EIEF WORKING PAPER SERIES

EIEF

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 20/16

June 2020

Divided We Stay Home: Social Distancing and Ethnic Diversity

by

Georgy Egorov
(Northwestern University–Kellogg School of Management and NBER)

Ruben Enikolopov
(New Economic School, ICREA-UPF, Barcelona IPEG, and Barcelona GSE)

Alexey Makarin
(EIEF and CEPR)

Maria Petrova
(ICREA-UPF, Barcelona IPEG, Barcelona GSE, and New Economic School)

Divided We Stay Home: Social Distancing and Ethnic Diversity*

Georgy Egorov† **Ruben Enikolopov**‡
Alexey Makarin§ **Maria Petrova**¶

May 2020

Abstract

Voluntary social distancing plays a vital role in containing the spread of the disease during a pandemic. As a public good, it should be more commonplace in more homogeneous and altruistic societies. However, for healthy people, observing social distancing has private benefits, too. If sick individuals are more likely to stay home, healthy ones have fewer incentives to do so, especially if the asymptomatic transmission is perceived to be unlikely. Theoretically, we show that this interplay may lead to a stricter observance of social distancing in more diverse and less altruistic societies. Empirically, we find that, consistent with the model, the reduction in mobility following the first local case of COVID-19 was stronger in Russian cities with higher ethnic fractionalization and cities with higher levels of xenophobia. For identification, we predict the timing of the first case using pre-existing patterns of internal migration to Moscow. Using SafeGraph data on mobility patterns, we confirm that mobility reduction in the United States was also higher in counties with higher ethnic fractionalization. Our findings highlight the importance of strategic incentives of different population groups for the effectiveness of public policy.

Keywords: COVID-19, pandemic, social distancing, self-isolation, quarantine, fractionalization, diversity, altruism, xenophobia, Russia
JEL Classification: D64, D74, I12

---

*We are grateful to Nadezhda Belyakova for excellent research assistance.
†Northwestern University–Kellogg School of Management and NBER, g-egorov@kellogg.northwestern.edu
‡New Economic School, ICREA-UPF, Barcelona IPEG, and Barcelona GSE, renikolopov@nes.ru
§Einaudi Institute for Economics and Finance (EIEF) and CEPR, alexey.makarin@eief.it
¶ICREA-UPF, Barcelona IPEG, Barcelona GSE, and New Economic School, petrova.ma@gmail.com

1 Introduction

Prosocial behavior may become commonplace in society either through government regulation or through voluntary adherence to social norms. Social distancing and self-isolation during a pandemic is one example of such prosocial behavior, as it plays a key role in slowing down the spread of the infection. During the COVID-19 pandemic, governments in almost all affected countries imposed restrictions aimed at promoting social distancing. However, enforcement of these restrictions is very costly, both logistically and politically. Thus, the effectiveness of these measures, to a large extent, depends on voluntary observance of social distancing by the population. Generally, informal social norms are more difficult to sustain in ethnically diverse societies (Alesina and La Ferrara, 2000; Algan et al., 2016; Goette et al., 2006; Miguel and Gugerty, 2005; Putnam, 2007). This paper challenges this conventional wisdom by showing that ethnic diversity increased socially beneficial behavior during the COVID-19 pandemic in Russia and the United States, and proposes a theoretical mechanism to explain these findings.

We start with a simple observation: at least in the beginning of the COVID-19 pandemic, most people considered themselves healthy. This could be because they had not traveled abroad, had not contacted the ones who had the disease, remembered early suggestions that human-to-human transmission was unlikely, or showed no symptoms. For such individuals, the decision to stay home is driven more by the fear of getting infected than by the desire to avoid infecting others. The likelihood of getting infected is higher if sick people cannot be expected to self-isolate, which, in turn, depends on their prosocial considerations. If people are subject to out-group biases and care less about people from other groups, then the sick are less likely to engage in social distancing in more diverse places. This makes people who consider themselves healthy more likely to self-isolate. Since healthy people constitute a majority, at least at the early stages of the pandemic, we expect to see more social distancing in more diverse societies. Generally speaking, in these circumstances, the private benefits of those who consider themselves healthy are aligned with social objectives. In this paper, we formalize this argument and provide the causal evidence on the differential decline of social distancing in ethnic diversity in Russia and the United States.

We develop a model where people can belong to one of two ethnic groups and have one of the following health statuses: they can be sick, healthy, or asymptomatic carriers. Sick people know they are sick, which means they cannot be infected, and the only reason to self-isolate is their concern for other members of the community. Healthy and asymptomatic carriers do not know whether they are infected, and their reasons to self-isolate are twofold. First, they may be healthy, and self-isolation allows them to remain healthy. Second, if they are asymptomatic carriers and believe they can transmit the disease, they may have altruistic reasons to self-isolate. Suppose that asymptomatic transmission is being underestimated or dismissed. Then in a more diverse society, where sick individuals care less about others and are therefore less likely to self-isolate, the decision of healthy and asymptomatic individuals will be driven by private benefits, which will induce them to self-isolate, given that the sick fail to do so. As long as most people are healthy, which is likely the case at the beginning of the pandemic, more ethnically diverse places should exhibit more compliance with self-isolation. In contrast, if asymptomatic transmission is known to be the main risk, then prosocial incentives of the healthy (and asymptomatic carriers) determine overall compliance with social distancing. In this case, smaller tolerance to out-group members, implied by high fractionalization, should decrease social distancing.

Our main empirical hypothesis is that, once the pandemic starts and the threat of getting infected becomes real, people will be more likely to minimize their day-to-day movements in places with higher ethnic diversity. To identify the effect in the Russian data, we rely on a discontinuous jump in the perceived threat of getting infected after the first case of COVID-19 in a given locality is reported.[1] However, the timing of the official reporting of the first local case is potentially endogenous. It may be affected by the quality of the medical system (e.g., its capacity to diagnose) or by the officials' willingness to publicly admit the problem, both of which can have implications for the citizens' decision to observe social distancing.[2] To deal with this potential endogeneity problem, we use the fact that pre-existing internal migration patterns predict travel flows in 2020. Therefore, how soon the virus spreads to different locations can be predicted by internal migration (Valsecchi, 2020; Mikhailova and Valsecchi, 2020). In Russia, the coronavirus spread primarily from Moscow, which methodologically allows us to use two-stage least squares. First, to predict the timing of the first case, we relate it to internal migration flows to Moscow. Next, following the literature on migration in labor economics (e.g., Altonji and Card, 1991; Card, 2001), we use a shift-share instrument for internal migration. In particular, we combine the data on migration from a given region to Moscow during the 1990s with the nationwide domestic out-migration from that region in more recent years (2015–2018) to instrument for more recent migration flows to Moscow. We then predict the timing of the first local case using the instrumented migration flows from a given region to Moscow. Finally, we use the timing of the first predicted coronavirus case in the region in a difference-in-differences framework, comparing people's behavior before and after the predicted discovery of the first case in places with different levels of ethnic diversity. Note that internal migration to other large cities does not significantly explain the timing of the first COVID-19 case in a region, consistent with the disproportionately high penetration of the virus in Moscow.[3]

[1]E.g., Barrios and Hochberg (2020) document a significant increase in COVID-19 Google searches on the day of the announcement and the following day in the U.S.

[2]The virus was also more likely to hit more densely populated and economically developed places first: see, e.g., www.nytimes.com/2020/03/23/nyregion/coronavirus-nyc-crowds-density.html.

[3]Moscow has had more than 50% of all reported cases in Russia, see Figure 4.

We use the data on people's movements in Russia provided by the largest Russian technology company Yandex, which tracks individuals' cell phones that use its mobile apps.[4] We find that people are more likely to engage in social distancing after the first local COVID-19 case report in more ethnically diverse places. Numerically, we find that a one standard deviation increase in ethnic fractionalization can explain 5.7% of mobility reduction after the first case report. This magnitude corresponds to 4.7% of the average weekday-weekend gap in mobility. Importantly, these magnitudes do not change much after controlling for the introduction of mobility restrictions by the government.

To provide additional evidence on the mechanisms behind our findings, we test another prediction of the model, which is that intolerance between different ethnic groups may have an additional distancing effect on top of ethnic diversity. To measure ethnic tensions, we use data on xenophobic online searches and the number of ethnic hate crimes in a city in the recent years. The results confirm that the reduction in mobility after the first reported case is stronger in places with a higher number of xenophobic searches, as well as in places with a higher number of hate crimes, even taking ethnic fractionalization into account. The magnitudes imply that the additional reduction in mobility in places with one standard deviation higher xenophobia accounts for 2.2% of the average mobility reduction after the first report or, alternatively, 1.8% of weekday-weekend gap in an average locality. Similarly, the additional mobility reduction for places with one standard deviation more ethnic hate crimes accounts for 2.8% of the average mobility reduction after the first report or, alternatively, 2.3% of weekday-weekend gap for an average locality. These reductions are on top of the differential decline by ethnic fractionalization documented earlier.

