*** (0.00448)	0.0239*** (0.00637)	0.00781*** (0.00164)
Observations	2,767	2,767	2,767	2,767	2,767	2,766
F first stage	27.04	25.03	25.08	15.73	14.09	22.55
Mean of dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477	0.0494
Excluding SSEs in prisons	Y	Y	Y	Y	Y	Y
Control SSE in county	Y	Y	Y	Y	Y	Y
Measure	linear	square	linear	square	linear	square
SSE probability weighting	Y	Y	Y	Y	Y	Y
All controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Results on the robustness of our instrument. Panel A shows the full sample. Panel B shows the sub-sample of counties with no prior BLM event. Column 1 corresponds to our baseline specification. Column 2 excludes SSEs that took place in prisons. In column 3, a control is added for the number of SSEs within the county 6 weeks before the murder of George Floyd. Columns 4 and 5 weigh the effect of SSEs by distance with smaller weights given to more distant SSEs. Weights are applied linearly (column 5), or quadratically (column 6). In column 6, observations are weighted by the inverse probability of observing a SSE affecting the county if a SSE is observed, no SSE if no SSE is observed. All specifications include the whole set of controls and state fixed effects. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.A.4: Robustness of main results - sample composition and variable definition

	Presence of BLM events								
	3 weeks	Past events	3 weeks	3 weeks	2 weeks	6 weeks	8 weeks	3 weeks	3 weeks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Panel A: All counties</b>									
IV: COVID (deaths/1000)	0.215*	-0.0523	0.253	0.324	0.0694	0.289**	0.244**	0.233*	0.233*
	(0.121)	(0.215)	(0.185)	(0.279)	(0.122)	(0.108)	(0.119)	(0.129)	(0.129)
IV: COVID (cases/1000)								0.0135*	0.0135*
								(0.00694)	(0.00694)
<b>Panel B: Counties with no BLM protest before</b>									
IV: COVID (deaths/1000)	0.404**	0.531**	0.344*	0.440**	0.575***	0.470***	0.451**	0.451**	0.451**
	(0.187)	(0.219)	(0.173)	(0.194)	(0.171)	(0.166)	(0.166)	(0.207)	(0.207)
IV: COVID (cases/1000)								0.0312***	0.0312***
								(0.0116)	(0.0116)
Observations									
F first stage	2,767	2,616	1,697	2,767	2,767	2,767	2,767	2,767	2,767
Mean of dep. var.	27.04	45.85	157.2	27.04	27.04	27.04	27.04	27.04	27.04
	0.0477	0.0428	0.0371	0.0354	0.0665	0.0763	0.0763	0.0477	0.0477
Excluding coastal									
All controls (except past BLM)	Y	Y	Y	Y	Y	Y	Y	Y	Y
COVID deaths in past 7 days									
Past BLM events	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Results on the robustness of our main results. Panel A shows the full sample, Panel B shows the sub-sample of counties with no prior BLM event. Column 1 corresponds to our baseline specification. Column 2 predicts past BLM events (likelihood of observing a BLM event between 2014 and 2019) with (instrumented) COVID deaths just before the murder of Floyd. Columns 3 and 4 exclude coastal counties and states. In columns 5, 6 and 7, the presence of BLM events is measured in the 2, 6 and 8 weeks following May 25th, e.g. Floyd's murder. Column 8 looks at the effect of COVID cases instead of deaths. Column 9 includes as an additional control the number of new COVID-19 related deaths in the 7 days leading up to Floyd's murder. All specifications include the whole set of controls and state fixed effects, except column 2 where past BLM events are removed as a control. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level.

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

Table 3.A.5: Robustness of main results - estimation method, protest propensity, spatial correlation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Panel A: All counties</b>									
IV: COVID (deaths/1000)	0.215* (0.121)		0.249** (0.114)	0.242** (0.116)	0.116 (0.129)	0.1000 (0.125)	0.215** (0.101)	0.215** (0.103)	0.215** (0.0883)
IV Probit: COVID (deaths/1000)		0.344 (0.215)							
Observations	3,106	3,106	3,002	3,106	3,106	3,106	3,106	3,106	3,106
F first stage	36.05	276.3	37.47	37.15	37.98	48.63			206.6
Mean of dep. var.	0.0988	0.113	0.102	0.0988	0.0988	0.0988	0.0988	0.0988	0.0988
<b>Panel B: Counties with no BLM protest before</b>									
IV: COVID (deaths/1000)	0.404** (0.187)		0.423** (0.185)	0.405** (0.184)	0.341* (0.186)	0.338* (0.182)	0.404* (0.234)	0.404** (0.205)	0.404*** (0.128)
IV Probit: COVID (deaths/1000)		0.878*** (0.230)							
Observations	2,767	2,767	2,663	2,767	2,767	2,767	2,767	2,767	2,767
F statistic	27.04	127.7	28.56	28.20	26.64	27.08			72.09
Mean of dep. var.	0.0990	0.0992	0.102	0.0990	0.0990	0.0990	0.0990	0.0990	0.0990
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Propensity to protest			Y						
Propensity to protest group: size		Y	Y	Y	Y	Y	Y	Y	
State clustering				Y	Y	Y	Y	Y	
Spatial clustering					Y	Y	Y	Y	
State fixed effects						Y	Y	Y	Y
Note: Results on the robustness of our main results. Panel A shows the full sample, Panel B shows the sub-sample of counties with no prior BLM event. Column 1 correspond to our baseline specification. Column 2 estimates the second stage with an IV Probit model (with an OLS in the first stage) and omits state fixed-effects. Column 3 adds a control for the propensity to protest derived from our LASSO selection model. Columns 4 to 6 add fixed effects for groups of propensity to protest of size 1000, 100 and 10 respectively. Column 7 and 8 replace the state clustering by spatial clustering, allowing correlation in a 50 km radius for column 7, and between neighbors for column 8. Columns 9 omits clustering altogether. All specifications include the whole set of controls, except column 2 where state fixed effects are removed. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level except for columns 7 to 9. *** p<0.01, ** p<0.05, * p<0.1									

## B. Appendix: Alternative Estimation Strategies

### B.1. Alternative Instrument: Florida Spring Break

In our preferred empirical strategy, we chose smaller and decentralized SSEs to argue for a causal relationship between COVID-19 and BLM protests. Here, we add another cross-sectional instrumental variable: the spatial distribution of touristic flows originating in major Florida Spring Break destinations during March of 2020. Instead of collecting information on multiple independent SSEs as in the previous section, we now focus on one single, large-scale event that is known to have contributed substantially to the spread of COVID-19 Mangrum and Niekamp (2020).

Despite the fact that COVID-19 infections had surged in Florida's main spring break destinations and despite the fact that the Center for Disease Control had issued multiple warnings, Florida Governor DeSantis failed to implement social distancing orders until April 1st 2020<sup>29</sup>. We exploit this unique, large scale event to track the diffusion of COVID-19 infections that originated in Florida during spring break and then spread across the United States. To track these movements we benefit from exceptionally rich data on cell phone mobility provided by SafeGraph. We can identify spring breakers' home counties – locations that they most likely returned to after vacationing in highly infectious spring break locations.

Specifically, we pick three Florida vacation destinations: Miami Beach, Panama Beach and Fort Lauderdale. In early March these three destinations caught the attention of the media, which reported congestion of tourists who did not respect social distancing measures (BBC, CNN). We are using anonymised mobile data for the period from March, 1, 2020 to April 1, 2020, covering the majority of spring break periods across the country. With the help of the Monthly Patterns data (MP), we measure unique devices that visited specific «points of interest» in one of three popular spring break destinations.

The SafeGraph data provides us with a rich set of points of interest, which include more than 3000 places such as restaurants, bars, hotels, gyms, public parks, malls and other establishments. Using this data, we measure the number of devices that «pinged» in each point of interest during March, 2020. The MP data also allows us to observe home locations on the level of the US Census Block Groups (CBG). An individual “home” is defined as a place where a user's devices pinged most often in the night time between 6 PM and 7 AM during the baseline 6-week period determined by SafeGraph.

Using this information, we calculate the number of unique visitors to points of interest in three cities in Florida and group this number by device home counties. Given that cell

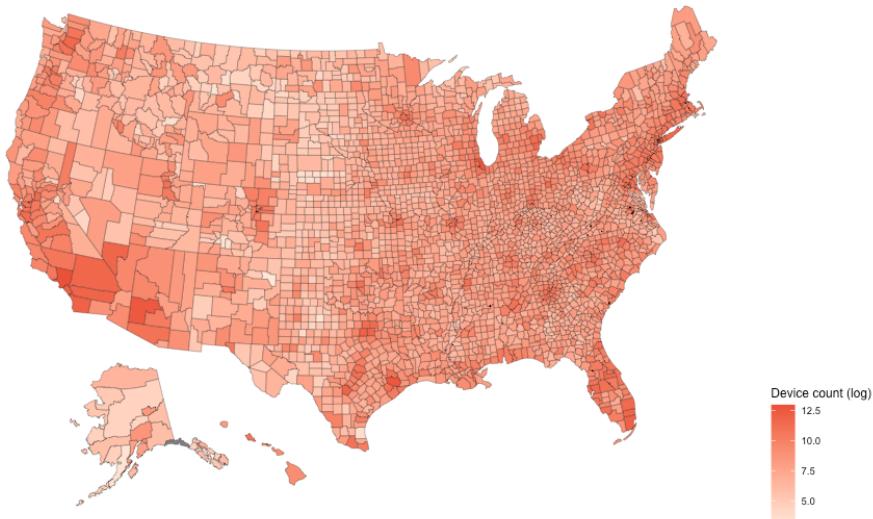
---

<sup>29</sup>Local officials had started to close some of the beaches for public access in mid March

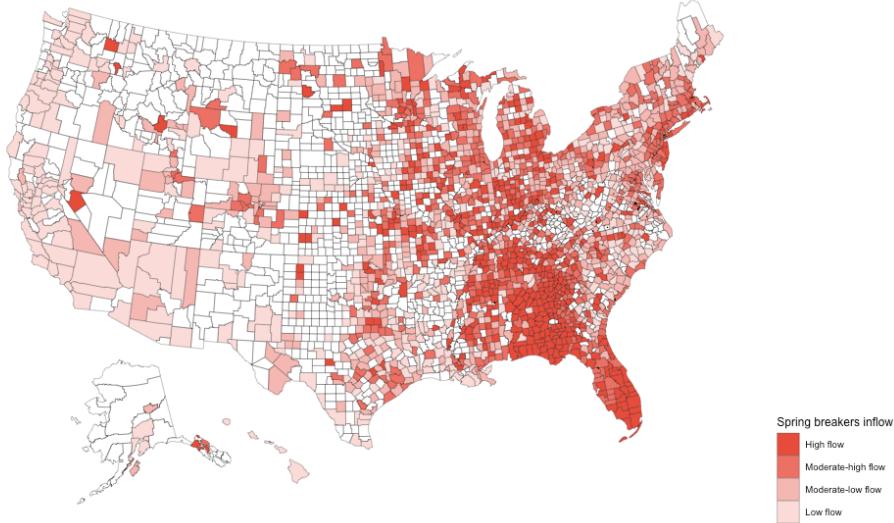
phone data is anonymized, each device is counted as many times as it has visited different places (such as restaurants and shops) in a given tourist destination. Therefore, this measure captures both intensity of tourism flow from the county and mobility of these tourists during their spring break. Since higher mobility is associated with higher chances of disease contraction, our variable captures both extensive and intensive margins of COVID-19 spread. We see this variable as an improvement over ones used in literature examining stay at home behaviour (Abouk and Heydari (2020); Lasry et al. (2020); Friedson et al. (2020); Dave et al. (2020); Dave et al. (2021)). The exposure to COVID-19 is therefore instrumented by the number of spring-break tourists.

$$Z_c = \frac{\sum_{POIs} pings_{POI,c}}{devices_c} \quad (3.6)$$

Number of devices (log) by US counties pinged during March 1st, 2020



Spring Breakers by US counties. Own visualization based on SafeGraph data.



We normalise this variable calculating a ratio of the total number of devices detected in spring breakers' home counties at March 1, 2020 to account for differences in population size and differences in resident device coverage between counties in the SafeGraph data. In Figure 3.B.1 the map of (log) number of devices by counties is presented. Figure 3.B.2 shows our resulting measure of "spring breakers" inflow split into five categories: high flow, moderate-high flow, moderate-low flow, low flow, no flow (missing).

We use the same set of controls and connotations as in our baseline cross-sectional estimation. Our estimating equation is:

$$BLM_c = \beta_0 + \beta_1 \widehat{Covid}_{cs} + \mathbf{X}_c \boldsymbol{\beta}_{\mathbf{X}} + \delta_s + \epsilon_{cs}$$

We present our 2SLS results in Table 3.B.1. We use the same set of controls as in the previous cross-sectional estimations, successively introducing socio-economic, demographic and political control variables. The inclusion of the Black population rates and Black poverty index in column 3 substantially decreases the F-Statistic (see First Stage results in Table 3.B.1). When including the full set of controls, the instrument remains at 7.3, well below the conventional threshold. However, for all specifications we find a positive coefficient for COVID-19 on the presence of a BLM event and where the first stage is sufficiently strong, we find a positive and statistically significant sign.