To ensure that our results are not specific to Russia, we further investigate whether similar effects are observed in the United States. Since the epidemics in the U.S. started in several different locations, we lack a similar source of variation in the timing of the spread of the virus. Thus, we rely on a standard difference-in-differences approach and compare the behavior of people before and after the actual discovery of the first case, in places with different levels of ethnic diversity. Using data on mobile devices from SafeGraph at the county level,[5] we show that the reduction in mobility in the U.S. following the report on the first case in the state is indeed stronger in more ethnically fractionalized counties. The magnitudes imply that a one standard deviation increase in ethnic fractionalization is associated with a 0.46 percentage point larger increase in the share of people staying home. Put differently, the difference between the counties with highest and the lowest fractionalization can explain 6.1% of average mobility reduction after the discovery of the

[4] The company offers many diverse products to its customers and claims to be the Russian Google, Amazon, Uber, and Spotify at the same time (https://www.datacenterdynamics.com/en/analysis/cloud-russia/). Its mobile apps include a web browser, a search engine, a map app, a traffic monitoring app, an Uber-type service (Yandex bought the Russian branch of Uber), a mobile payment app, and many others. Its website was the most visited in Russia as of March 2020 (https://www.similarweb.com/top-websites/russian-federation).

[5] https://www.safegraph.com/dashboard/covid19-commerce-patterns

first case or, alternatively, 8.2% of the weekday-weekend gap for an average locality. These findings are entirely consistent with the results that we get in the Russian case.

To put our estimates in a perspective, we produce a back-of-the-envelope calculation of how many lives might have been saved by a stronger social distancing response in communities with higher levels of diversity. For our calculations, we rely on two estimates of the effect of social distancing on the eventual number of COVID-19 deaths: one coming from a mainstream epidemiological model by Walker et al. (2020) and one from the local average treatment effect estimated by Kapoor et al. (2020) based on a rainfall IV strategy. We consider the elasticity produced by Walker et al. (2020) as an upper bound, as they take into account all potential future deaths from the disease that evolves according to their model. In contrast, we consider the elasticities in Kapoor et al. (2020) as the lower bound, as they study a temporary reduction in social distancing on one particular weekend and because they only take into account data available by the time of writing that article. Based on these two studies, we calculate that a one standard deviation increase in ethnic fractionalization is associated with a range from 570 to 22,250 fewer deaths in Russia and from 2,000 to 40,000 fewer deaths in the United States.

Our paper contributes to the literature on the role of voluntary adherence to social norms in establishing order in a society (Ostrom, 1990; Ellickson, 1994). Cooperation based on other-regarding preferences plays a vital role in sustaining informal institutions and social norms, which greatly enhance the possibilities of collective action (Fehr and Gächter, 2000). A vast existing literature suggests the informal social norms are more difficult to maintain in ethnically diverse societies (Alesina and La Ferrara, 2000; Miguel and Gugerty, 2005; Goette et al., 2006; Putnam, 2007; Algan et al., 2016). Our paper shows that, in contrast with this conventional wisdom, voluntary social distancing during the pandemic may be higher in more diverse places, due to the co-existence of both public and private benefits from the prosocial action.

The paper also relates to the literature on the impact of diversity on development outcomes. Ethnic diversity is often found to be detrimental for outcomes such as economic growth (Easterly and Levine, 1997; Alesina and Ferrara, 2005), public good provision (Alesina et al., 1999), and civil conflicts (Montalvo and Reynal-Querol, 2005; Rohner et al., 2013; Arbatli et al., 2020).[6] However, in recent years there was some evidence that diversity can also be beneficial for productivity (Ottaviano and Peri, 2006; Peri, 2012), innovation (Lee, 2015), and economic development (Alesina et al., 2016; Ager and Brueckner, 2018; Montalvo and Reynal-Querol, 2020). Governments often blame ethnic cleavages for preventing them from reaching their policy goals. In this paper, we show that group heterogeneity can help governments reach their policy goal of imposing social distancing through better individual adherence to this behavior.

[6] Inter-group tensions induced by conflict are also found to decrease inter-ethnic team performance (Hjort, 2014) and inter-group trade (Korovkin and Makarin, 2019).

Finally, we also contribute to the emerging literature on the determinants of social distancing and compliance with stay-at-home orders during the COVID-19 pandemic. This literature is mostly based on a difference-in-differences analysis with social distancing as a function of stay-at-home orders and some third variable. For instance, focusing on the United States, several studies independently find that Republican-leaning counties are less compliant with social distancing recommendations and quarantine orders (Andersen, 2020; Allcott et al., 2020; Barrios and Hochberg, 2020; Engle et al., 2020; Painter and Qiu, 2020; Wright et al., 2020). Other factors that were found predictive of lower compliance with social distancing include local infection rates and older population (Engle et al., 2020), poverty (Wright et al., 2020), as well as higher trust in science and higher education levels (Brzezinski et al., 2020). However, given that both stay-at-home measures and coronavirus spread are unlikely to be random, identification remains an important concern. The counterfactual is not observed: due to the unprecedented nature of the crisis, it is not clear what the trends in people's behavior would be in this new, unusual situation that has not been experienced in the last 100 years. On top of that, policies and the spread of the virus could depend on the dynamics of capacity of the healthcare system and testing policies. In contrast to this literature, we rely on the fact that, in Russia, the virus mostly spread from a single location, and we thus improve on the identification by using pre-existing patterns of internal migration as an instrument. GE: removed "variable" - it's not one variable.

The literature also studies the impact of persuasion on people's mobility. For example, Simonov et al. (2020) and Ananyev et al. (2020) independently show that higher Fox News viewership led to a significantly lower propensity to stay at home during the pandemic. Bursztyn et al. (2020) show that, even conditional on viewing Fox News, watching TV hosts that were more concerned about COVID-19 (Tucker Carlson) led to fewer coronavirus cases and deaths. Finally, there is mixed evidence of the effect of social capital and trust on social distancing. While Borgonovi and Andrieu (2020) and Durante et al. (2020) document evidence of a larger drop in social mobility in areas with higher social capital in the U.S. and Italy, respectively, Doganoglu and Ozdenoren (2020) provide cross-country evidence that generalized trust is associated with *less* social distancing. We contribute to these studies by providing evidence that social diversity is an important determinant of voluntary compliance with social distancing norms in a pandemic.

The rest of the paper is organized as follows. Section 2 contains some background information about Russia and its response to COVID-19. Section 3 presents a theoretical model and discusses its implications. Section 4 then presents our empirical strategy, while Section 5 discusses our data. We present our main results in Section 6 and additional results in Section 7. In Section 8, we show that our main results hold in the case of the U.S. as well. We discuss the implications of our results in Section 9. Section 10 concludes.

2 Background

Starting in early January 2020, the world has been hit with one of the biggest pandemics in history—the COVID-19 pandemic. After an initial period when the virus originated and spread mostly in China, the novel coronavirus started to spread through the rest of the world. In the Western world, the pandemic first broke out in Italy in late February-early March 2020 and then spread in the rest of Europe, and them shortly to the United States and Russia. As of May 2020, these two countries have the largest numbers of detected COVID-19 cases.

In Russia, travel restrictions from China were imposed starting January 31, 2020, and the virus did not begin to spread until it was brought from Italy in early March. Moscow became the main epicenter of the pandemic, and other Russian regions typically got the disease from people arriving from Moscow. Despite the preponderance of international news and evidence, Russian citizens were generally skeptical of the coronavirus threat and did not trust the media and the government with the information about the pandemic. Thus, discovery of regional COVID-19 cases played a big role in informing the local population of the reality and severity of the virus.[7]

Although commonly perceived as relatively ethnically homogeneous, Russia is a multi-national country and home to dozens of ethnic minorities. According to the 2010 Census, ethnic minorities comprise 19.1% of the Russian population. Moreover, there is plenty of regional heterogeneity in ethnic composition. For instance, Yaroslavl and Novgorod oblasts are relatively homogeneous and have 96% and 93% of Russians, respectively. At the same time, the Republic of Tatarstan is a highly ethnically heterogeneous with 115 different ethnicities, a Tatar majority (53.2%) but a sizeable Russian minority (39.7%).

3 Theoretical Framework

3.1 Setup

Consider a simple one-period model. The society is a unit continuum of individuals G and it consists of two ethnic groups G_1 (share $g_1 \in \left(0, \frac{1}{2}\right]$) and G_2 (share $g_2 = 1 - g_1$). In the beginning of the game, each individual may be either healthy (subset H), sick (subset S), or an asymptomatic carrier (subset C). These states are mutually exclusive, and the shares of healthy (h), sick (s), and carrier (c) individuals are all positive and sum to 1; furthermore, we assume that health status is independent of ethnicity. We will denote infected people as $I = C \sqcup S$ and individuals that do not

[7] A survey revealed that 60% of Russians trust information about the coronavirus from the doctors they personally know, while only 8% trust the information coming from the Russian Ministry of Health (https://www.rbc.ru/society/17/04/2020/5e998b669a794768d09da79e).

exhibit symptoms as $N = H \sqcup C$. In other words,

$$G = Group_1 \sqcup Group_2 = \overbrace{Healthy \sqcup \underbrace{Carrier}_{}}^{No\ symptoms} \underbrace{\sqcup\, Sick}_{Infected}.$$

Individuals observe whether or not they are sick, i.e., i knows if $i \in S$ or $i \in N$. However, if they do not exhibit symptoms ($i \in N$), they do not know if they are healthy ($i \in H$) or are asymptomatic carriers ($i \in C$). With this information in hand, all individuals make, simultaneously and independently, a binary decision $d_i \in \{0,1\}$, where 1 is interpreted as self-isolation and 0 as refusal to do so (i.e., going out). Self-isolation does not produce any direct costs or benefits. Going out has a direct benefit b_i; we assume that $b_i \sim U[0, W_N]$ if $i \in N$ and $b_i \sim U[0, W_S]$ if $i \in S$ (and it is independent from ethnicity). It might be natural to think that $W_S \leq W_N$, as sick individuals may have a less desire to go out, but nothing substantive changes in the model if we assume $W_S = W_N = W$. The cost of going out depends on one's health status. A healthy person may become infected, and anyone who is infected in the end of the period gets disutility $-L$ (where $L > 0$). An infected person might infect someone else, leading to a psychological cost $M > 0$ per each healthy person infected as long as this person is from the same ethnic group; the cost of infecting an outgroup person is tM, where $t \in [0,1]$ captures tolerance towards individuals from the other ethnic group (i.e., lack of a negative out-group bias).

Consider the following simplified model of interactions during a pandemic. Suppose that all individuals are matched in pairs, and let $m(i)$ denote the match of individual i. Assume that if everyone goes out, then each i would come into close proximity of exactly one other person, their match $m(i)$. If one or both of two matched individuals decide to stay home, there is no transmission of the infection between them, and the same is true if both are healthy or both are infected (regardless if they are carriers or are sick). If one is healthy and the other is infected, the healthy one becomes infected with probability q if the infected person is sick and r if the infected person is a carrier.[8] Naturally, $r > 0$ reflects the possibility of asymptomatic transmission.

When deciding whether to self-isolate or not, individuals do not know who they are matched with, but know the distribution of types. Thus, individuals that show no symptoms ($i \in N$) choose

[8]The probability of getting infected is thus proportional to the mass of infected individuals who go out, weighted by their contagiousness. In practice, this relationship may be more complex. For example, it may be concave because of the possibility of getting infected by multiple individuals, or it might be convex, for example, because close interactions are easier to avoid when few sick people are out. We adopt the simple proportionality assumption for simplicity.

d_i to maximize their expected utility:

$$U_N = -\frac{c}{h+c}L + b_i \mathbf{1}_{d_i=0} - \frac{h}{h+c}\left(q\mathbf{1}_{m(i)\in S} + r\mathbf{1}_{m(i)\in C}\right) L\mathbf{1}_{d_i=d_{m(i)}=0} \\ - \frac{c}{h+c} r\mathbf{1}_{m(i)\in S}\left(M\mathbf{1}_{G(i)=G(m(i))} + tM\mathbf{1}_{G(i)\neq G(m(i))}\right)\mathbf{1}_{d_i=d_{m_i}=0}, \quad (1)$$

while sick individuals ($i \in S$) maximize

$$U_S = -L + b_i \mathbf{1}_{d_i=0} - q\mathbf{1}_{m(i)\in h}\left(M\mathbf{1}_{G(i)=G(m(i))} + tM\mathbf{1}_{G(i)\neq G(m(i))}\right)\mathbf{1}_{d_i=d_{m(i)}=0}. \quad (2)$$

We are interested in Perfect Bayesian equilibria of this game. To focus on the interesting case, we maintain the following assumption.

Assumption 1. *$W_N < qL$ and $W_S > qM$.*

This first part of the condition is satisfied if the disutility of getting infected L is high enough. Specifically, it states that if a healthy person were certain to encounter a sick individual (and thus get infected with probability q), this person would prefer to stay home. The second condition suggests that altruism M is not too high. This condition means that at least some sick individuals (those with b_i sufficiently high) would go out even if they were certain to encounter a healthy individual. If this condition were to fail, altruism would keep all sick individuals at home at least when most people are healthy. This upper boundary on M also happens to be sufficient (though not necessary) to guarantee existence and uniqueness of an equilibrium.

3.2 Analysis

A Perfect Bayesian equilibrium of this game is characterized by four cutoffs, $\beta_{N1}, \beta_{N2} \in [0, W_N]$ and $\beta_{S1}, \beta_{S2} \in [0, W_S]$, such that individual i with health status $j \in \{N, S\}$ from ethnic group G_k, $k \in \{1, 2\}$, self-isolates if $b_i < \beta_{jk}$ and goes out if $b_i > \beta_{jk}$. The following Proposition characterizes the equilibrium.

Proposition 1. *If $W_N > q\frac{h}{h+c}sL$, there is a unique interior equilibrium, in which $0 < \beta_{N1} < \beta_{N2} < W_N$ and $0 < \beta_{S1} < \beta_{S2} < W_S$ (provided that $g_1 < \frac{1}{2}$ and $t < 1$).[9] Otherwise, in the unique equilibrium,*

[9] A closed form solution exists but is too cumbersome. For example, in the extreme case $r = 0$ (no asymptomatic transmission), we would have $\beta_{N1} = \beta_{N2}$ as people without symptoms would not be concerned about infecting anyone else, and the solution would be given by

$$\begin{aligned} \beta_{Nk} &= W_N\left(1 - W_S\frac{(c+h)W_N - qhsL}{(c+h)W_N W_S - q^2h^2sLM(1-2g_1g_2(1-t))}\right); \\ \beta_{Sk} &= W_S\frac{qhM(1-(1-g_k)(1-t))((c+h)W_N - qhsL)}{(c+h)W_N W_S - q^2h^2sLM(1-2g_1g_2(1-t))}. \end{aligned}$$

$\beta_{N1} = \beta_{N2} = W_N$, $\beta_{S1} = \beta_{S2} = 0$*, so all people without symptoms self-isolate and all sick people go out.*

The coefficient $q\frac{h}{c+h}s$ in the first condition is the probability that a person without symptoms will get infected by a sick person if all sick people go out. If this probability is sufficiently low, then at least some people without symptoms will go out (the first person to do so will not be afraid of getting infected by another such person, so the possibility of asymptomatic transmission does not enter this condition). For example, this condition is guaranteed to hold if $s = 0$, i.e., in the beginning of the pandemic. In equilibrium, people from the ethnic minority are less likely to self-isolate, because the person they might infect is likely to be from the majority group, whereas the probability of getting infected is the same for healthy individuals from both ethnic groups.

We now turn to the comparative statics results.

Proposition 2. *Suppose that* $W_N > q\frac{h}{h+c}sL$*, so the equilibrium is interior. Then an increase in the size of the minority group* g_1*, a decrease in altruism* M *or a decrease in tolerance* t *all decrease self-isolation by sick individuals. The effect on overall self-isolation is ambiguous: it is increasing as a result of either of these changes if* $r < q\frac{s}{c}\frac{qhL-W_N}{W_S+q\frac{h}{h+c}sL}$*, and decreases if the converse is true.*

In the light of Assumption 1, the right-hand side of the last condition is positive for h close to 1, i.e., in the beginning of the pandemic. This means that the comparative statics critically depends on the likelihood of asymptomatic transmission. If it is small, then higher fractionalization implies less self-isolation by sick individuals, but more self-isolation overall, because healthy individuals are concerned of getting infected by sick ones who self-isolate less. If the likelihood of asymptomatic transmission is large, then people higher fractionalization also means that people without symptoms have less concern of infecting healthy ones, and thus overall self-isolation may decrease. As h becomes small (e.g., later in the pandemic), the comparative statics becomes driven solely by sick individuals, and fractionalization will imply less self-isolation. The effect of a decrease in altruism or tolerance is similar.

Proposition 2 implies, in particular, that we should expect fractionalization to have a positive effect on self-isolation in the beginning of the pandemic (h close to 1) and in cases where asymptomatic transmission is believed to be impossible or unlikely (r close to 0). Of course, in the extreme, if $h = 1$ (i.e., before the pandemic), there is no self-isolation, and this does not depend on fractionalization or tolerance.

4 Empirical Strategy

Our theory predicts that in places where the likelihood of asymptomatic transmission is perceived low, when the probability of getting infected becomes non-trivial, people engage more in social distancing in places with higher ethnic fractionalization. To test this prediction, we report the results of two different estimation strategies. First, we report the difference-in-differences estimates, comparing cities with higher and lower level of ethnic fractionalization before and after the first reported case of COVID-19 infection in their region. Second, we combine the difference-in-differences approach with two-stage least squares approach, in which the timing of the first reported case is instrumented using pre-existing migration measures.