Table 3.B.1: Spring breakers IV: Covid-19 deaths on the presence of BLM events, 2SLS

	(1)	(2)	(3)	(4)	(5)
Presence of BLM events					
<b>Panel A: IV</b>					
COVID (deaths/1000)	0.614*** (0.218)	1.854** (0.876)	1.859* (1.011)	1.441 (0.908)	0.832 (0.697)
<b>Panel B: OLS</b>					
COVID (deaths/1000)	0.203** (0.0831)	0.158** (0.0638)	0.0758* (0.0435)	0.0382 (0.0289)	0.0323 (0.0264)
<b>Panel C: First stage</b>					
Visits per device	1.239*** (0.168)	0.595*** (0.165)	0.494*** (0.159)	0.452*** (0.158)	0.430*** (0.159)
State fixed effects		Y	Y	Y	Y
Demographic controls			Y	Y	Y
Economic controls				Y	Y
Political controls					Y
Observations	3,039	3,039	3,039	3,039	3,038
F first stage	54.41	13.06	9.677	8.223	7.305

Cross-sectional 2SLS estimation of the effect of the cumulative number of COVID-19 related deaths per thousand population the day before the death of George Floyd on the likelihood of having at least one BLM event during the first three weeks after George Floyd's death. Each column include sequentially different sets of additional controls. Demographic controls: share of Black population, urban (category [1-6]). Economic controls: median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience. Political controls: Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations), number of past BLM events between 2014 and 2019, deadly force used by police against Black people. Cross-sectional data at the county level. We report Cragg-Donald Wald F statistic. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## B.2. Difference in Differences: Notable Deaths Sample

With this empirical approach, we use data on BLM at the county-week level starting in 2014 and exploit differences in protest behavior following what we call a "notable" death. Deaths of Black people at the hands of the police have been - not only in the case of George Floyd - a trigger for BLM protests across the country. Roughly, more than 300 Black people die each year in the US either due to police brutality or under police custody. However, not all of these deaths result in media coverage, which is crucial for generating public discourse or action. Many of these events only received public traction since they were - mostly by chance - recorded on a phone camera. We construct a data set of all police-related Black deaths since July 2014 that were covered in a major national daily newspaper like the Washington Post, received TV coverage by CNN and/or has a dedicated Wikipedia page.

We now exploit the full potential of our panel data by interacting our main COVID-19 variable with a dummy variable for a notable death occurring in a certain week. Following the sample selection of our baseline estimation, we use information on BLM protests in counties in the 3 weeks after the recorded notable death (we can reduce this to 2 weeks and expand it to 4 weeks without significantly changing the first and second stage results). This data set structure allows us to observe counties' protest behavior after a protest trigger. Following a difference in differences logic, we then look at whether the reaction following this trigger differs in counties that were more exposed to the COVID-19 pandemic. Again, we use the SSE IV to account for the fact that COVID-19 exposure may be endogenous to past and present protest behavior.

$$Covid_{ct} = \zeta_0 + \zeta_1 Notable\_deaths + \zeta_2 Z_{cst} + \zeta_3 Notable\_deaths \times Z_{cst} + \mathbf{X}_{cs} \zeta_{\mathbf{X}} + \gamma_c + \theta_{st} + \eta_{cst}, \quad (3.7)$$

$$Z_{cst} = \sum SSE_{cst}^{neighbor} \quad (3.8)$$

The second stage is written as:

$$\begin{aligned} BLM_{cst} = & \beta_0 + \beta_1 Notable\_deaths_t + \widehat{\beta_2 Covid_{cst}} \\ & + \beta_3 Notable\_deaths_t \times \widehat{Covid_{cst}} + \mathbf{X}_{cs} \zeta_{\mathbf{X}} + \mu_c + \delta_{st} + \epsilon_{cst} \end{aligned}$$

where,  $Notable\_deaths_{cst}$  is a dummy variable that takes the value of one in the three weeks following a nationally covered death and zero otherwise. We include county and state-week fixed effects, as well as all police-related deaths of Black people at the county level.

This is a crucial control as it allows us to exploit the "extra" trigger that nationally covered deaths create, above and beyond the local level of deadly force used by local police. The key coefficient of interest is  $\beta_3$  which is the difference in differences estimator.

Table 3.B.2 shows the results of this estimation. Columns 1 and 3 report the effect of notable deaths up to 4 weeks after they occurred and columns 2 and 4 report for up to 3 weeks. In both cases we find that the effect of notable deaths in predicting the likelihood of observing a BLM protest is significantly higher in the presence of COVID death burden. The results control for county specific time trends as shown in columns 3 and 4.

It is important to mention, particularly in the light of new literature on generalized difference in differences - especially the designs that use two way fixed effects like our estimation model - that the underlying assumption for causal interpretation of  $\beta_3$  is that the effect of treatment, which in our case is occurrence of notable death, is homogeneous across space and time Roth et al. (2022); De Chaisemartin and d'Haultfoeuille (2020); Marcus and Sant'Anna (2021). The assumption of a homogeneous effect of notable deaths relies on the fact the occurrence of these deaths is random and their location and time cannot be predicted. Therefore, each county of the country has an equally likely probability of being affected by this. While the exposure to COVID-19 is staggered in time across the USA, in this estimation we assume all counties to be equally exposed to the COVID-19 pandemic since it broke out in the US in January 2020.

Table 3.B.2: Notable Deaths Regression

	(1)	(2)	(3)	(4)
	Presence of BLM			
Covid deaths per thousand	0.0595*** (0.0166)	0.0597*** (0.0166)	0.0450*** (0.0116)	0.0451*** (0.0116)
Notable deaths × Covid deaths	1.4926*** (0.1053)	2.0714*** (0.1095)	1.4935*** (0.1057)	2.0707*** (0.1102)
Notable deaths	-0.0389*** (0.0125)	-0.0391*** (0.0128)	-0.0410*** (0.0127)	-0.0412*** (0.0130)
Black police-related deaths	Y	Y	Y	Y
Unemployment	Y	Y	Y	Y
Weeks post Notable Death	4	3	4	3
County FE	Y	Y	Y	Y
State-Week FE	Y	Y		
County Week Trend			Y	Y
Observations	96286	96286	96329	96329
F First Stage (COVID)	18.03	17.92	32.23	32.09
F First Stage (Interaction)	13.05	13.87	14.59	14.97

Note: Estimation of the effect of Notable deaths and COVID-19 deaths on different Black Lives Matter measures. This table presents 2SLS results, using the cumulative number of all super-spreader events in neighbouring counties (50km radius) as an instrument. Columns (1) and (3) presents the effect of instrumented cumulative number of COVID-19 deaths and notable deaths on the likelihood of having a BLM event in the county within 4 weeks of the notable death. Column (2) and (4) presents the effect of instrumented cumulative number of COVID-19 deaths and notable deaths on the likelihood of having a BLM event in the county within 3 weeks of the notable death. All specifications include county fixed effects and two time varying controls (the number of black police-related deaths and the unemployment rate both at a county level) along with either state-week fixed effects or county week time trend to increase precision. Weekly data by county from year 2014 until the 14th June 2020. Standard errors clustered at the county level are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

### B.3. LASSO Matching: Propensity to Protest

We again exploit data on past protests, this time to predict the propensity of a county to protest in response to a notable death using a wide variety of observable county characteristics.

More precisely, we start by estimating the following logit model:

$$\log \frac{\Pr(BLM_{ci} = 1)}{1 - \Pr(BLM_{ci} = 1)} = \beta_0 + \beta_1 X_c + \varepsilon_{ci}$$

We select the most relevant subset of variables with LASSO regression Tibshirani (1996). This avoids overfitting and gives confidence in using the model to predict the propensity to react to another notable death. This model is estimated on the subset composed on all counties, and we compute the estimated propensity to protest for each county.

We then perform a propensity score matching-like estimation: we consider the binary treatment where counties are considered treated if they had at least one COVID-19 related death on or before May 24th. We match counties with similar historical propensities to protest, and consider as the outcome where these counties held a BLM protest in the 3 weeks following the murder of George Floyd. The results are presented in Table 3.B.3 for the whole sample, and the subsamples of counties that did and did not protest before. For each of these samples, the propensity-to-protest model is estimated on the whole sample. The results in each case are positive and significant; their magnitude is not comparable with our main specification as the treatment is different. Unlike our main specification, with this estimation strategy, the effect on counties that had BLM events is significant and much higher in magnitude than the effect on counties that did not have BLM events before. This might be consistent with a multiplicative effect of protest: the relative increase (relative to the probability of having a BLM event after the death of George Floyd) is roughly similar.

Note that this is not a proper propensity score matching Rosenbaum and Rubin (1983): we are matching not on the propensity to have a COVID death but on the (past) probability to hold a protest. With an usual propensity score matching, we would need to be concerned about unobservable characteristics of the county that affect both the treatment probability and the outcome. In this case, we can also get bias from observable characteristics of the counties that may influence the probability of treatment and protests, but did not influence the past propensity to protest as much. One such example would be the quality of the health system: it raises both the probability of deaths from COVID, and people are likely more concerned about the quality of the health care system than they were for past protests. In the robustness checks section, we use this propensity as a control in our main specification instead.

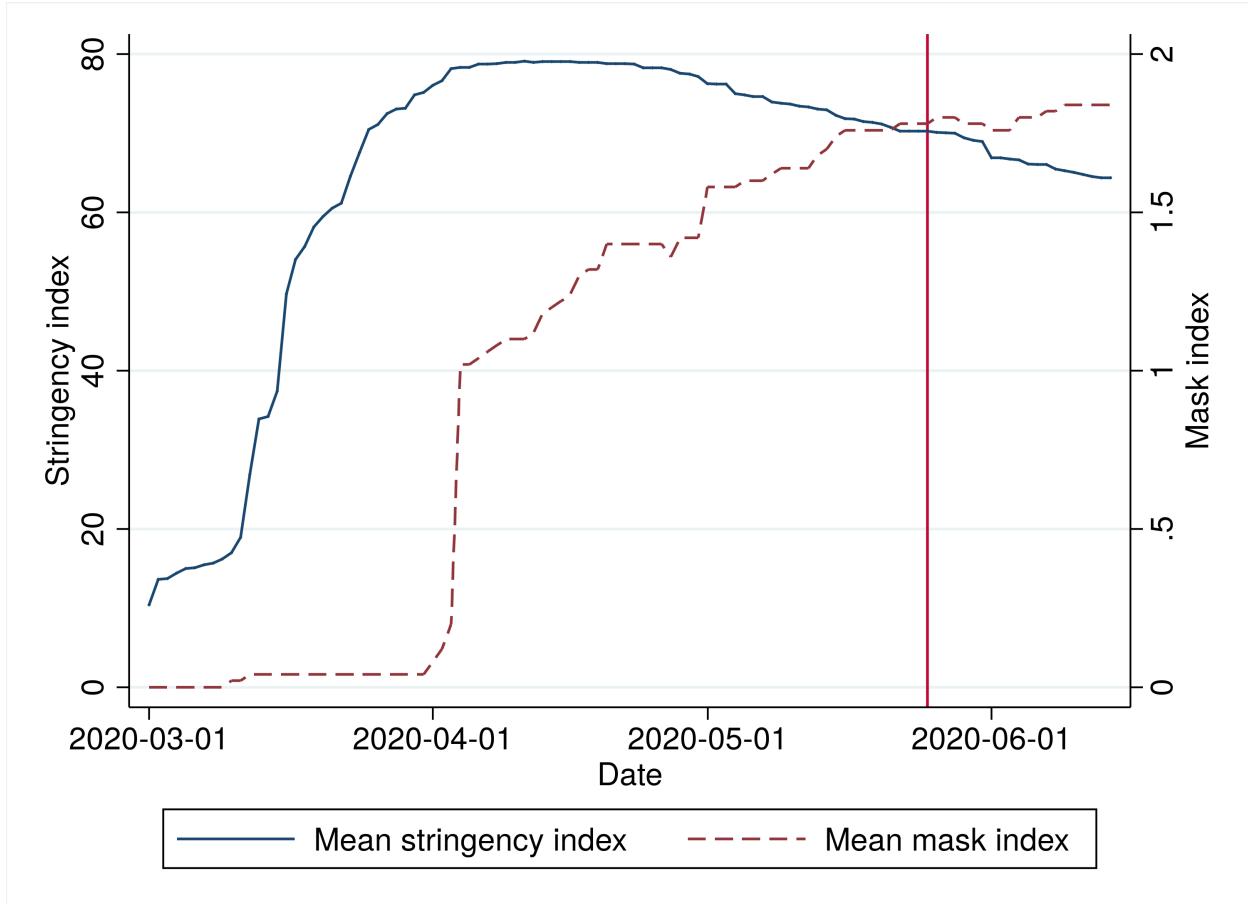
Table 3.B.3: Matching on past propensity to protest

	(1)	(2)	(3)
	Presence of BLM events		
	All counties	Never protested before	Protested before
Average Treatment Effect	0.117*** (0.0110)	0.0439*** (0.00866)	0.324*** (0.0537)
Observations	3,108	2,768	340
Mean of dep. var.	0.0994	0.0477	0.521
Propensity to protest	Y	Y	Y