More specifically, we aim to estimate the following regression specification:

$$SocialDistance_{irt} = \alpha_i + \theta_t + \gamma FirstCase_{rt} + \beta FirstCase_{rt} \times Ethnic_i + \mathbf{X}_{irt}\delta + \varepsilon_{it}. \tag{3}$$

Here, $SocialDistance_{irt}$ is a measure of people's mobility/staying at home in locality i in region r at time t; $FirstCase_{rt}$ is an indicator variable equal to 1 after the first reported case of COVID-19 in region r (first predicted case in the region in case of IV estimation); $Ethnic_i$ is a measure of ethnic fractionalization in locality i; X_{irt} is a vector of controls that includes some interactions of $FirstCase_{rt}$ with the baseline locality characteristics; α_i are the locality fixed effects, which control for any time-invariant locality characteristics, such as population, population density, baseline levels of health, etc.; and θ_t are the day fixed effects which account for country-wide shocks.

In the OLS specifications, we estimate equation (3) using the actual data on the dates of the first case. The identifying assumption is that of parallel trends, i.e., that in the absence of coronavirus, social distancing patterns in places with high and low ethnic diversity would have followed parallel trends. One potential concern with this approach, however, is that the timing of the first case is not fully random. For example, regions that reported their first COVID-19 case later that others could have done that because of lower medical capacity that did not allow them to identify the virus correctly in time, or their testing policies could be different, or their administration could have been more prone to conceal the first cases for longer. To deal with these potential confounds, we predict the timing of the first case in equation (3) in a two-stage least squares framework.

More specifically, we use the fact that social connections between various cities and the place of the original major outbreak (Moscow) could affect the timing of the first case in their respective regions. We rely on internal migration as a proxy for these type of connections (Valsecchi, 2020; Mikhailova and Valsecchi, 2020). We then estimate the following regression specification for the timing of the first case at the regional level:[10]

$$FirstCase_r = \alpha_0 + \alpha_1 MigToMoscow_r + \eta_r. \tag{4}$$

Here $MigToMoscow_r$ stands for recent migration flows from region r to Moscow, while $FirstCase_r$

[10] We only have dates of the first case and internal migration flows data at the regional rather than the city level, thus we can only estimate this equation at the regional level.

is the date of the first case in this region.[11]

Next, we predict the timing of the first case from equation (4), create a dummy that is equal to 1 after the date of the predicted first case, and finally plug this variable into the equation (3) to estimate the second stage. Moreover, following migration literature, to consistently estimate equation (4), we create a shift-share instrument for internal cross-regional migration. More specifically, we compute the following term:

$$\frac{EarlyMigrationToMoscow_r}{\sum_i EarlyMigrationToRegion_i} \times RecentTotalMigationFromRegion_r \tag{5}$$

and then use it to predict $MigToMoscow_r$ in equation (4). Since this is not a standard IV procedure, for the second stage estimation, which combines IV with difference in differences, we use the bootstrap method to compute standard errors. The identifying assumption behind this identification is that the migration to Moscow from a particular region during the 1990s, interacted with recent (2015–2018) total outflow of migration for this region and further interacted with ethnic fractionalization in a city, only affects isolation through the timing of the first case interacted with ethnic fractionalization (conditional on city and day fixed effects).

5 Data

5.1 Social Distancing Indicators

As the main measure of people's movements in Russia, we use daily averages of Yandex Isolation Index, compiled based on mobile app data.[12] This index aggregates all the data on people's movements at the city level, available from various Yandex applications. Yandex is the largest telecom company in Russia, and its main website Yandex.ru is the most visited website in Russia and the twelfth most visited website in the world. Yandex applications include Yandex browser, Yandex search engine, Yandex Maps (with traffic monitoring), Yandex Cash for payments, Yandex Taxi for taxi rides (in fact, Yandex bought the Russian part of Uber in 2019), Yandex Weather, etc. This data is similar to Google Mobility Index[13] or the data on mobility in China provided by the largest Chinese search engine in China, Baidu (Xiao, 2020). The index is calibrated for each city to be 0 for the busiest hour of the working day, and 5 for the quietest hour of the night before the coronavirus outbreak. For example, Fig. 1 shows the change in isolation index for the city of

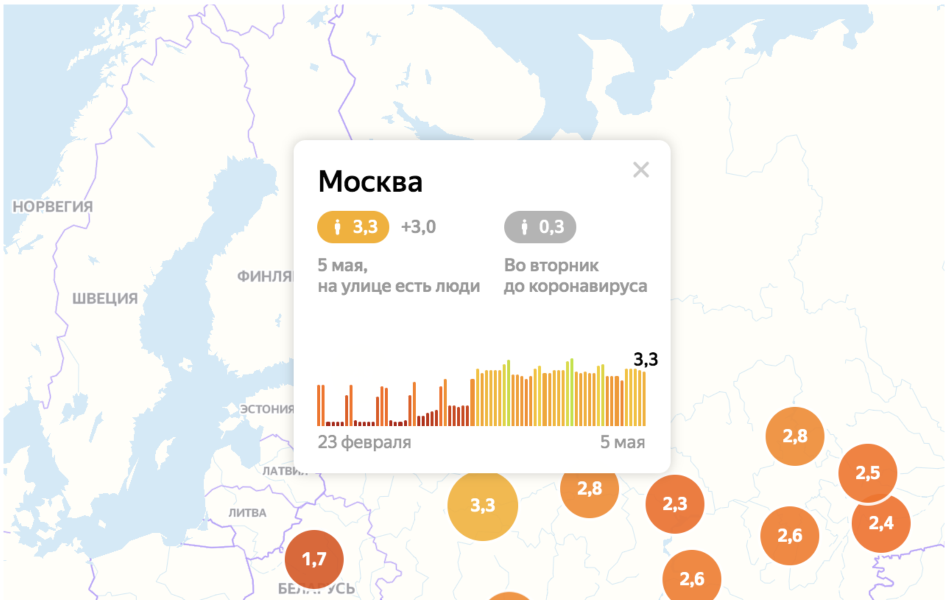
[11] Note that both Moscow and Saint Petersburg have regional status in the Russian administrative division, in contrast to most other cities, which are administratively parts of their region. Thus regional statistics on internal migration includes the data on migration to Moscow.

[12] More specifically, all the data comes from https://yandex.ru/maps/covid19/isolation.

[13] https://www.google.com/covid19/mobility/

Moscow between February 23 and May 5.

Figure 1: Illustration of the Yandex Isolation Index and Its Evolution for the City of Moscow between February 23 and May 5.



Source: Yandex 2020.

We use daily data for all the cities available, e.g. 302 cities with population over 50,000, from February 23, 2020, till April 20, 2020. As one can see, some decline in people's movements began even before the week of March 29 when the first stay-at-home order was issued (see Figure 2). Note that in our subsequent analysis we exclude the data on Moscow and Saint Petersburg from the sample as these are clear outliers in many respects, with Moscow being the place of the largest outbreak in the country (see more on that below).

5.2 Data on COVID-19 Cases

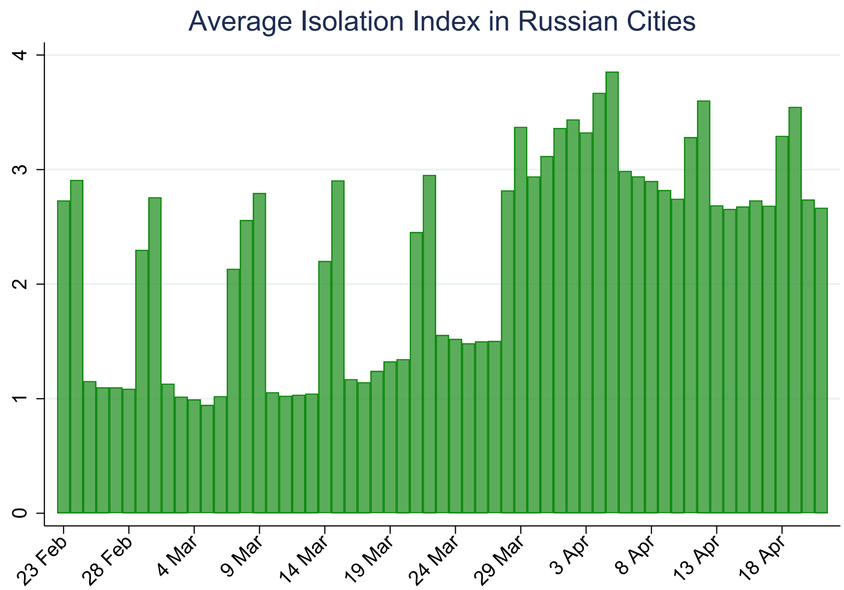
We use the official statistics on the daily number of coronavirus cases by region from the website that contains the official information about the coronavirus and policies enacted by the Russian government to fight it.[14] Figure 3 reports the distribution of the dates of the first case in our data.

Importantly for our identification strategy, even though COVID-19 spread across the country, it started in Moscow (the first case was confirmed in a traveler from Italy on March 1),[15] and it

[14] The source of data is *Rospotrebnadzor*, the government agency responsible for the epidemiological surveillance. As the website does not report historical information, we obtain the actual data from Yandex coronavirus page, which uses this website as a source.

[15] Prior to that, four Russians were diagnosed with coronavirus, three from Diamond Princess cruise ship and one transit passenger flying from Iran to Azerbaijan. In addition, two Chinese citizens were diagnosed with COVID-19 as early as on January 31st, but they were quickly isolated without further documented spread.

Figure 2: Average Isolation Index Across All Russian Cities Over Time.



Source: Yandex 2020.

still accounts for more than half of all cases in Russia. The dynamics of the number of coronavirus cases in Russia and in Moscow is summarized in Figure 4.

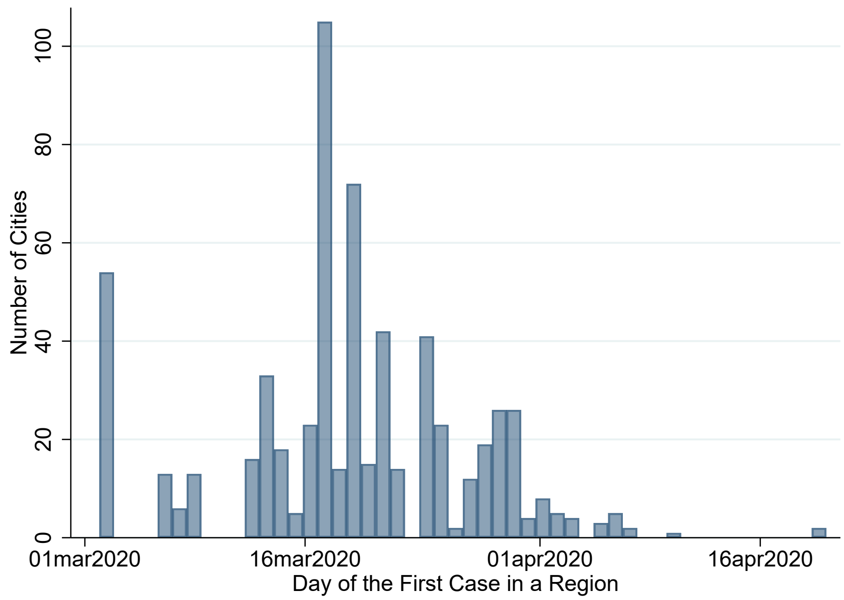
5.3 Other Data

Migration. The data on cross-regional migration comes from the Russian Statistical Agency, *RosStat*. For our empirical exercise, we distinguish between early migration (1990-1997, before the crisis of 1998) and recent migration (2015-2018). Note that in all the years migration to Moscow, as summarized in Figure 5, constituted a much smaller share of overall migration as compared to Moscow's share of coronavirus cases (see Figure 4).[16] That implies that it is unlikely that migration to Moscow accounted for the vast majority of internal migration in Russia, thus our empirical approach, based on shift-share instrument (5) makes sense.

Xenophobia. We use two alternative measures of xenophobia in a city, based on online searches and on the number of hate crimes. The first measure is based on the relative numbers of explicitly xenophobic Internet searches, coming from Yandex WordStat, which is similar to Google Search

[16] Note that RosStat data counts only "official" migration with the change in registration address, so migration to Moscow might be underestimated. However, in any case it is unlikely that migration to Moscow, on average, was even close to its share of coronavirus cases.

Figure 3: Distribution of dates of the first case of coronavirus by region, Moscow excluded.



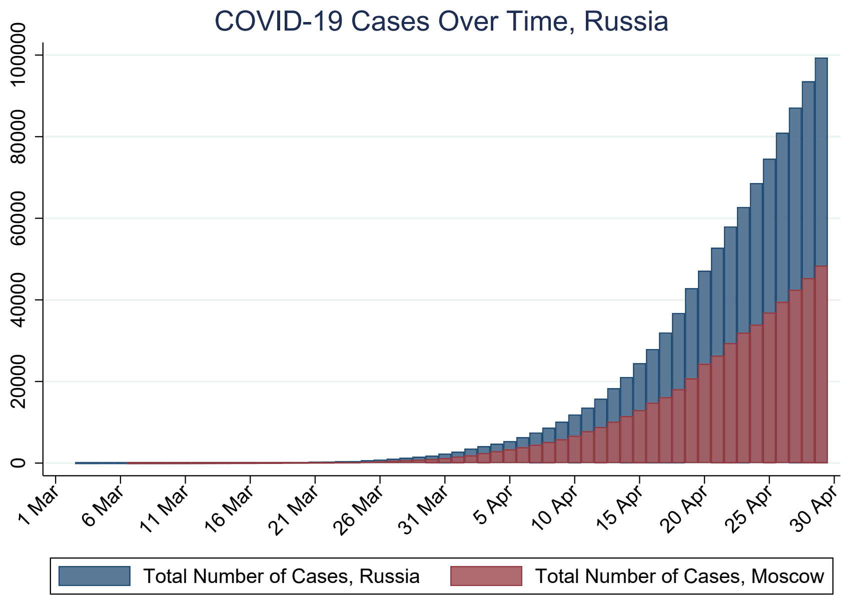
Source: RosPotrebNadzor.

Volume Index (SVI).[17] The data is analogous to the search-based measures of xenophobia or racism are increasingly used in the literature (Stephens-Davidowitz, 2014; Chetty et al., 2019; Ross, 2015). The second measure is based on the city-level data on ethnic hate crime from the database compiled by SOVA Center for Information and Analysis.[18] This is a Moscow-based Russian independent nonprofit organization providing information related to hate crimes that is generally considered to be the most reliable source of information on this issue. The dataset covers incidents of hate crimes and violent acts of vandalism, as well as convictions under any article of the Criminal Code related to "extremism." These data are collected consistently starting 2007, with some incomplete data for 2004–2006. In the analysis we use data from 2007–2015. We classify all hate crimes as "ethnic" or "non-ethnic" based on the type of victim reported in the database. Based on the textual description of each incident in the database we manually coded the number of perpetrators for all the incidents. More details are available in Bursztyn et al. (2019).

[17]There are two main differences between Google SVI and Yandex WordStat. Most importantly, Yandex measure does show the relative numbers of searches per city even if their absolute numbers are small. In fact, Yandex does not have a minimal number of searches for the statistics to be shown, and even a single search is shown. Second, Yandex measure is easily available at the city level, while Google SVI does not report city-level searches for most requests in Russia.

[18]The database can be found at `https://www.sova-center.ru/en/database/`

Figure 4: Number of Cases Over Time.



Source: RosPotrebNadzor.

Other Data The city-level data on population, age, education, and ethnic composition come from the Russian Censuses of 2002 and 2010. The data on the average wage and municipal budgets come from the Russian Federal State Statistics Service (or *RosStat*). Additional city characteristics (latitude, longitude, year the city was founded, and locations of administrative centers) come from the national encyclopedia of Russian cities and regions.[19]

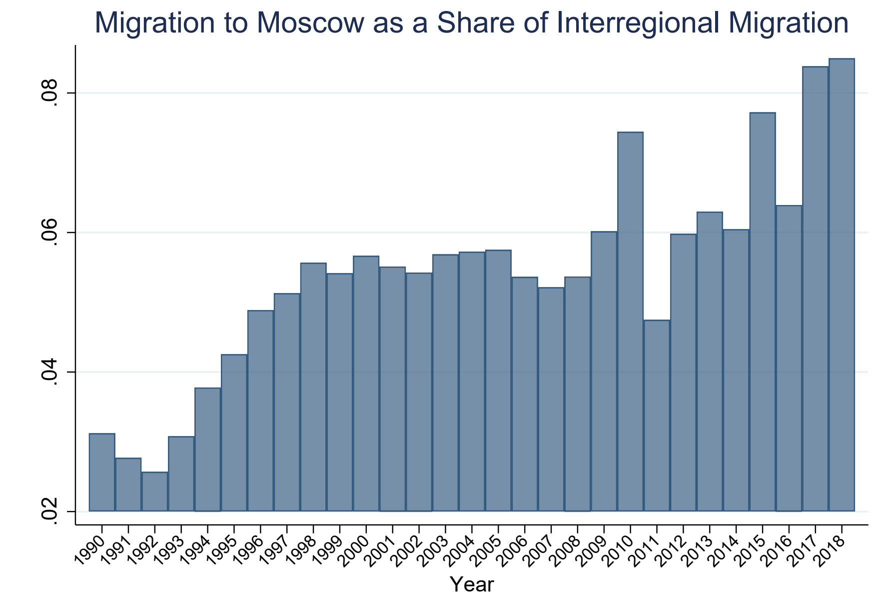
6 Empirical Results

Parallel Trends. Identification in the OLS estimation of equation (3) relies on the parallel trends assumption. That assumption implies that in the absence of COVID-19, the patterns of people's movements around or staying at home would evolve in parallel fashion for places with different levels of ethnic fractionalization. This assumption is not testable, but we can provide some supportive evidence by examining pre-trends. Figure 6 summarizes the patterns of people's movements before and after the first case in a region. It shows the evolution of the isolation index conditional on city and day of the week fixed effects around the day of reporting of the first case of coronaviruis in the region.