Note: Estimation of the effect of having at least one COVID-19 death on presence of BLM protests. The average treatment effect is evaluated by matching on the past propensity to protest after a notable death. Column 1 presents the results for the whole sample, column 2 for counties that never protested before and column 3 for counties that did protest before. Propensity-to-protest model estimated on the full sample using logit LASSO regression using all available controls. Standard errors (in parentheses) are not clustered. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

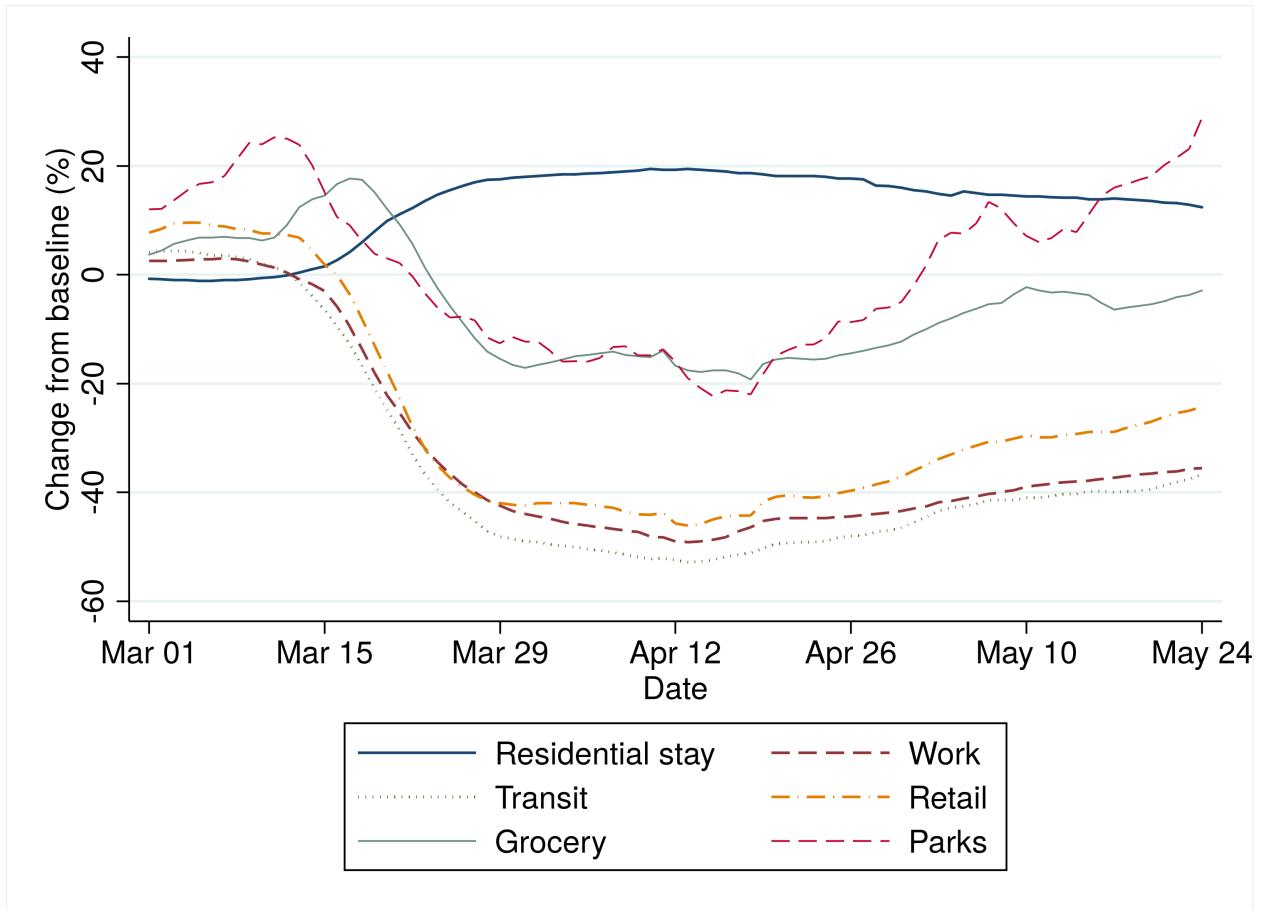
## C. Appendix: Additional Figures and Tables

Evolution of lockdown stringency index, and masks recommendations



Note: This graph represents two indicators of average health and lockdown measures in the US over the period from March 1st to June 14th 2020. The blue continuous lines represent the mean lockdown stringency index. The red dashed lined isolates only the indicator for mask recommendations and mandates. The vertical line corresponds to the murder of George Floyd.

## Evolution of mobility index



Note: This graph represents the components of the Google Community Mobility index: residential stay, and mobility to different types of places, between March 1st and May 24th, 2020. The index is relative to the average mobility to these places in the same day of the week between January 3 and February 6, 2020. The displayed value is an average of the 7 previous days.



Table 3.C.1: Summary statistics - counties with and without prior BLM event

	From 25th of May to 14th of June 2020:				All counties				No BLM event before				Has BLM event before			
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	
Presence of BLM events	3106	0.099	0.298	0.000	1.000	2768	0.048	0.213	0.000	1.000	338	0.518	0.500	0.000	1.000	
Number of BLM events	3106	0.250	1.348	0.000	36.000	2768	0.064	0.322	0.000	5.000	338	1.778	3.642	0.000	36.000	
Participants in BLM events	3106	270.759	5968.521	0.000	323687.500	2768	1.026	172.090	0.000	5500.000	338	2315.911	17979.700	0.000	323687.500	
Participants per event	3106	539.141	878.429	0.000	8991.319	132	355.115	621.538	0.000	5500.000	175	677.149	1010.498	0.000	8991.319	
Tweets mentioning BLM	3106	819.502	7187.496	0.000	243596.000	2768	322.706	4818.930	0.000	243506.000	338	4887.938	16330.354	0.000	177550.000	
New users tweeting about BLM	3106	4.586	53.812	0.000	2442.000	2768	1.809	25.763	0.000	1306.000	338	27.331	143.697	0.000	2442.000	
Followers of @BkLivesMatter created during the pandemic	3106	1.540	11.207	0.000	453.000	2768	0.441	2.167	0.000	78.000	338	10.536	32.057	0.000	453.000	
Tweets mentioning #AllLivesMatter	3106	134.741	833.066	0.000	28943.000	2768	47.488	326.063	0.000	15659.000	338	849.290	2224.119	0.000	28943.000	
Tweets mentioning #BlueLivesMatter	3106	17.753	113.478	0.000	4117.000	2768	6.125	36.206	0.000	1647.000	338	112.976	312.535	0.000	4117.000	
Neighbor protested first	3106	0.348	0.477	0.000	1.000	2768	0.334	0.472	0.000	1.000	338	0.464	0.499	0.000	1.000	
<b>On the 25th of May 2020:</b>																
COVID deaths (total)	3106	24.461	141.132	0.000	3304.000	2768	8.366	46.396	0.000	1025.000	338	156.266	382.483	0.000	3304.000	
COVID cases (total)	3106	459.678	2438.202	0.000	72010.000	2768	164.485	663.300	0.000	15169.000	338	2877.112	6677.133	0.000	72010.000	
COVID deaths (per 1000)	3106	0.113	0.248	0.000	2.935	2768	0.099	0.230	0.000	2.935	338	0.224	0.345	0.000	2.010	
COVID cases (per 1000)	3106	2.791	5.664	0.000	145.513	2768	2.596	5.662	0.000	145.513	338	4.391	5.430	0.000	40.048	
Superspread events, 6+ weeks ago, neighbouring	3106	3.070	9.790	0.000	143.000	2768	2.327	7.564	0.000	143.000	338	9.154	19.279	0.000	140.000	
Black death burden	3106	1.346	0.963	0.000	4.104	2768	1.344	0.985	0.000	4.104	338	1.363	0.755	0.000	4.104	
Lockdown stringency index	3106	68.445	8.508	47.220	89.810	2768	68.210	8.543	47.220	89.810	338	70.367	7.969	47.220	89.810	
<b>Before the 25th of May 2020:</b>																
Google searches for Twitter	3056	61.265	11.222	17.000	100.000	2731	60.523	10.941	17.000	100.000	325	67.505	11.623	41.000	100.000	
Residential stay	1348	10.633	3.387	3.600	26.286	1023	10.006	3.016	3.600	26.286	325	12.606	3.722	4.429	26.143	
<b>Later outcomes:</b>																
Followers of @BkLivesMatter*	3106	63.198	495.174	0.000	20058.000	2768	16.186	94.110	0.000	3304.000	338	448.195	1421.137	0.000	20058.000	
Street art count	3106	0.703	26.735	0.000	1467.000	2768	0.099	0.206	0.000	10.000	338	6.382	80.924	0.000	1467.000	
<b>County characteristics:</b>																
Black police-related deaths (2014-2019)	3106	0.677	3.207	0.000	84.000	2768	0.207	0.724	0.000	15.000	338	4.527	8.589	0.000	84.000	
Black police-related deaths (2020)	3106	0.047	0.301	0.000	6.000	2768	0.014	0.131	0.000	3.000	338	0.314	0.783	0.000	6.000	
Unemployment rate (year average)	3106	4.691	1.550	0.708	19.650	2768	4.713	1.575	0.708	17.442	338	4.506	1.323	2.492	19.650	
Black population share	3106	0.100	0.147	0.000	0.875	2768	0.093	0.146	0.000	0.875	338	0.157	0.142	0.009	0.727	
Non-white population share	3106	0.144	0.162	0.000	0.928	2768	0.134	0.160	0.000	0.928	338	0.231	0.150	0.014	0.801	
Large cities	3106	0.020	0.140	0.000	1.000	2768	0.001	0.027	0.000	1.000	338	0.178	0.383	0.000	1.000	
Suburban areas	3106	0.118	0.323	0.000	1.000	2768	0.105	0.307	0.000	1.000	338	0.228	0.420	0.000	1.000	
Smaller towns	3106	0.234	0.423	0.000	1.000	2768	0.201	0.400	0.000	1.000	338	0.506	0.501	0.000	1.000	
Rural areas	3106	0.628	0.483	0.000	1.000	2768	0.694	0.461	0.000	1.000	338	0.089	0.285	0.000	1.000	
BLM events (2014-2019)	3106	0.617	4.183	0.000	117.000	2768	0.000	0.000	0.000	0.000	338	5.672	11.510	0.000	117.000	
Black poverty rate	3106	0.281	0.225	0.000	1.000	2768	0.283	0.236	0.000	1.000	338	0.263	0.099	0.000	0.600	
Population share with 3+ risk factors	3106	25.899	5.019	10.685	48.448	2768	25.957	5.066	10.685	48.448	338	25.428	4.600	11.763	39.453	
Vote share for republicans (2016)	3106	0.633	0.156	0.083	0.960	2768	0.656	0.141	0.083	0.960	338	0.446	0.143	0.084	0.818	
Vote share for republicans (2012)	3106	0.596	0.148	0.060	0.959	2768	0.614	0.140	0.060	0.959	338	0.456	0.131	0.092	0.823	
Median household income (2016)	3106	48795.991	13277.575	20170.891	129150.343	2768	47521.697	12362.349	20170.891	129150.343	338	59231.630	15713.932	28625.618	12036.813	
Social capital	3106	1.384	0.705	0.000	6.887	2768	1.426	0.726	0.000	6.887	338	1.037	0.334	0.274		
Distance to Minneapolis	3106	1216.679	555.825	11.998	6474.706	2768	1192.851	538.680	34.438	6474.706	338	1411.818	648.904	11.998	6395.519	
Notable Deaths	3106	0.010	0.116	0.000	3.000	2768	0.001	0.033	0.000	1.000	338	0.080	0.331	0.000	3.000	
Log(SXSW followers created before March 2017)	3106	0.114	0.258	0.000	1.474	2768	0.090	0.228	0.000	1.427	338	0.311	0.378	0.000	1.474	
Log(SXSW followers created during March 2017)	3106	0.193	0.350	0.000	1.658	2768	0.157	0.312	0.000	1.658	338	0.489	0.482	0.000	1.619	

Table 3.C.2: Summary statistics for super spreading events by their type

Type of SSE event	Total events	Total Events 6 weeks before GF's murder	Mean	Standard Deviation	Total Cases
Community	11	9	1.364	0.505	504
Development Center	12	12	3.833	1.404	1612
Event/group gathering	21	13	3	1.549	1083
Industry	125	87	15.656	8.642	17825
Medical	140	134	36.586	17.037	13731
Nursing Home	273	261	80.597	37.073	26684
Prison	193	187	45.487	19.674	49747
Rehabilitation / Medical	262	251	89.618	41.009	26979
Restaurant/Bar	8	4	1.5	0.535	1306
Retail	5	0	1	0	68
School	7	2	1.286	0.488	218
Other	20	15	2.5	1.051	1592

All super spreading (SSE) in the USA by their type. Total events are total number of SSE event of each type occurring till 29 August. Total Events 6 weeks before GF's murder is sum of all SSE events by their type that occurred 6 weeks before GF's death. Total cases is sum of all reported COVID-19 positive cases attributed to each type of SSE event.

Table 3.C.3: First stage

	COVID (deaths/1000)				
	(1)	(2)	(3)	(4)	(5)
<b>Panel A: All counties</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.0114*** (0.00201)	0.0114*** (0.00201)	0.0105*** (0.00201)	0.00935*** (0.00151)	0.00930*** (0.00155)
Observations	3,107	3,107	3,107	3,106	3,106
F statistic	31.92	31.92	27.44	38.38	36.05
Mean dep. var.	0.114	0.114	0.114	0.113	0.113
<b>Panel B: Counties with no BLM protest before</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00881*** (0.00132)	0.00881*** (0.00132)	0.00797*** (0.00147)	0.00772*** (0.00146)	0.00751*** (0.00144)
Observations	2,767	2,767	2,767	2,767	2,767
F statistic	44.53	44.53	29.25	27.95	27.04
Mean dep. var.	0.0990	0.0990	0.0990	0.0990	0.0990
<b>Panel C: Counties with BLM protest before</b>					
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.0126*** (0.00205)	0.0126*** (0.00205)	0.0121*** (0.00214)	0.00942*** (0.00174)	0.00961*** (0.00181)
Observations	334	334	334	333	333
F statistic	37.56	37.56	32.01	29.27	28.09
Mean dep. var.	0.233	0.233	0.233	0.227	0.227
State fixed effects		Y	Y	Y	Y
Demographic controls			Y	Y	Y
Economic controls				Y	Y
Political controls					Y

Note: Estimation of the SSE in neighbouring counties (50km radius) six weeks prior to George Floyd's murder on COVID-19 deaths. Panel A presents estimation for all US counties. Panel B presents these results for the sub-sample of counties with no BLM protest before the murder of George Floyd. Panel C presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. Each column include sequentially different sets of additional controls. Demographic controls: share of Black population, urban (category [1-6]). Economic controls: median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience. Political controls: Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations), number of past BLM events between 2014 and 2019, deadly force used by police against Black people. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.C.4: COVID deaths interacted with county characteristics - Counties without BLM events before

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Sample: Counties with no BLM protest before</b>									
COVID (deaths/1000)	0.404** (0.187)	-0.776 (0.920)	0.421 (0.294)	-0.214 (0.257)	1.083* (0.618)	-0.0160 (0.117)	0.134 (0.221)	0.457** (0.201)	0.433** (0.195)
... × Non-Black population share	1.305 (1.048)								
... × Non-white population share		-0.0523 (1.003)							
... × Median household income			0.00703** (0.00294)						
... × Vote Republican 2016				-1.386 (1.236)					
... × Not large cities					0.430** (0.0955)				
... × Suburban areas						0.333* (0.176)			
... × Smaller towns							-0.320 (0.237)		
... × Rural areas								-0.179 (0.187)	
Interacting variable									
	-0.143 (0.309)	-0.102 (0.0822)	0.00131 (0.000856)	-0.328* (0.183)	0.227** (0.0731)	-0.0416 (0.0284)	0.0701** (0.0303)	-0.0186 (0.0228)	
Observations	2,767	2,767	2,767	2,767	2,767	2,767	2,767	2,767	2,767
F COVID	27.04	15.24	11.86	45.98	14.34	114.1	18.50	23.50	12.70
F interaction		14.71	11.17	96.93	18.67	159.4	71.80	92.71	6.328
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 inhabitants on first-time BLM protest, interacted with county characteristics. We present results for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and all standard controls. Standard errors (in parentheses) are clustered at the state level. \*\* p<0.01, \* p<0.05, \* p<0.1

Table 3.C.5: Alternative Mechanisms

	Presence of BLM			Other Protests	COVID-19 Protests	
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Sample: All counties</b>						
COVID (deaths/1000)	0.279** (0.119)	0.570* (0.289)	0.252 (0.424)	0.890 (1.066)	0.180 (0.138)	0.225 (0.104)
... × Black death burden	1.017 (0.888)					
... × Google BLM search		-0.015 (0.010)				
... × Unemployment			0.006 (0.030)			
... × Stringency				-0.007 (0.0146)		
Interacting variable	-0.195 (0.176)	0.001 (0.001)	0.008* (0.005)	0.001 (0.0013)		
Observations	3,106	3,056	1,351	3107	3,106	3,106
F stat COVID	25.59	22.14	27.49	96.71	31.4	31.4
F stat Interaction	12.46	58.19	27.49	96.04		
Mean of dependent variable	0.099	0.099	0.099	0.099	0.081	0.030
All controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y		Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on presence of BLM protest. Column 1 shows estimates for instrumented COVID deaths. Columns 2 to 4 show heterogeneous effects for Black death burden weeks prior to GF's murder, Google searched for BLM 3 weeks prior to GF's murder, unemployment and stringency 3 weeks after GF's murder. Column 5 presents results for other protests. Panel A presents 2SLS estimation for all counties. Panel B presents these results for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.C.6: Effect of SXSW users on Twitter presence

VARIABLES	(1) Log(Preexisting users)	(2) Log(New users)	(3) Presence of BLM events
Log(SXSW users)	0.373*** (0.103)	0.193*** (0.0505)	0.0151 (0.0175)
SSE			-0.00117 (0.00257)
× SXSW users			0.00439** (0.00172)
Mean of dep. var	1.738	0.420	0.0477
F first stage	13.02		
Observations	2,767	2,767	2,767
<hr/>			
Instruments			
All controls	Y	Y	Y
Pre-SXSW users	Y	Y	Y
State fixed effects	Y	Y	Y

Note: Column 1 shows the first stage regression for predicting existing Twitter users at the end of 2019 in the county using SXSW followers that joined Twitter during the festival in the county and its neighboring counties. Column 2 shows the same effect on the users created during COVID-19. Column 3 shows the reduced-form effect of SXSW followers interacted with superspread event on the presence of protest. We present results for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and all standard controls. Standard errors (in parentheses) are clustered at the state level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3.C.7: Principal component analysis of online presence

(a) Correlation between measures						
	New Twitter accounts	New Twitter accounts (log)	Google searches for Twitter	Google searches for Twitter	Residential stay	New accounts following @BlklivesMatter
New Twitter accounts	1					
New Twitter accounts (log)	.379***	1				
Google searches for Twitter	.0558*	.234***	1			
Residential stay	.0770*	.355***	.520***	1		
New accounts following @Blklivesmatter	.397***	.450***	.127***	.215***	1	
New accounts following @Blklivesmatter (log)	.279***	.735***	.225***	.404***	.567***	1

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(b) Principal components						
	Eigenvalue	Difference	Proportion	Cumulative		
PC1	2.764871	1.494786	0.4608	0.4608		
PC2	1.270085	.5323919	0.2117	0.6725		
PC3	.7376932	.187934	0.1229	0.7954		
PC4	.5497592	.0989178	0.0916	0.8871		
PC5	.4508414	.2240914	0.0751	0.9622		
PC6	.22675	.	0.0378	1.0000		

(c) Factor loadings						
	PC1	PC2	PC3	PC4	PC5	PC6
New Twitter accounts	.3087237	-.4361973	.7768168	-.2448021	.1515153	.1676077
New Twitter accounts (log)	.4987554	-.1210586	-.2494509	-.5052231	-.2346866	-.6033522
Google searches for Twitter	.2796974	.6289325	.387592	.1363476	-.5946134	.0618868
Residential stay	.3655117	.5431075	.0268323	-.0207581	.7495347	-.0920787
New accounts following @Blklivesmatter	.42358368	-.3098153	-.0348433	.8105414	.0093471	-.2537088
New accounts following @Blklivesmatter (log)	.5126036	-.0925705	-.4268119	-.0939284	-.0806687	.7288118

Note: The first table reports the correlation between the online presence measures. The second table reports the eigenvalues of the six principal components. The third table reports the loading of the different components.

## D. Data Appendix

**Super Spreader Events.** Our identification strategy relies on records of Super Spreader Events in the early stages of the pandemic. In this section, we discuss the limitations of the SSE data set and how we address these in the empirical section. The data set is collected from various sources by researchers from the London School of Hygiene and Tropical Medicine and published as a free access data base for researchers and the media under the *SARS-CoV-2 Superspreading Events from Around the World* Project.

A main challenge in the construction of this data base is that there is no standard definition of a Super Spreader Event. The data base mainly refers to "outbreak" and "clusters" for which they use the UK Government Public Health Definition: "two or more test-confirmed cases of COVID-19 among individuals associated with a specific non-residential setting with illness onset dates within a 14-day period." The outbreak definition is expanded to "identified direct exposure between at least 2 of the test-confirmed cases in that setting (for example under one metre face to face, or spending more than 15 minutes within 2 metres) during the infectious period of one of the cases when there is no sustained local community transmission - absence of an alternative source of infection outside the setting for the initially identified cases."

The data base draws from one main source: Leclerc et al. (2020) who performed a systematic review of available literature and media reports to find settings reported in peer reviewed articles and media with "outbreak" or "cluster" characteristics. There were various extensions to this data set, using articles of journalists, expanding that data set to second and third generation events by Swinkels (2020), and including the Western Pacific Region for a project of the World Health Organisation (under the project lead of Fatim Lakha, also from the London School of Tropical Medicine and Hygiene). We will primarily draw from Leclerc et al. (2020), as we focus on SSEs in the United States during the early stages of the pandemic.

There are various limitations in the measurement of SSEs. First, there exists some uncertainty about the exact date of the SSE. If, for instance, there was a COVID-19 cluster at a worker dormitory, the exact date of the transmission event is difficult to narrow down. In these cases, researchers make an approximation based on the timing of tests and overall case numbers. We address this concern by using the cumulative number of SSEs until a certain cut-off date (first week of April in the baseline version of the instrument), thereby not relying on the specific timing of the SSE. Second, for many SSEs it is not known exactly how many people were infected (either directly at the SSE or by somebody who was infected at the initial SSE). The database always uses the lowest number cited in the articles about

the SSE but actual numbers can be much higher. The actual detected number of cases will be related to testing capacity and potentially other unobserved factors at the county level. For this reason, we use the most simple version of the instrument, i.e. counting the number of SSEs rather than using the cases associated with the SSE. Third, the GPS coordinates of SSEs are almost always approximate. For instance, when an SSE occurred somewhere in city A, typically the database uses GPS coordinates for a random location within that city, not the for precise location. In a robustness check, we make sure that our results are not sensitive to changing the radius around SSEs to account for potential measurement error. Overall, the measurement error in Super Spreader Events would only bias our results if it is somehow related to the counties' overall propensity to protest (and is not captured in the set of controls or state fixed effects). One important exercise, addresses this concern: SSEs do not predict past BLM events. If SSEs were disproportionately recorded in places with a higher likelihood of a BLM event occurring, we should see a systematic relationship to previous BLM protest, which is not the case.

**Twitter usage during the protests.** Twitter data is an important source when studying social events and protests. Previous work on BLM events has used this data Ince et al. (2017). We collected tweets using the Twitter Academic Research API. In particular, we collected all tweets that contain the keywords “BLM”, “Black Lives Matter”, “Black Life Matters” or “George Floyd”,<sup>30</sup> including retweets, between May 25 and June 14. For each tweet, we extract the time and text of the tweet, the user, the user’s stated location, and account creation date. We present a selection of tweets that are part of our sample in Table 3.D.4.

**Geo-location of tweets.** We follow the literature in assigning the location of a tweet or a user by extracting information on their self-reported location from their Twitter profile Enikolopov et al. (2020); Takhteyev et al. (2012); Müller and Schwarz (2020). Not all users report a location and among those who do, not all state a valid location (e.g., “in the heart of Justin Bieber”) so we restrict the sample to the users that state a valid location that can be matched to a USA county (in particular, we exclude users whose location only mentions a state). The location is an arbitrary text field which is not meant to be machine-readable. We use the Nominatim geocoding engine (based on the Open Street Map database) to find the coordinates of the most likely match for the location. We then filter out all locations outside the US and all locations that are too vague (i.e. that map the whole country or a whole state). Finally, we map these coordinates to counties using the US Census Bureau

---

<sup>30</sup>These keywords are considered both in when appearing separated with space, or without spaces as a hashtag (e.g. #BlackLivesMatter)

cartographic boundary files. Across our different tweet collections, we end up with 23.3 million tweets. This approach has clear limitations as it relies only on self-reported locations and may not be representative of the whole Twitter universe. We report summary stats on the counties for which we were able to assign tweets and compare them to the characteristics of the full set of counties in Table 3.D.3. We would be particularly concerned if counties with geolocalizable tweets were substantially different from other counties. Reassuringly, counties without localizable tweets only form a tiny minority: out of the 3106 counties in our universe, only 21 (0.7%) are not attributed any tweet.

**Pre-existing Twitter usage and instrument.** For the study of mechanisms, we use a proxy of pre-existing Twitter usage measured in December 2019. This is measured by sampling all tweets containing the word "the" during random intervals in one week of December 2019. One million tweets were collected from 765 000 users. Users were attributed to counties using the location in their profile. To study causally the effect of pre-existing Twitter usage on the reaction to COVID-19, we collected data to reproduce the SXSW instrument used by Müller and Schwarz (2019): we collected in November 2021 the locations of all 639 915 followers of the @SXSW Twitter account as well as the date they joined the network.