As one can see from Figure 6, there is no visible difference in the behavior of people in the two groups of cities before the first coronavirus case. In both groups of cities people engage more in

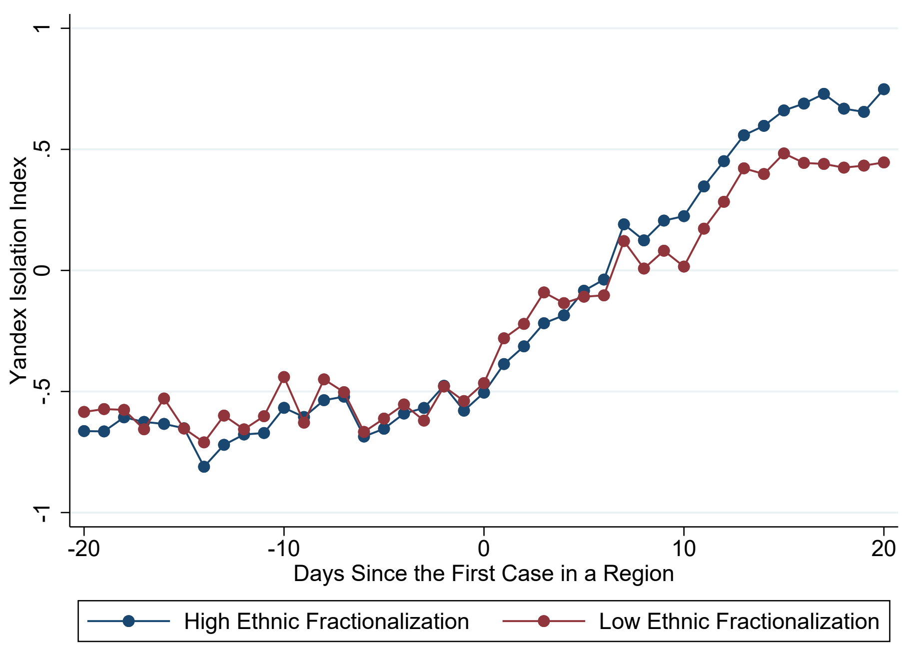
[19] Available at http://www.mojgorod.ru/.

Figure 5: Migration to Moscow over Time.



Source: RosStat.

Figure 6: Isolation over time for places with high and low ethnic fractionalization, Russian data.



Notes: The Yandex isolation index is demeaned by city and day of the week fixed effects. Source: Authors' calculations.

social distancing after the discovery of the first case. However, there seem to be a marked difference in social distancing after the first coronavirus case is reported, with people in more fractionalized cities becoming more likely to stay home. These results are consistent with the parallel trends identifying assumption for (3). This preliminary evidence already points out in favor of our main empirical hypothesis.

Baseline difference-in-differences results. Here, we report the results of estimation of equation (3) using ordinary least squares. Table 1 summarizes these results. Column 1 reports the basic specification with city fixed effects, day of the week fixed effects, and calendar week fixed effects included. Column 2 adds several additional controls on top of that, specifically the interactions of the *Post First Case* dummy with shares of people with higher education, average wage, and population density. Columns 3-4 report the same specifications with day fixed effects included instead of day of the week and calendar week fixed effects. The results indicate that the coefficient for the interaction between the *Post First Case* dummy and ethnic fractionalization is consistently positive and significant in all the specifications. The magnitude of the coefficient goes slightly down from 0.38 to 0.32 with additional interactions, but remains statistically significant at the 1% level. This reduction is smaller than the standard error for both coefficients, and we cannot reject the hypothesis of the equality of the coefficients in a seemingly unrelated regressions framework. Thus, we conclude that the coefficient is robust to inclusion of additional controls. Overall, the results in table 1 are consistent with our theoretical prediction: we indeed observe more social distancing in more ethnically diverse places.

IV estimation. First stage. As discussed above, the OLS estimates from the previous subsection could be biased because of the endogeneity of reporting of the first case in a region, which would lead to the violation of the parallel trends assumption. In what follows, we proceed to estimate equation (3) using the IV approach. We first check whether our logic for the first stage holds, and internal migration to Moscow indeed predicts the timing of the first case in the region. In particular,we estimate equation (4) using OLS and IV, using the shift share instrument (5) to predict migration in the latter case.

The results of these estimations are summarized in Table 2. Columns 1-2 present the results of the OLS estimation, and columns 3-4 present the results of the IV estimation with migration to Moscow being instrumented with the shift-share instrument (5). Columns 1 and 3 present the results without additional controls, while columns 2 and 4 contain the results with basic controls such as population density, income, and education. The results suggest that migration to Moscow has a large negative effect on the timing of the first case. The coefficient is remarkably stable when extra controls are added. IV coefficients are slightly larger than OLS ones, with the magnitudes of

Table 1: Social Distancing, First Case, and Ethnic Fractionalization. OLS.

	Yandex Isolation Index			
VARIABLES	(1)	(2)	(3)	(4)
Post First Case x Ethnic Fractionalization	0.378***	0.318***	0.380***	0.324***
	[0.111]	[0.078]	[0.113]	[0.091]
Post First Case	-0.037	1.233**	-0.095*	0.808
	[0.068]	[0.515]	[0.050]	[0.593]
Post First Case x Education		1.880***		1.818***
		[0.263]		[0.266]
Post First Case x Average Wage		-0.180***		-0.142**
		[0.055]		[0.063]
Post First Case x Population Density		0.003**		0.003**
		[0.001]		[0.001]
City Fixed Effects	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effect	Yes	Yes		
Day Fixed Effects			Yes	Yes
Observations	17,817	17,817	17,817	17,817
R-squared	0.816	0.820	0.944	0.948

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are clustered by region. Isolation index is the aggregate measure of staying at home based on mobile app data. The sample includes 302 Russian cities with population of 50,000 and above. The time period is 23/02/2020–20/04/2020.

coefficients going from -58.93 for OLS to -66.80 for IV. These magnitudes imply that one standard deviation in internal migration to Moscow led to the first case reported 4.6 days earlier according to the OLS estimates, or 5.2 days earlier according to the IV estimates. Another important predictor of the timing of the first reported COVID-19 case is the average wage. According to the estimates, one standard deviation of average wage led to the first case 2.5 days earlier. The results, summarized in Table 2 essentially represent the first stage of the IV estimation.

We also check whether migration to Moscow indeed played a special role in spread of the virus across the country, as compared with that to other big cities. As Figures 4 and 5 suggest, Moscow accounted for the disproportionately large share of all COVID-19 cases, as compared with its share in internal migration (as well as its share in the country's population, which is around 10%). This implies that while other large cities could play a similar role, their actual importance in spreading the virus is likely to have been smaller. In Table 3, we report the results of this estimation. As one can see, the coefficients at migration to the regions with other large cities are smaller in magnitude and flip signs if additional controls are added. Neither the OLS nor the IV coefficients are significant in this estimation, despite the fact that our shift-share instrument still works reasonably well, with the corresponding Kleibergen-Paap statistics around 200 (see columns 3 and 4). Overall, the results of Table 3 confirm the special role of Moscow and regional

Table 2: Timing of First Case and Internal Migration to Moscow in 2015–2018.

	Date of the First Covid-19 case in a Region			
	OLS		IV	
VARIABLES	(1)	(2)	(3)	(4)
Migration to Moscow in 2015-2018	-59.676***	-58.934***	-68.697***	-66.979***
	[11.314]	[9.176]	[5.869]	[5.805]
Average Wage		-6.106**		-5.957**
		[2.451]		[2.416]
Education		15.645		17.042*
		[9.623]		[9.634]
Population Denisty		-0.018		0.011
		[0.043]		[0.047]
Observations	302	302	302	302
R-squared	0.372	0.410	0.364	0.406
Kleibergen-Paap F-statistic			2,032	4,102

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are clustered by region. The sample includes 302 Russian cities with population of 50,000 and above. In columns (3) and (4), migration to Moscow is predicted with a shift-share instrument, using earlier pre-1998 migration to Moscow combined with recent 2015–2018 aggregate outflow of internal migration from a region.

links to Moscow in the spread of the virus, which is consistent with the idea that tighter migration connections to Moscow resulted in regions getting the coronavirus earlier.