**BLM account followers.** As an additional outcome, we use the number of all followers of the official BLM account @Blklivesmatter. We collected the followers and their geolocation in February 2022. This gap between the period of analysis and the date of data collection can lead to measurement error because we do not know the starting date of following. Accounts that followed the official BLM account may stop following it and accounts that are computed as followers may start following just a few hours before the collection. Similarly, geolocation of accounts may have changed between the period of study and the date of data collection. Using this data we also compute the number of accounts created between the first COVID-19 death in the USA and the 24th of May (the day before the murder of George Floyd) that are followers of the account @Blklivesmatter.

**Google Searches.** We also use the Google Trends data to analyze patterns of search activity before and after the death of George Floyd. Each variable is a normalized index of search activity for a given search term. The indices are specified on a Nielsen's Designated Market Area (DMA) level. A DMA is a region of the United States that consists of counties and ZIP-codes. There are 210 DMA regions covering the US. Search activity is averaged across the period of interest: each observation is a number of the searches of the given term divided by the total searches of the geography and time range, which is then normalized

between regions such that the region with the largest measure is set to 100. The important limitation of the Google Trends data is that an index of search activity is an integer from zero to one hundred with an unreported privacy threshold. The search terms that were used in the analysis are presented in Table 3.D.2.

**SafeGraph.** We rely on two data sets provided by SafeGraph. Both of them are based on anonymized mobile data. SafeGraph aggregates data from around 45 million smartphones on the level of US Census Block Groups. With the help of the first data set, Monthly Patterns (MP), we can answer such questions as: who visited each «point of interest», where they came from and where they go to. The set of «points of interests» consists of millions of places such as hotels, restaurants, public parks, malls and other establishments. The MP data allows us to observe home locations at the level of the US Census Block Group, which we can use to construct our variable of touristic flows out of spring break locations in March 2020. In our alternative identification strategy we employ an instrumental variable based on data provided by the data company SafeGraph. The SafeGraph data is GPS location data that reveal the spatial mobility of population between the points of interest. For the region of interest (three vacation destinations in Florida: Miami Beach, Panama Beach and Fort Lauderdale) the SafeGraph data provide rich set of points of interest, which include more than 3000 places such as restaurants, bars, hotels, gyms, public parks, malls and other establishments. Using this data, we measure the number of devices that “pinged” in each of the point of interest during March, 2020. Using these data we can also observe home locations on the level of the US Census Block Groups (CBG). An individual “home” is defined as a place where user’s devices pinged most often in the night time between 6 PM and 7 AM during the baseline 6-week period determined by the SafeGraph.

**Elephrame.** Elephrame is a crowd-sourced platform that collects data on Black Lives Matter and other protests. It provides information on the place and date of each BLM protest and estimated number of participants, as well as a link to a news article covering the protest. We extracted all protest records from June 2014 to September 2020 and geo-coded their location. The observation period starts with the first BLM demonstration for Eric Garner on 7/19/2014 and consists of any public demonstration or public art installation focused on “communicating the value of a Black individual or Black people as a whole”. Each observation is manually collected by the creator of Elephrame, Alisa Robinson, from sources that include press, protest organizers, participants and observers.

**Lockdown stringency.** We use data from the Oxford COVID-19 Government Response Tracker Hale et al. (2020) to measure the restrictiveness of the government’s pandemic policy. Use of this data is inspired by recent work which shows that stringent policies lead to lower mortality, mobility and consequently spread of infection during the pandemic Jinjarak et al. (2020); Askitas et al. (2020). This data provides four key indices (i) an overall government response index, (ii) a containment health index, (iii) an economic support index, and (iv) an original stringency index which captures the strictness of lockdown-style policies. Each of this indices reports values between 1 and 100 and varies across states and weeks.

**Community Resilience.** One of the most important COVID-19 related control variables used in our empirical analysis is the ability of counties to cope with the pandemic. This variable comes from the United States Census Bureau. These estimates measure the capacity of individuals and households to absorb, endure, and recover from the health, social, and economic impacts of a disaster such as a hurricane or a pandemic. For each county the population living under each of 11 risk factors is estimated and these factors are aggregated into 3 composite risk factors: (i) population with 0 risk factors; (ii) population with 1-2 risk factors, and (iii) population with 3 or more risk factors. These risk factors are based on households’ and individuals’ socio-economic and health conditions. Risk factors include: Income-to-Poverty Ratio, single or zero caregiver household, unit-level crowding defined as  $\geq 0.75$  persons per room, communication barriers (defined as either limited English-speaking households or no one in the household over the age of 16 with a high school diploma), no one in the household is employed full-time, disability posing constraint to significant life activity, no health insurance coverage, being aged 65 years or older, households without a vehicle and households without broadband Internet access. For our analysis we look at populations within each county that are classified as living under 1-2 risk factors and 3 or more risk factors.

**Notable Deaths.** We collect data on all notable Black deaths that have occurred in the country since 2014. Notable deaths are defined as deaths of Blacks at the hands of a police officer and which are covered in national media and/or have a dedicated Wikipedia page. This data set includes personal information about the victim like their name, age, sex and race. It also has details about the event, like the county and zip code of the place where shooting took place, cause of death, whether the victim was armed, if a video of the incidence was taken by onlookers and if the police officer wore a body camera. We also collect information on date of the shooting, date of the official verdict from this incident and whether the police officer was convicted. From 2014 till 2020, we have 34 notable deaths from all over the

country. Average age of victim is 34 years, 31 out of 34 are men. All victims in our data are Black.

**Use of deadly force by police.** We obtain this from the collaborative platform Fatal Encounters. This data is collected by a multi-disciplinary team at the University of Southern California. The results are published as part of the *National Officer-Involved Homicide Database*. The data is available from 2000 onward and contains the name, gender, race, and age of each victim and the specific address where the death occurred, among other variables.

**George Floyd Street Art.** We extract information on the location of street art representing or referring to George Floyd from the Urban Art Mapping George Floyd and Anti-Racist Street Art database. The crowd-sourced website run by researchers from the University of St. Thomas documents street art from around the world created in the aftermath of the murder of George Floyd. Their archive is a repository of images made available for research and education. The website contains geo-tagged information and images of George Floyd related street art, which we match to counties. The data does not contain time stamps and has no information on when these images were added. For this reason, we can only interpret the street art as cross-sectional snapshots at time of accessing the website in January of 2022. Overall, we record 2183 images across 70 counties. Most of the images (1467) are recorded in Minneapolis.

Table 3.D.1: Description of variables and data sources

Variable	Source	Start of period	End of period
<b>BLM related variables</b>			
Presence of BLM events	Elephrame	June 2014	September 2020
Total participants	Elephrame	June 2014	September 2020
Participants per event	Elephrame	June 2014	September 2020
Number of events	Elephrame	June 2014	September 2020
Use of deadly force	Fatal Encounters	January 2000	January 2022
Notable deaths	Self-collected	July 2014	May 2020
<b>COVID-19</b>			
COVID deaths (per 1000)	NYTimes data SSEs Database SafeGraph data	January 2020 January 2020 March 2020	May 2020 May 2020
SSEs event	CDC	January 2020	May 2020
Spring breakers (Visits per device)	OxCGRT	January 2020	May 2020
Black death burden			
Stringency			
<b>Internet and Social Media</b>			
New Twitter accounts	Twitter API	5 weeks before George Floyd's death	
Predicted new Twitter accounts	Twitter API	5 weeks before George Floyd's death	
Google searches for Twitter	Google Trends	5 weeks before George Floyd's death	
Google BLM search	Google Trends	5 weeks before George Floyd's death	
<b>Demography, Income, Unemployment and Votes</b>			
Urban	2013 NCHS Urban-Rural Classification	2013	
Rural areas	2013 NCHS Urban-Rural Classification	2013	
Suburban areas	2013 NCHS Urban-Rural Classification	2013	
Not large cities	2013 NCHS Urban-Rural Classification	2013	
Black population	American Community Survey	2018	
Non-black population share	American Community Survey	2018	
White population share	American Community Survey	2018	
Non-white population share	American Community Survey	2018	
Black poverty	American Community Survey	2018	
Median hh income	U.S. Census	2019	
3+ risk factors	Local Area Unemployment Statistics	January 2020	May 2020
Unemployment	American Community Survey	2012, 2016	
Republican vote shares	? (2014)	2014	
Social capital	Google mobility report	January 2020	May 2020
Residential stay			

Table 3.D.2: Search terms used in indices of search activity

<b>Keywords</b>	<b>Start of period</b>	<b>End of period</b>	<b>Duration</b>
twitter	2020-01-01	2020-05-25	6 months
twitter	2020-04-20	2020-05-25	5 weeks
blm	2020-05-25	2020-06-15	3 weeks
floyd	2020-05-25	2020-06-15	3 weeks
george floyd	2020-05-25	2020-06-15	3 weeks
blm + black lives matter + floyd + george floyd	2020-05-25	2020-06-15	3 weeks
blm + black lives matter + floyd + george floyd	2020-04-20	2020-05-25	5 weeks

Note: The Google Trends data is generated on a designated market area (DMA) level. Keywords are case-independent. The resulting outcomes are normalised measures generated by Google Trends.

Table 3.D.3: Summary statistics - counties with and without geo-localized tweets

	All counties						Counties with tweets						Counties without tweets							
	N		Mean		SD		Min		Max		N		Mean		SD		Min		Max	
<b>From 25th of May to 14th of June 2020:</b>																				
Presence of BLM events	3106	0.298	0.000	1.000	0.000	36.000	3085	0.100	1.000	3085	0.252	1.352	0.000	1.000	0.000	0.000	0.000	0.000	0.000	
Number of BLM events	3106	0.250	1.348	0.000	0.000	323687.500	3085	272.602	5088.765	3085	539.141	878.429	0.000	323687.500	21	0.000	0.000	0.000	0.000	
Participants in BLM events	3106	270.759	5968.521	0.000	0.000	8991.319	307	825.005	7211.615	3085	243596.000	2442.000	0.000	243596.000	21	11.048	27.558	0.000	124.000	
Participants per event	307	539.141	878.429	0.000	0.000	8991.319	307	14.616	53.994	3085	135.658	835.824	0.000	2442.000	21	0.143	0.478	0.000	2.000	
Tweets mentioning BLM	3106	819.502	7187.496	0.000	0.000	2442.000	3106	28943.000	28943.000	3085	4117.000	113.478	0.000	28943.000	21	0.000	0.000	0.000	0.000	
New users tweeting about BLM	3106	4.586	53.812	0.000	0.000	28943.000	3106	4.117	113.478	3085	17.753	113.478	0.000	4117.000	21	0.000	0.000	0.000	0.000	
Tweets mentioning #AIIlivesMatter	3106	134.741	833.066	0.000	0.000	4117.000	3106	0.348	0.477	3085	0.350	0.477	0.000	4117.000	21	0.143	0.359	0.000	1.000	
Neighbor protested first																				
<b>On the 25th of May 2020:</b>																				
COVID deaths (total)	3106	24.461	141.132	0.000	0.000	3304.000	3106	24.627	141.597	3085	462.751	2446.204	0.000	3304.000	21	0.095	0.301	0.000	1.000	
COVID cases (total)	3106	459.678	2438.202	0.000	0.000	72010.000	3106	2.935	0.113	3085	0.249	0.000	2.935	72010.000	21	8.190	18.878	0.000	66.000	
COVID deaths (per 1000)	3106	0.113	0.248	0.000	0.000	145.513	3106	5.664	2.801	3085	5.676	0.000	145.513	21	1.340	3.365	0.000	0.182		
COVID cases (per 1000)	3106	2.791	14.513	0.000	0.000	3085	3106	9.790	0.000	3085	9.816	0.000	143.000	21	1.238	4.392	0.000	15.273		
Superspread events, 6+ weeks ago, neighboring	3106	3.070	143.000	0.000	0.000	4.104	3106	0.963	0.000	3085	1.350	0.960	0.000	4.104	21	0.682	1.094	0.000	4.104	
Black death burden	3106	1.346	0.963	0.000	0.000	47.220	3106	8.508	89.810	3085	68.502	84.59	47.220	89.810	21	60.007	11.416	49.070	78.700	
Lockdown stringency index																				
<b>Before the 25th of May 2020:</b>																				
Google searches for Twitter	3056	61.285	11.222	17.000	100.000	3037	3106	61.317	11.217	3085	11.217	17.000	100.000	19	53.053	9.132	41.000	65.000		
Residential stay	1348	10.633	3.387	3.600	26.286	1347	3106	10.633	3.388	3.600	26.286	1	11.000	11.000	-	11.000	11.000	11.000		
<b>Later outcomes:</b>																				
Followers of @BMLivesMatter	3106	63.198	495.174	0.000	0.000	20058.000	3106	63.628	496.830	3085	1.550	11.245	0.000	20058.000	21	0.000	0.000	0.000	0.000	
Followers of @BMLivesMatter created during the pandemic	3106	1.540	11.207	0.000	0.000	453.000	3106	26.735	0.000	3085	0.708	26.825	0.000	453.000	21	0.000	0.000	0.000	0.000	
Street art count	3106	0.703	0.000	0.000	0.000	1467.000								1467.000	21					
<b>County characteristics:</b>																				
Black police-related deaths (2014-2019)	3106	0.677	3.207	0.000	0.000	84.000	3106	0.682	3.217	3085	0.682	1.545	0.000	84.000	21	0.048	0.218	0.000	1.000	
Black police-related deaths (2020)	3106	0.047	0.301	0.000	0.000	6.000	3106	1.550	0.708	3085	4.692	0.302	0.047	6.000	21	0.000	0.000	0.000	0.000	
Unemployment rate (year average)	3106	4.691	1.550	0.708	0.000	19.650	3106	0.147	0.875	3085	0.147	0.708	0.000	19.650	21	4.456	2.275	1.642	10.267	
Black population share	3106	0.147	0.000	0.000	0.000	0.875	3106	0.162	0.928	3085	0.144	0.160	0.000	0.875	21	0.058	0.156	0.000	0.602	
Non-white population share	3106	0.144	0.162	0.000	0.000	0.928	3106	0.140	0.000	3085	0.020	0.140	0.000	0.928	21	0.242	0.302	0.001	0.928	
Large cities	3106	0.020	0.140	0.000	0.000	1.000	3106	0.118	0.323	3085	0.119	0.324	0.000	1.000	21	0.000	0.000	0.000	1.000	
Suburban areas	3106	0.118	0.323	0.000	0.000	1.000	3106	0.234	0.423	3085	0.235	0.424	0.000	1.000	21	0.095	0.301	0.000	1.000	
Smaller towns	3106	0.628	0.483	0.000	0.000	1.000	3106	0.617	4.183	3085	0.626	0.484	0.000	1.000	21	0.857	0.359	0.000	1.000	
Rural areas																				
BLM events (2014-2019)	3106	0.281	0.225	0.000	0.000	1.000	3106	5.019	10.685	3085	0.281	0.224	0.000	1.000	21	0.239	0.370	0.000	1.000	
Black poverty rate	3106	25.899	5.019	10.685	48.448	3085	3106	0.156	0.960	3085	0.633	0.156	0.084	10.685	21	5.028	27.900	19.106	39.176	
Population share with 3+ risk factors	3106	0.633	0.156	0.083	0.060	1.048	3106	0.596	0.959	3085	0.596	0.959	0.147	0.959	21	0.678	0.214	0.083	0.960	
Vote share for republicans (2016)	3106	0.536	0.148	0.060	1.048	1.048	3106	1.3277.475	20170.891	129150.343	3085	48823.227	13286.361	20170.891	129150.343	21	0.612	0.203	0.060	0.857
Median household income (2016)	3106	1.384	0.705	0.000	6.887	3085	3106	1.384	0.689	3085	1.384	0.689	0.000	6.887	21	1.464	1.424	0.000	6.055	
Social capital	3106	1216.679	555.825	11.998	6474.706	3085	3106	1218.933	555.758	3085	1218.933	555.758	11.998	6474.706	21	885.570	468.940	324.576	1898.310	
Distance to Minneapolis	3106	0.010	0.116	0.000	3.000	1.474	3106	0.114	0.258	3085	0.115	0.259	0.000	3.000	21	0.000	0.000	0.000	0.000	
Notable Deaths																				
Log(SXSW followers created before March 2017)	3106	0.114	0.258	0.000	0.000	1.6558	3106	0.193	0.350	3085	0.194	0.351	0.000	1.6558	21	0.033	0.151	0.000	0.693	
Log(SXSW followers created during March 2017)																				

Table 3.D.4: Example Tweets

Date	Text
May 29, 2020	While #BlackLivesMatter is raising awareness on Twitter, it shouldn't stop there. While you're inside with your families, talk about racism and discrimination. Especially with older generations who don't use social media and don't see further than the national news's portrayal.
May 30, 2020	This is called UNITY. this is what white america doesn't want. they're afraid of the non racist whites to form partnership and unity with POC bc then they will be out numbered. I stand by my brothers #BlackLivesMatter <a href="https://t.co/EPYE9HKkBN">https://t.co/EPYE9HKkBN</a>
May 31, 2020	Reach out to black friends, peers, and social media connections to LISTEN to them with the understanding that I do not know what their struggles are like as a person that has lived with privilege. #BlackLivesMatter If it weren't for Twitter and social media the videos of George Floyd and Ahmed Arbery would have not been seen and murderers would have walked free. Fact. #BlackLivesMatter
Jun 2, 2020	(3/7) We will also be sharing courses made by the Arist community designed to educate allies. The first example: <a href="https://bit.ly/anti-racism101">https://bit.ly/anti-racism101</a> This 20-day text message course will teach you about systemic racism against Black people and how you can practice anti-racist allyship.#BlackLivesMatter
Jun 4, 2020	I made a decision when I came on twitter to keep it strictly for work. I have other social media for expressing personal and political views. However, given the events of the last week, I feel compelled to say something - so here is my bit #BlackLivesMatter #WhitePrivilege
Jun 6, 2020	#IAmASuburbanMom and Black Lives Matter to me! I just went to a rally in a suburb of Atlanta, and there are a lot of us moms who want racial justice and change!
Jun 7, 2020	White privilege means you CAN walk away from #BlackLivesMatter when you get weary and you go back to your regular routine. Our black and coloured allies don't have that privilege to simply walk away. It's their life. Recognizing our white privilege means refusing to walk away. 3 / 3
Jun 11, 2020	1 / I've been trying to learn more about all the complexities of everything going on lately, and how to be a better ally, better support the #blacklivesmatter movement & simply be an anti-racist. For what it's worth, here's a few things I've found to be especially helpful:
Jun 13, 2020	There was a #BlackLivesMatter car parade in my VERY white, VERY red suburban San Antonio neighborhood today. I was afraid was afraid we'd be the only car. There were 50 of us!!!

Note: Example of tweets containing #BLM or #BlackLivesMatter between May 25th and June 14



# Conclusion

This dissertation aims to understand how racial diversity affects the economy. Each of these chapters is an example of one of potential scenarios: exclusion, conflict, and inclusion. Although the contexts of these chapters are different, together they complement each other and reveal different facets of racial diversity in the economy.

In the first chapter I show that racial discrimination can generate significant racial disparities in economic outcomes: I find that an apartment with a discriminatory ad has 4% lower rent price than an identical, but non-discriminating apartment in the same building. This result complements well-established theoretical insights on how differential treatment can generate racial differentials in the housing market. While there are many channels through which racial differentials can occur, pure discrimination in the market remains important and requires further research.

This paper touches on the uncovered topic of the relationship between overt and subtle forms of discrimination. I analyse unique data from the Moscow rental housing, where landlords do not hide their racial preferences. I show that overt and subtle forms of discrimination are closely related. I find that they coexist in Moscow rental housing market and that their relative prevalence is stable across neighborhoods.

Finally, I borrow theoretical framework from the literature on labor search with discrimination and show how the racial rent differential can occur. I do heterogeneity analysis and find that the racial rent differential is higher in neighborhoods with a lower share of discriminating landlords. I show that this result can coincide with a random search model with discrimination by introducing the stylized version of neighborhood sorting

The second chapter studies the impact of tourism on urban amenities. Exploiting a large decline in international travel during the COVID-19 pandemic, we find that tourism decreases the perceived quality of restaurants among locals. We find suggestive evidence that the negative effect of tourism operates through direct aversion against the presence of tourists, rather than overcrowding or supply-side changes. The effect is concentrated in restaurants where the tourist clientele was from countries that have few social ties with the French population.

This paper contributes to an emerging literature on the effects of tourism on locals' welfare. While the existing literature emphasizes price channels, i.e. tourists driving up prices Allen et al. (2020) and endogenous adjustment of amenities Almagro and DominguezIino (2019), we show that tourism has an additional effect on existing amenities which lowers their experienced quality. While we do not aim to evaluate the overall welfare impact of tourism in this paper, we highlight an additional source of discontent that can be caused by tourism. This adds to the debate preceding the pandemic on limiting tourism inflows in some of the most popular tourist destinations. It remains an open question whether tourism will rebound to its pre-pandemic levels. If it does not, our paper provides a preview how persistently lower inflows may affect locals' quality of life.

The third chapter we shed light on the role of social media in generating mobilization in counties whose characteristics are closer to the median voter and where a larger part of the population is not directly impacted by the movement's grievances. We first document that around half of the protests following George Floyd's murder occur in counties that are hosting a BLM event for the first time. We next show that exposure to the pandemic increased protest behavior and that this effect is driven by those counties hosting a protest for the first time. We then turn to the study of the role of social media in explaining this effect. We first present evidence showing that the pandemic lead to an increase in the time spent on online activities and in the use of social media in all counties, and more so in counties hosting their BLM first event after George Floyd's murder. Then, we show that counties where social media was more widely used at the beginning of the pandemic and counties where a higher number of new Twitter users were created during the pandemic show a higher effect of COVID-19 on their protest behaviour. This differential effect is only present in counties with no prior BLM-related protest activity, which suggests that exposure to social media content related to a protest trigger can increase mobilization in parts of the population that were not yet conscious of the problems faced by the aggrieved minority.

**Institut d'études politiques de Paris**  
**ÉCOLE DOCTORALE DE SCIENCES PO**  
**Programme doctoral en économie**  
**Département d'Économie**  
**Doctorat en sciences économiques**

Essais sur l'économie de la discrimination et de  
la diversité

Vladimir Avetian

Thèse supervisée par  
Sergei Guriev, Professeur des universités d'Économie, Sciences Po  
Paris



# Resumé

Cette thèse se divise en trois parties distinctes ayant comme fil conducteur le sujet de la diversité. Je me concentre sur un type particulier de diversité : dans la race, l'identité, les attitudes et les croyances.

Depuis Becker (1957), la race et l'identité sont devenues des éléments légitimes du raisonnement économique. Dans son travail sur la discrimination, Becker a considéré une situation où des travailleurs de deux races coexistent sur le marché et où certains employeurs ont un "dégoût" pour les travailleurs d'une race. Les travaux précurseurs de Becker peuvent être considérés comme faisant partie d'une question plus large : "Que se passe-t-il lorsque des agents de races ou d'identités différentes opèrent dans la même économie ?". Dans les trois chapitres de cette thèse, j'envisage trois scénarios différents qui peuvent se produire.

Le premier scénario, qui a déjà été mentionné, est la discrimination - c'est-à-dire l'exclusion du marché.

Le second scénario est un conflit - lorsqu'aucun groupe n'est en mesure d'en exclure un autre du marché, mais qu'ils continuent à avoir un comportement de rejet. Un exemple serait la ségrégation des consommateurs (Davis et al. (2019)).

Enfin, le troisième scénario, l'inclusion est également possible lorsque des groupes rejoignent une coalition, ou lorsqu'il y a transmission culturelle. Les chapitres présentés ici doivent être considérés comme des exemples, et non comme des généralisations. Dans l'introduction, je me concentrerai sur la littérature à propos de ces trois cas.

La discrimination raciale est un exemple clé d'exclusion. Une vaste littérature économique a été développée pour examiner la discrimination sur différents marchés: travail, logement, consommation, crédit, scolarité, etc<sup>1</sup>.

Deux types de discrimination sont devenus l'épitomé de la littérature théorique : la discrimination basée sur le goût et la discrimination statistique. La premier est déterminée par les préférences des agents (Becker (1957); Arrow (1972); Black (1995)). Pour la second est différente. Elle ne suggère pas que les agents ont des préjugés. Au contraire, les agents sont

---

<sup>1</sup>Pour des analyses approfondies de la littérature, voir Lang and Lehmann (2012); Bertrand and Duflo (2017)

rationnels et utilisent l'identité de la contrepartie comme un indicateur de sa "performance" dans une situation d'asymétrie d'information. Si le groupe discriminé a une performance inférieure en moyenne, alors la discrimination apparaît comme un choix rationnel. Le modèle classique de discrimination statistique a été proposé par Phelps (1972). Un cadre plus complexe de ce modèle, tel qu'introduit par Tirole (1996), implique une étape préalable dans laquelle l'agent minoritaire peut choisir combien il ou elle veut investir dans le développement de la compétence qui détermine la performance future. Ensuite, la "mauvaise réputation" du groupe supprime l'incitation de l'agent à investir dans cette compétence. Il est important de noter que les deux formes de discrimination - statistique ou fondée sur le goût - correspondent à la définition de la discrimination des Nations unies et sont illégales dans de nombreux pays.

Les cadres de la discrimination fondée sur le goût et de la discrimination statistique n'épuisent pas ou ne représentent pas la multitude de mécanismes potentiels et de cadres institutionnels par lesquels la discrimination peut se produire. Small and Pager (2020) soulignent l'importance d'autres cadres et montrent comment ils peuvent compléter et étendre les approches traditionnelles. Ils mentionnent plusieurs directions. Certaines d'entre elles sont déjà apparues dans la littérature économique.

Premièrement, les gens peuvent faire de la discrimination sans s'en rendre compte, un phénomène qui a été appelé "discrimination implicite" dans Bertrand et al. (2005). Deuxièmement, la discrimination peut être renforcée par la structure organisationnelle, même sans l'intention des membres individuels. Troisièmement, la discrimination passée (parfois inscrite dans la loi) peut avoir une forte influence sur l'inégalité contemporaine. Par exemple, Aaronson et al. (2021) montre que le "redlining" des années 1930 a eu un effet socio-économique à long terme. Quatrièmement, des formes mineures de comportement discriminatoire peuvent avoir des conséquences importantes. Par exemple, un travailleur issu d'une minorité peut être embauché mais traité différemment sur le lieu de travail (il a une charge de travail plus élevée, il est plus étroitement surveillé). Enfin, dans l'ensemble, il faudra également tenir compte d'un ensemble plus large de conséquences, telles que la discrimination vécue et la tension émotionnelle.