IV estimation. Second stage. Once we predict the timing of the first case, as summarized in columns 3 and 4 in Table 2, we can now use these predicted values in the second stage estimation. We report the results of this estimation in Table 4 below. The results of the IV estimation are similar to the results of the OLS estimation, with higher ethnic fractionalization leading to more social distancing post-outbreak. The IV magnitudes are slightly smaller than the OLS ones, but we cannot reject the hypothesis of the equality of the coefficients. The magnitudes in Table 4 imply that one standard deviation increase in ethnic fractionalization leads to 3.1 % higher social distancing following the report of the first local COVID-19 case. In other words, a one standard deviation increase in ethnic fractionalization can explain 4.7 % of the average mobility reduction after the report of the first case or, alternatively, 3.8% of the weekday-weekend gap for an average locality. Overall, the results in Tables 1 and 4 are consistent with the main theoretical prediction that higher ethnic fractionalization increases social distancing once the threat of the virus becomes real.

Table 3: Timing of First Case and Internal Migration to Other Large Cities in 2015–2018.

	Date of the First Covid-19 case in a Region			
	OLS		IV	
VARIABLES	(1)	(2)	(3)	(4)
Migration to Other Large Cities in 2015-2018	0.631	-3.942	3.684	-0.259
	[6.923]	[4.147]	[7.658]	[4.897]
Average Wage		-7.323**		-7.201**
		[3.434]		[3.414]
Education		2.392		5.211
		[10.420]		[10.434]
Population Density		-0.251***		-0.237***
		[0.050]		[0.051]
Observations	302	302	302	302
R-squared	0.000	0.211	-0.009	0.199
Kleibergen-Paap F-statistic			198.5	183.1

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are clustered by region. The sample includes 302 Russian cities with population of 50,000 and above. Migration to Other Large Cities is computed as the aggregate migration to regions with cities with population of at least 1 million people. The list of regions includes Novosibirskaya oblast, Chelyabinskaya oblast, Sverdlovskaya oblast, Tatarstan Republic, Nizhegorodskaya oblast, Samarskaya oblast, Rostovksaya oblast, Bashkortostan Republic, Krasnoyarskyi krai, Permskyi krai, Voronezhskaya oblast, Volgogradskaya oblast, Krasnodarsky krai. In columns (3) and (4), migration to these regions is predicted with a shift-share instrument, using earlier pre-1998 migration to these regions, combined with recent 2015–2018 aggregate outflow of internal migration from a source region.

7 Additional Results and Mechanisms

Xenophobia. Our model suggests that a reduction in tolerance towards out-group members should lead to further increase in self-isolation, even holding the pre-existing levels of ethnic diversity fixed. We test this prediction using two distinct measures of xenophobia in Russian cities, one of which is based on the numbers of explicitly xenophobic Internet searches and the other one is based on the number of ethnic hate crimes in the earlier period. These results of these estimations are summarized in Table 5.

The results indicate that both xenophobia and the history of ethnic hate crime led to an increase in social distancing following the discovery of the first COVID-19 case in the region. Moreover, both of these effects coexist with the positive effect of ethnic fractionalization, without canceling each other. The coefficients for xenophobic searches (Panel A of Table 5) and ethnic hate crime (Panel B of Table Panel A of Table 5) go down substantially when additional interaction terms with control variables are included, but the main coefficient for ethnic fractionalization remains remarkably stable in terms of its magnitude.

Numerically, the estimates in Panel A of Table 5 imply that the difference in mobility between the place with the highest and the lowest level of xenophobia, on top of the impact of ethnic fractionalization, accounts for 2.2% of average mobility reduction following the report of the first case

Table 4: Social Distancing, First Case, and Ethnic Fractionalization. IV.

	Yandex Isolation Index			
VARIABLES	(1)	(2)	(3)	(4)
Post Predicted First Case x Ethnic Fractionalization	0.352***	0.293**	0.345***	0.285**
	[0.109]	[0.122]	[0.107]	[0.117]
Post Predicted First Case	-0.154**	0.893	-0.186***	0.793
	[0.069]	[0.547]	[0.065]	[0.540]
Post Predicted First Case x Education		1.798***		1.813***
		[0.288]		[0.289]
Post Predicted First Case x Average Wage		-0.156***		-0.151***
		[0.059]		[0.058]
Post First Case x Population Density		0.003**		0.003**
		[0.001]		[0.001]
City Fixed Effects	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effects	Yes	Yes		
Day Fixed Effects			Yes	Yes
Observations	17,817	17,817	17,817	17,817
R-squared	0.816	0.820	0.944	0.949

Notes: *** p<0.01, ** p<0.05, * p<0.1. Bootstrapped robust standard errors in brackets are clustered by region. Isolation index is the aggregate measure of staying at home based on mobile app data. The sample includes 302 Russian cities with population of 50,000 and above. The time period is 23/02/2020–20/04/2020. Predicted First Case is computed using the data on inter-regional migration, as summarized in the previous subsection.

or, alternatively, 1.8% of weekday-weekend gap for an average locality. Similarly, the estimates in Panel B of Table 5 suggest that the difference in mobility between the place with the highest level and the lowest levels of ethnic hate crime, on top of the impact of ethnic fractionalization, accounts for 2.8% of average mobility reduction following the report of the first case or, alternatively, 2.3% of weekday-weekend gap for an average locality.

Table 5: Social Distancing, First Case, and Xenophobia.

	Yandex Isolation Index							
	OLS				IV			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Post First Case x Xenophobic Searches	0.051***	0.023**	0.051***	0.021**	0.050***	0.022*	0.051***	0.021*
	[0.011]	[0.010]	[0.011]	[0.010]	[0.012]	[0.012]	[0.012]	[0.012]
Post First Case x Ethnic Fractionalization	0.367***	0.306***	0.366***	0.310***	0.343***	0.283**	0.336***	0.274**
	[0.099]	[0.077]	[0.103]	[0.091]	[0.103]	[0.123]	[0.101]	[0.118]
Post First Case	-0.075	1.333**	-0.132**	0.904	-0.193***	0.985*	-0.226***	0.883
	[0.073]	[0.536]	[0.056]	[0.616]	[0.073]	[0.570]	[0.069]	[0.562]
Observations	17,640	17,640	17,640	17,640	17,640	17,640	17,640	17,640
R-squared	0.817	0.820	0.944	0.948	0.817	0.820	0.945	0.949
Panel B								
Post First Case x Ethnic Hate Crime	0.090***	0.032***	0.090***	0.030**	0.088***	0.032**	0.089***	0.030**
	[0.010]	[0.012]	[0.010]	[0.012]	[0.012]	[0.013]	[0.012]	[0.014]
Post First Case x Ethnic Fractionalization	0.423***	0.340***	0.425***	0.345***	0.397***	0.316***	0.390***	0.306***
	[0.109]	[0.076]	[0.108]	[0.086]	[0.101]	[0.119]	[0.098]	[0.115]
Post First Case	-0.106	1.246**	-0.164***	0.820	-0.221***	0.910	-0.255***	0.809
	[0.069]	[0.523]	[0.051]	[0.597]	[0.072]	[0.558]	[0.068]	[0.553]
Observations	17,817	17,817	17,817	17,817	17,817	17,817	17,817	17,817
R-squared	0.818	0.820	0.946	0.948	0.818	0.820	0.946	0.949
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effects	Yes	Yes			Yes	Yes		
Day Fixed Effects			Yes	Yes			Yes	Yes
Additional conrols		Yes		Yes		Yes		Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are clustered by region. In columns (5)–(8) bootstrapped standard errors are reported. Isolation index is the aggregate measure of staying at home based on mobile app data. The sample includes 302 Russian cities with population of 50,000 and above. The time period is 23/02/2020–20/04/2020. Data on Internet xenophobic searches by city, as captured by Yandex, was collected in 2018. Data on ethnic hate crime by city comes from NGO SOVA (2008–2015). Additional controls include interactions of the dummy for post first case with measures of education attainment, average wage, and population density.

Table 6: Social Distancing, First Case, and Stay-at-Home Orders.

	Yandex Isolation Index			
VARIABLES	(1)	(2)	(3)	(4)
Post First Case x Ethnic Fractionalization	0.281**	0.235**	0.284***	0.240**
	[0.109]	[0.105]	[0.093]	[0.117]
Post First Case	-0.005	1.214**	-0.070	0.803
	[0.070]	[0.472]	[0.055]	[0.562]
Stay at Home Measures x Ethnic Fractionalization	0.033	0.019	0.076	0.064
	[0.191]	[0.190]	[0.145]	[0.140]
Stay at Home Measures	0.362**	0.353**	0.197***	0.187***
	[0.167]	[0.158]	[0.074]	[0.060]
Post First Case x Education		1.822***		1.786***
		[0.256]		[0.260]
Post First Case x Average Wage		-0.174***		-0.139**
		[0.050]		[0.060]
Post First Case x Population Density		0.003**		0.003**
		[0.001]		[0.001]
City Fixed Effects	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effects	Yes	Yes		
Day Fixed Effects			Yes	Yes
Observations	17,817	17,817	17,817	17,817
R-squared	0.819	0.823	0.945	0.949

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are clustered by region. Isolation index is the aggregate measure of staying at home based on mobile app data. The sample includes 302 Russian cities with population of 50,000 and above. The time period is 23/02/2020–20/04/2020.