Du point de vue de la littérature empirique sur la discrimination, le principal défi est que la discrimination est difficile à observer. Dans de nombreuses communautés, la discrimination est illégale et socialement inacceptable. Par conséquent, afin d'étudier la discrimination, nous devons d'abord apprendre à la détecter. Or, cela n'a pas toujours été le cas. Par exemple, aux États-Unis, avant la loi sur les droits civils de 1964, la discrimination raciale était manifeste et répandue. Les offres d'emploi publiées dans le New York Times contenaient régulièrement des exigences raciales explicites (Darity and Mason (1998)). Les complexes immobiliers informaient publiquement les locataires de la politique "pas de

Noirs". Mais surtout, la discrimination à cette époque n'était pas étudiée avec les outils statistiques disponibles aujourd'hui. Une façon d'identifier la discrimination consiste à comparer les résultats économiques de différents groupes raciaux. Cette approche a donné lieu à une littérature qui estime les écarts raciaux en utilisant la décomposition par régression. Les écarts raciaux sur le marché du logement sont bien documentés, la plupart des études portant sur les États-Unis : Ihlanfeldt and Mayock (2009); Bayer et al. (2017); Yinger (1997); Early et al. (2019). Plus précisément, pour le marché américain du logement locatif, Early et al. (2019) montre que les Noirs paient 0,6 à 2,4 % de plus que les Blancs pour un logement identique dans un quartier identique.

On peut toutefois se demander si ces résultats tiennent lorsque tous les contrôles nécessaires sont inclus. Neal and Johnson (1996) montrent que l'écart salarial racial se réduit ou même disparaît lorsqu'une variable mesurant les compétences cognitives d'un demandeur d'emploi est incluse dans l'équation<sup>2</sup>. Cela a conduit les chercheurs à s'interroger : peut-être que les écarts constatés précédemment dans les études ne sont pas le résultat d'une discrimination, mais reflètent des différences entre les groupes avant leur entrée sur le marché. Selon cette logique, les différences de capital humain avant l'entrée sur le marché peuvent expliquer les disparités raciales en matière de salaires, et les différences de compétences en matière de négociation peuvent expliquer les disparités en matière de logement. En s'appuyant sur la décomposition de la régression, il est difficile de dire dans quelle mesure les différences raciales sont causées par la discrimination. Les études qui peuvent aborder cette question de manière empiriquement rigoureuse sont rares (Fryer et al. (2013)).

Depuis le début des années 2000, un autre volet de la littérature est apparu. Afin de révéler l'existence d'un traitement différentiel, les chercheurs ont commencé à mener des expériences de correspondance. Dans leur travail de référence, Bertrand and Mullainathan (2004) ont envoyé des paires de CV fictifs avec des noms à consonance noire ou blanche à des employeurs de Boston et de Chicago, en randomisant les autres caractéristiques. Cette approche leur a permis d'identifier un traitement différentiel : les candidats dont le nom avait une consonance noire avaient moins de chances d'être rappelés par un employeur potentiel. Les expériences de correspondance ont attiré l'attention des chercheurs. Baert (2018) discute de son efficacité et de ses lacunes. Les expériences de correspondance ont révélé des discriminations sur de nombreux marchés, éliminant certains des angles morts caractéristiques des études précédentes sur la discrimination raciale.

En même temps, les expériences par correspondance n'explorent pas clairement le lien entre la discrimination et les écarts raciaux. Dans le premier chapitre, j'identifie ce lien

---

<sup>2</sup>Neal and Johnson (1996) mesurent les compétences avec l'Armed Forces Qualification Test (AFQT), un test utilisé pour déterminer la qualification pour l'enrôlement dans les forces armées des États-Unis

en m'appuyant sur le contexte unique du marché du logement locatif de Moscou, où les propriétaires pratiquent ouvertement la discrimination : environ 20 % des propriétaires moscovites du marché en ligne *Cian* incluent des exigences raciales dans leurs annonces de location. Je vais résumer brièvement ce chapitre dans l'introduction.

Le deuxième chapitre illustre un autre scénario courant : un conflit entre des consommateurs de groupes différents qui se rencontrent dans le même environnement économique sans discrimination du côté de l'offre.

Dans ce chapitre, qui est basé sur un travail commun avec Stefan Pauly, nous examinons la concurrence intra-urbaine entre les touristes et les résidents pour les équipements urbains.

Comme le souligne Faber and Gaubert (2019), "le tourisme implique l'exportation de services locaux autrement non commercialisés en déplaçant temporairement les consommateurs à travers l'espace, plutôt qu'en expédiant des marchandises". En s'appuyant sur les enseignements de la littérature commerciale, Faber and Gaubert (2019) effectue une analyse structurelle des avantages économiques du tourisme. Lanzara and Minerva (2019) se penchent sur les interactions entre le tourisme et les équipements, et examinent les conséquences en termes de bien-être. L'insatisfaction à l'égard du tourisme a rarement été explorée dans la littérature économique. L'exception rare est Takahashi (2019) qui examine les effets négatifs du tourisme d'un point de vue théorique.

Plusieurs facteurs sont à prendre en compte : les touristes, en tant que consommateurs importés, peuvent avoir des préférences et des attitudes différentes de celles des résidents, ils peuvent exercer une pression supplémentaire sur les infrastructures et les services locaux, et enfin, les résidents peuvent avoir des attitudes négatives envers les touristes. Tous ces aspects sont abordés dans le deuxième chapitre, et un bref résumé est présenté plus loin dans l'introduction.

La littérature sur l'économie urbaine contient d'autres exemples, plus touristiques, de conflits entre différents groupes. Dans de nombreuses villes, différents groupes raciaux coexistent, interagissent et consomment dans le même environnement. Mazzolari and Neumark (2012) observent que la diversité parmi les résidents est corrélée à la diversité de la consommation. Ceci est également cohérent avec les preuves de Schiff (2015) concernant l'attrait de la densité en ville. Parallèlement, on sait qu'il peut y avoir une ségrégation dans la consommation en ville. Davis et al. (2019) examine la ségrégation dans la consommation dans la ville de New York, ajoutant à la notion traditionnelle de ségrégation résidentielle dans la littérature.

Le troisième chapitre, rédigé en collaboration avec Annal'i Casanueva Artís, Sulin Sardoschau et Kritika Saxena, jette un éclairage sur un autre scénario potentiel : l'inclusion. Lié à l'économie politique de la protestation, ce chapitre met en lumière un aspect crucial

de la diversité : la capacité de différents groupes à former une coalition pour apporter un changement politique.

Ce chapitre se distingue également des deux autres car il est lié à la littérature qui examine le rôle de l'information et des médias dans l'économie. Des travaux antérieurs ont montré que les médias sociaux peuvent résoudre le problème de l'action collective et de la coordination pour les individus déjà sympathisants d'une cause politique : Enikolopov et al. (2018); Manacorda and Tesei (2020). En revanche, nous nous concentrons sur le rôle des médias sociaux en tant qu'outil permettant d'élargir la coalition et de mobiliser de nouveaux manifestants.

Les études qui examinent l'impact d'Internet et des nouveaux médias ont tendance à utiliser un changement du côté de l'offre dans les premiers stades de l'adaptation d'Internet ou des médias sociaux : Guriev et al. (2019); Müller and Schwarz (2021); Enikolopov et al. (2018); Manacorda and Tesei (2020). À notre connaissance, nous sommes les premiers à étudier le rôle des médias sociaux dans l'élargissement des coalitions politiques par la persuasion, plutôt que par la mobilisation d'individus déjà sensibles aux doléances du mouvement.

Un autre thème qui unit ces chapitres est celui de l'économie numérique. Tous les chapitres bénéficient des nouvelles données issues des plateformes numériques. La consommation, le logement, le transport se sont déplacés en ligne (Goldfarb and Tucker (2019)). Les informations politiques et socialement pertinentes se répandent à travers les médias sociaux. Cela crée une empreinte numérique qui peut être utilisée par les chercheurs. Les économistes du passé accordaient moins d'attention à des questions telles que l'inégalité, non pas parce que ces questions ne présentaient pas d'intérêt social. Au contraire, elles ont toujours été d'un intérêt primordial, mais les données étaient difficiles à obtenir.

Dans les parties suivantes de cette introduction, je résumerai les principaux résultats de chacun des chapitres de la thèse.

## **Chapitre 1 : Considérons les Slaves : Discrimination ouverte et disparités raciales dans les logements locatifs**

Aujourd'hui, la discrimination est le plus souvent subtile. Son impact est donc difficile à mesurer. Ce chapitre tente de surmonter ce défi en s'appuyant sur le contexte unique du marché du logement locatif de Moscou, où les propriétaires font preuve de discrimination. Moscou, où les propriétaires pratiquent ouvertement la discrimination. Ils incluent des exigences raciales dans les annonces, en utilisant des phrases telles que "l'offre est réservée aux locataires slaves", où "slave" désigne les locataires d'origine russe ou les locataires d'origine

russe. russes ou des locataires d'apparence ethniquement russe.

Plus précisément, j'étudie comment la discrimination sur le marché du logement locatif peut générer un différentiel de loyer racial.

Je collecte de nouvelles données sur les annonces provenant de la principale place de marché immobilière en ligne russe, cian.ru. L'ensemble de données comprend toutes les annonces disponibles sur une période d'environ six mois. Je classe les annonces en fonction de la présence d'exigences raciales et je les combine avec d'autres caractéristiques observables des appartements et des quartiers. Environ 20 % des annonces comportent des exigences raciales. Ce paramètre me permet donc d'estimer l'effet de la discrimination sur l'écart de loyer racial. Pour identifier cet effet de manière causale, j'inclus des effets fixes au niveau de l'immeuble dans le modèle afin d'absorber toute caractéristique géographique et au niveau de l'immeuble.

Je constate que la discrimination génère un écart de loyer racial significatif et important : si l'on compare des appartements du même immeuble présentant des caractéristiques observables identiques, les appartements non discriminatoires affichent un prix supérieur de 4 %. Cet article examine également la relation entre les formes manifestes et subtiles de discrimination. Je réalise des expériences classiques de correspondance, en envoyant des messages avec des noms à consonance russe et non russe à un sous-ensemble aléatoire d'annonces en ligne. Cette expérience me permet d'établir un lien entre les résultats obtenus à partir de l'étude d'observation et l'ensemble des preuves existantes dans la littérature expérimentale. Je constate que des formes subtiles et manifestes de discrimination coexistent sur le marché du logement locatif à Moscou. Leur prévalence relative est constante d'un quartier à l'autre.

Enfin, j'emprunte un cadre théorique à la littérature sur la recherche de travail avec discrimination ? et l'applique au contexte des logements locatifs à Moscou. Je démontre que le modèle basé sur la recherche peut expliquer l'existence du différentiel de loyer racial. L'intuition est la suivante : lorsque la recherche est coûteuse et que les minorités ont plus de chances d'être rejetées, elles sont plus susceptibles que la majorité d'accepter une offre défavorable. Alors, les propriétaires non discriminants qui l'anticipent augmenteront le prix du loyer à l'équilibre.

Cependant, le modèle standard basé sur la recherche ne peut pas expliquer les résultats de l'analyse de l'hétérogénéité. Je constate que dans les quartiers (et les immeubles) où la proportion d'appartements discriminants est plus élevée, le différentiel de loyer racial est plus faible. À première vue, cela contredit l'implication du modèle, qui dit qu'avec une plus grande proportion d'appartements discriminants, l'écart devrait se creuser. Cependant, ce point de vue suppose que les quartiers sont des marchés différents et isolés, alors qu'en fait

les locataires potentiels font un tri (mais pas nécessairement une forte ségrégation) entre les quartiers. J'inclus une étape de choix du quartier dans le modèle basé sur la recherche pour expliquer les résultats obtenus dans l'analyse de l'hétérogénéité.

## Chapitre 2 : Aménagements urbains et tourisme : Les données de Tripadvisor

*Ce chapitre est co-écrit avec Stefan Pauly.*

Dans cet article, nous estimons l'effet du tourisme sur la satisfaction des résidents à l'égard des restaurants et d'autres équipements urbains. Nous utilisons des données sur les critiques de restaurants provenant de Tripadvisor - la plateforme qui regroupe le contenu généré par les utilisateurs sur les restaurants et autres expériences de voyage. Nous construisons des données de panel uniques sur la consommation et les commodités dans la ville. Ces données nous permettent d'atteindre plusieurs objectifs en même temps.

Premièrement, nous les utilisons pour produire une mesure très granulaire du tourisme. La part des non-français parmi l'ensemble des avis sert de proxy proche de la présence des touristes, que nous validons à l'aide de plusieurs autres mesures. L'avantage de cette mesure est qu'elle peut être définie à un niveau très granulaire, celui du restaurant lui-même. En outre, alors que de nombreuses études se concentrent sur l'emplacement où les touristes passent la nuit pour étudier l'impact, la mesure utilisée ici permet d'étudier le lieu où les touristes consomment.

Deuxièmement, les données d'évaluation et les notes données par les locaux peuvent être utilisées comme un indicateur de la satisfaction des locaux quant à l'expérience du restaurant. Plus généralement, il s'agit d'une mesure de la satisfaction à l'égard des aménagements urbains, qui varie dans l'espace et dans le temps. La littérature montre que cet indicateur est significatif : Par exemple, Kuang (2017) constate que les évaluations des restaurants sont fortement corrélées aux prix de l'immobilier.

Nous associons les données sur les restaurants à une autre source d'information sur la qualité de vie des résidents : le nombre de plaintes sur la plateforme de crowd-sourced DansMaRue. Cette plateforme est fournie par la mairie de Paris. Les utilisateurs peuvent signaler tout problème lié à l'espace public (déchets abandonnés, tags, affichage sauvage, etc.) via l'application mobile ou le site web. L'administration municipale analyse ensuite les rapports et tente de résoudre les problèmes. Nous considérons cette mesure de désamour comme un autre résultat pertinent pour notre étude.

Nous documentons d'abord deux faits stylisés. Premièrement, les restaurants les plus

touristiques sont moins bien notés par les habitants, ce qui suggère une éventuelle nuisance liée à la demande des touristes. Deuxièmement, les quartiers touristiques présentent une plus faible variété d'équipements, ce qui peut indiquer que les touristes accordent moins d'importance à la variété que les habitants. En utilisant la pandémie comme source de variation exogène des arrivées de touristes internationaux, nous constatons que la baisse du tourisme a entraîné une augmentation de la satisfaction des résidents à l'égard des équipements urbains, à la fois en termes d'évaluation des restaurants et de diminution du nombre de plaintes sur DansMaRue. En particulier, le restaurant moyen augmente sa note de près de 10 % d'un écart-type en l'absence de touristes. d'écart-type en l'absence de touristes et le nombre de plaintes dans le voisinage direct du restaurant moyen diminue d'au moins 8 %.

Il est important de noter que notre effet n'est pas unique au déclin du tourisme induit par le lockdown. Nous trouvons des preuves similaires en utilisant les attaques terroristes qui ont eu lieu en novembre 2015. Nos résultats sur résultats sont également robustes à l'utilisation de mesures du tourisme basées sur la localisation auto-déclarée des utilisateurs plutôt que sur la langue.

Ensuite, nous examinons trois mécanismes potentiels à l'origine de nos résultats : la surpopulation, les changements du côté de l'offre et l'aversion des résidents pour le tourisme. Notre analyse ne trouve que du soutien pour le mécanisme d'aversion. Tout d'abord, nous constatons que le nombre d'avis mentionnant explicitement le tourisme (qui sont souvent négatifs) diminue. Deuxièmement, en se basant sur un indicateur des liens sociaux entre les pays, dérivé des données Facebook, nous constatons que les restaurants dont la clientèle a peu de liens avec la France voient leur note augmenter davantage après la fermeture. Cela suggère que les Parisiens sont moins gênés par les touristes provenant de pays avec lesquels ils ont des liens sociaux forts.

## **Chapitre 3 : Devenir viral dans une pandémie : Les médias sociaux et l'allié dans le mouvement Black Lives Matter**

*Ce chapitre est co-écrit avec Annalí Casanueva Artís, Sulin Sardoschau et Kritika Saxena.*

Qu'est-ce qui a conduit à l'élargissement de la coalition du mouvement Black Lives Matter pendant la pandémie ? Nous abordons cette question en deux parties. Tout d'abord, nous établissons un lien de causalité entre l'exposition à COVID-19 et la participation aux manifestations au niveau du comté, en utilisant les événements de super propagation comme source de variation exogène. Nous montrons que l'exposition à COVID-19 est associée à une

augmentation du comportement de protestation, mais uniquement dans les comtés qui n'ont jamais manifesté pour une cause liée au BLM auparavant.

Ensuite, nous développons un nouvel indice de pénétration des médias sociaux au niveau du comté pour montrer que cet effet est dû à une plus grande utilisation des médias sociaux pendant la pandémie mais avant le déclenchement de la protestation. Bien que nous ne puissions pas totalement exclure que d'autres mécanismes soient en jeu, nous montrons des preuves que des explications alternatives telles que *i*) une augmentation de l'importance de l'inégalité raciale induite par la pandémie, *ii*) des coûts d'opportunité inférieurs de la protestation, *iii*) une propension globale plus élevée à protester et *iv*) une protestation dispersée plutôt qu'étendue ne sont pas à l'origine de nos résultats.

Notre identification est basée sur une petite fenêtre entre la fin mars et la mi-avril 2020, pendant laquelle COVID-19 était suffisamment répandu mais la rigueur du verrouillage suffisamment faible pour permettre à ce qu'on appelle les Super Spreader Events (SSE) de se produire. Ces événements sont caractérisés par la présence d'un individu hautement infectieux (un super diffuseur) et ont eu lieu principalement lors de fêtes d'anniversaire, de maisons de retraite ou de prisons. Nous exploitons la variation transversale du nombre d'ESS dans un rayon de 50 kilomètres de la frontière du comté mais pas dans le comté 6 semaines avant le meurtre de George Floyd pour construire notre instrument d'exposition au COVID-19 au niveau du comté. Nous incluons des effets fixes d'état et un vaste ensemble de contrôles au niveau du comté, plus particulièrement le nombre d'événements historiques de BLM entre 2014 et 2019, ainsi que des variables socio-démographiques et des proxies pour le penchant politique et le capital social.

Nous trouvons des preuves solides que l'exposition à COVID-19 a augmenté les protestations de BLM. Nous estimons qu'une augmentation d'un écart-type du nombre de décès liés au COVID-19 dans un comté au moment du meurtre de George Floyd (environ 25 décès pour 100K habitants), augmente de 5% la probabilité qu'un événement BLM se produise dans les trois semaines suivant le meurtre. Notre résultat de base est entièrement déterminé par les comtés sans manifestations antérieures de BLM et l'effet double en taille et est estimé plus précisément pour ce sous-échantillon.

En outre, nous proposons trois stratégies d'identification alternatives et montrons que nos résultats se répliquent. Premièrement, en utilisant les données de mobilité des téléphones mobiles à grande échelle par *SafeGraph*, nous instrumentons l'exposition à la pandémie avec les flux touristiques vers l'une des plus grandes ESS aux États-Unis - le spring break de Floride en mars 2020. Deuxièmement, nous employons une approche de différence dans les différences, pour laquelle nous grattons des informations sur tous les déclencheurs de protestation BLM similaires depuis 2014 afin d'estimer la réponse différentielle à un déclencheur

de protestation avant et après la pandémie. Troisièmement, nous utilisons une approche d'appariement basée sur LASSO, en comparant les comtés ayant des probabilités de protestation similaires avant la pandémie.

Dans une étape suivante, nous étudions diverses sources d'hétérogénéité et montrons que - conformément à l'idée d'un mouvement d'élargissement - nos résultats de base sont portés par des comtés plus blancs, plus aisés et suburbains.

Dans la deuxième partie du document, nous cherchons à savoir si l'utilisation des médias sociaux peut expliquer l'élargissement du mouvement BLM induit par la pandémie. Nous commençons par répéter l'analyse ci-dessus, en utilisant cette fois un nouvel indice de pénétration des médias sociaux comme principale variable de résultat. Nous constatons que la pandémie a un effet positif et significatif sur notre indice de médias sociaux et que cet effet est entièrement dû au sous-échantillon de comtés qui n'ont jamais manifesté auparavant. Par exemple, nous montrons qu'une augmentation d'un écart-type de l'exposition à la pandémie a entraîné un doublement des comptes Twitter parmi les comtés n'ayant jamais manifesté pour le BLM, sans affecter les comtés qui manifestent traditionnellement.

Dans un deuxième temps, nous examinons le rôle de Twitter dans la mobilisation des manifestants de BLM. Tout d'abord, nous faisons interagir la pénétration de base de Twitter (avant la pandémie) avec l'exposition à COVID-19. Nous répondons à la préoccupation selon laquelle nos résultats pourraient capturer des facteurs sous-jacents qui déterminent à la fois la pénétration de Twitter et la participation aux manifestations, en reproduisant l'instrument SXSW pour la pénétration de base de Twitter utilisé par Müller and Schwarz (2020). Nous montrons que les comtés ayant une pénétration de base de Twitter plus élevée réagissent davantage à l'exposition à la pandémie. En outre, nous interagissons l'exposition à la pandémie avec la pénétration contemporaine de Twitter et nous constatons que l'effet de COVID-19 sur la protestation est entièrement déterminé par les comtés ayant une plus grande pénétration de Twitter pendant la pandémie.

Dans la dernière partie de notre article, nous examinons les mécanismes concurrents. Naturellement, la pandémie a affecté un certain nombre de dimensions importantes qui ne se limitent pas à une plus grande utilisation des médias sociaux. Tout d'abord, nous envisageons la possibilité que nos résultats soient dus à une dispersion plutôt qu'à un élargissement des protestations de BLM. Plus précisément, nous vérifions que l'effet n'est pas dû à une substitution de certains lieux à d'autres. Deuxièmement, la pandémie peut avoir augmenté la saillance globale de l'inégalité raciale *avant* le meurtre de George Floyd. Nous testons cette hypothèse en faisant interagir COVID-19 avec un indicateur de la charge disproportionnée des décès sur les Noirs et le nombre de termes de recherche liés à BLM sur Google avant le déclenchement de la manifestation. Troisièmement, nous cherchons à savoir si la pandémie a

diminué le coût d'opportunité des protestations. Nous interagissons COVID-19 avec le taux de chômage au niveau du comté et la rigueur au niveau de l'état. Si les individus choisissent de protester au lieu d'aller travailler ou de s'engager dans des activités sociales, nous devrions constater un effet plus important dans les comtés où le taux de chômage est plus élevé ou les mesures de rigueur plus strictes. Troisièmement, nous examinons l'effet de COVID-19 sur d'autres protestations. Si la pandémie a augmenté l'agitation générale et la propension à protester, nous nous attendrions à ce que cela soit également vrai pour d'autres causes que le BLM. Nous montrons qu'il est peu probable que ces canaux soient à l'origine de nos résultats.

# References

- Aaronson, D., Faber, J., Hartley, D., Mazumder, B., and Sharkey, P. (2021). The long-run effects of the 1930s home “redlining” maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, 86:103622.
- Arrow, K. J. (1972). Models of job discrimination. *Racial discrimination in economic life*, 83.
- Baert, S. (2018). Hiring discrimination: an overview of (almost) all correspondence experiments since 2005. In *Audit studies: Behind the scenes with theory, method, and nuance*, pages 63–77. Springer.
- Bayer, P., Casey, M., Ferreira, F., and McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, 102:91–105.
- Becker, G. S. (1957). *The economics of discrimination: an economic view of racial discrimination*. University of Chicago.
- Bertrand, M., Chugh, D., and Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, 95(2):94–98.
- Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 1:309–393.
- Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.
- Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, 13(2):309–334.
- Darity, W. A. and Mason, P. L. (1998). Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives*, 12(2):63–90.

- Davis, D. R., Dingel, J. I., Monras, J., and Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, 127(4):000–000.
- Early, D. W., Carrillo, P. E., and Olsen, E. O. (2019). Racial rent differences in us housing markets: Evidence from the housing voucher program. *Journal of Regional Science*, 59(4):669–700.
- Enikolopov, R., Petrova, M., and Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics*, 10(1):150–74.
- Faber, B. and Gaubert, C. (2019). Tourism and economic development: Evidence from mexico's coastline. *American Economic Review*, 109(6):2245–93.
- Fryer, R. G., Pager, D., and Spenkuch, J. L. (2013). Racial disparities in job finding and offered wages. *The Journal of Law and Economics*, 56(3):633–689.
- Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.
- Guriev, S., Melnikov, N., and Zhuravskaya, E. (2019). 3g internet and confidence in government. Available at SSRN 3456747.
- Ihlanfeldt, K. and Mayock, T. (2009). Price discrimination in the housing market. *Journal of Urban Economics*, 66(2):125–140.
- Kuang, C. (2017). Does quality matter in local consumption amenities? an empirical investigation with yelp. *Journal of Urban Economics*, 100:1–18.
- Lang, K. and Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4):959–1006.
- Lanzara, G. and Minerva, G. A. (2019). Tourism, amenities, and welfare in an urban setting. *Journal of Regional Science*, 59(3):452–479.
- Manacorda, M. and Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567.
- Mazzolari, F. and Neumark, D. (2012). Immigration and product diversity. *Journal of Population Economics*, 25(3):1107–1137.
- Müller, K. and Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4):2131–2167.

- Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, 104(5):869–895.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661.
- Schiff, N. (2015). Cities and product variety: evidence from restaurants. *Journal of Economic Geography*, 15(6):1085–1123.
- Small, M. L. and Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, 34(2):49–67.
- Takahashi, T. (2019). Battles between residents and tourists: On the welfare effects of growing tourism.
- Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.
- Yinger, J. (1997). Cash in your face: The cost of racial and ethnic discrimination in housing. *Journal of Urban Economics*, 42(3):339–365.