Stay-at-home orders. The results in Tables 1 and 4 can potentially reflect the fact that following the coronavirus outbreak, many regions introduced stay-at-home orders. If so, ethnic fractionalization could be related to the enforcement of these restrictions, rather then voluntary observance of social distancing as described in our theoretical model.

To test this alternative explanation we explicitly account for the introduction of the restrictive measure by the regional governments. In Table 6, we report what happens if both the dummy for the report of the first case of COVID-19 and the dummy for the introduction of the local stay-at-home orders are included. As one can see, even though the introduction of stay-at-home measures led to a clear increase in social distancing, there is no differential effect of the stay-at-home measures in places with high and low ethnic fractionalization. Unfortunately, we do not have a convincing instrument for the stay-at-home measures, so we report only the results of the OLS estimation.

8 Empirical Results for the United States

To provide evidence that the results reported in the previous section are not specific to Russia, we repeat the analysis using data from the United States. More specifically, we use county level data on mobile devices from SafeGraph to see if the results of estimation of equation (3) are consistent with the ones that we get based on Russian data.

Background The initial spread of COVID-19 in the U.S. occurred almost simultaneously in several states, in particular, California, New York, and Washington. Eventually, New York became the hardest hit state in the U.S. However, due to the multiple initial epicenters, the predictive power of inter-state migration patterns with New York on the initial COVID-19 spread is much lower as compared to the case of Moscow and Russia. For this reason we are not able to use the IV approach in the U.S. setting and have to rely on the OLS estimation only.

The issue of ethnic fractionalization is also highly relevant for the United States, which has one of the most ethnically diverse populations in the world. Typically, however, in the American context, instead of ethnicities, diversity is discussed in terms of races. According to the 2010 Census, 72.4% of the U.S. population are white, 16.3% are Hispanic, 12.6% are African American, and around 4.8% are Asian. Still, states and counties vary drastically in their levels of ethnic (or racial) diversity: for instance, on the one side of the spectrum, 93% of Maine's population is white, while, on the other side of the spectrum, Texas is split 40%–40%–12% among whites, Hispanics, and African Americans.[20] For historical reasons, however, the U.S. population is highly segregated, and ethnic fractionalization correlates with many county-level characteristics, such as population density. For this reason, in our analysis, we do our best to control for various confounders of ethnic diversity.

Data As a measure of social distancing in the United States, we use the social distancing metrics compiled and released by SafeGraph.[21] The data are generated using a panel of GPS pings from anonymous mobile devices. Similar to much of the literature (see, e.g., Chiou and Tucker, 2020; Kapoor et al., 2020), we use the share of devices remaining completely home on a given day in a given county as the dependent variable. For each device, 'home' location is determined by SafeGraph as the common nighttime location of each mobile device over a 6 week period. Since the data are presented at the census block level, we aggregate them up to the county level by taking

[20] https://www.kff.org/other/state-indicator/distribution-by-raceethnicity/

[21] SafeGraph is a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places. To enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group. For details on this particular dataset, see: https://docs.safegraph.com/docs/social-distancing-metrics.

Table 7: Social Distancing, First Case, and Ethnic Fractionalization. U.S. data.

	% Staying Home			
VARIABLES	(1)	(2)	(3)	(4)
Post First Case x Ethnic Fractionalization	3.615***	2.545**	3.599***	2.526**
	[1.280]	[1.198]	[1.295]	[1.225]
Post First Case x Education		0.115***		0.114***
		[0.026]		[0.026]
Post First Case x Median HH Income (in '000s)		0.134***		0.135***
		[0.027]		[0.027]
Post First Case x Population Density		0.220**		0.223**
		[0.091]		[0.092]
County Fixed Effects	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effects	Yes	Yes		
Day Fixed Effects			Yes	Yes
Days	115	115	115	115
Counties	3 138	3 137	3 138	3 137
Observations	360,870	360,755	360,870	360,755
R-squared	0.719	0.742	0.786	0.809

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in brackets are clustered at the state level. Percentage of people staying home is calculated based on the number of mobile devices never leaving house divided by the total number of mobile devices observed in the county that day. The time period is 01/01/2020–24/04/2020. Post first case indicator is equal to one after a county's state already had its first COVID-19 case, and zero otherwise.

a sum of the total number of devices and of the total number of devices remaining completely home. We then calculate the county-level daily share by dividing the latter number by the former.

We use data on COVID-19 cases and deaths over time from the New York Times open repository on coronavirus cases.[22] From this source, we obtain data on the daily total number of cases and deaths in each county and state. We accessed these data on May 5, 2020.

We obtain the data on counties' ethnic compositions from the 2010 Census, based on which we calculate the standard measure of ethnic (racial) fractionalization. For other county-level controls, such as population density, median household income, and the share of adults with a BA degree, we rely on the county-level benchmark indicators from the Social Capital Project.[23] Finally, we obtain data on state-level stay-at-home measures from Raifman et al. (2020).

Empirical Results We report the results of the difference-in-differences exercise, similar to Table 1 for the Russian case. These findings are summarized in Table 7 below.

The results are largely consistent with the Russian case. The magnitudes imply that following the discovery of the first case, the share of those staying at home increased by 1.9 percentage points

[22] `https://github.com/nytimes/covid-19-data`

[23] Available at `https://www.lee.senate.gov/public/index.cfm/scp-index`.

Table 8: Social Distancing, First Case, Stay-at-Home Orders, and Ethnic Fractionalization. U.S. data.

	% Staying Home			
VARIABLES	(1)	(2)	(3)	(4)
Post First Case x Ethnic Fractionalization	2.875***	2.138*	2.810***	2.072
	[0.876]	[1.241]	[0.899]	[1.257]
Post First Case	-1.350***	-9.950***	-1.247***	-9.876***
	[0.305]	[1.040]	[0.336]	[1.059]
Stay at Home Measures x Ethnic Fractionalization	1.630	0.980	1.764	1.110
	[2.000]	[1.928]	[1.949]	[1.881]
Stay at Home Measures	2.091***	1.960***	2.053***	1.918***
	[0.556]	[0.611]	[0.548]	[0.609]
Post First Case x % Education		0.106***		0.105***
		[0.026]		[0.027]
Post First Case x Median HH Income (in '000s)		0.135***		0.136***
		[0.026]		[0.027]
Post First Case x Population Density		0.202**		0.205**
		[0.087]		[0.088]
City Fixed Effects	Yes	Yes	Yes	Yes
Day of the Week and Calendar Week Fixed Effects	Yes	Yes		
Day Fixed Effects			Yes	Yes
Days	115	115	115	115
Counties	3 138	3 137	3 138	3 137
Observations	360,870	360,755	360,870	360,755
R-squared	0.724	0.745	0.791	0.812

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in brackets are clustered at the state level. Percentage of people staying home is calculated based on the number of mobile devices never leaving house divided by the total number of mobile devices observed in the county that day. The time period is 01/01/2020–24/04/2020. Post first case indicator is equal to one after a county's state already had its first COVID-19 case, and zero otherwise.

on average for the most fractionalized county as compared with the least fractionalized county. In other words, the difference between the counties with highest and the lowest fractionalization can explain 6.1% of average mobility reduction after the discovery of the first case or, alternatively, 8.2% of weekday-weekend gap for an average county.

Similarly to Table 6, we report regressions that include interaction terms both with the report of the first case in the state and with the state-level stay-at-home orders. We summarize these results in Table 8. We find that, similar to the Russian case, there is no differential effect of statewide stay-at-home orders on the likelihood of staying at home depending on the level of ethnic fractionalization. At the same time, even in this demanding specification, the coefficient for the interaction between the dummy for the first reported case and ethnic fractionalization remains positive and significant in three out of four specifications.

9 Implications

While the differential reduction in mobility by ethnic fractionalization is important to document on its own, it is also of interest to see the implications of this differential effect for the spread of the disease. To this end, we produce some back-of-the-envelope estimates of how many deaths may have been saved by greater social distancing in more diverse communities. Because the elasticity of deaths with respect to social distancing is unknown at this point, we rely on two estimates—one from a widely cited epidemiological study, and one based on the local average treatment effect estimated in the economic literature.

Elasticity of COVID-19 deaths with respect to social distancing Based on an epidemiological model, Walker et al. (2020) predict that a uniform 45% reduction in interpersonal contact rate within a country would lead to a 50% reduction in mortality rate in Europe and North America, from eight deaths per 1,000 people to four deaths per 1,000 people.

In economics, Kapoor et al. (2020) use variation in rainfall the weekend prior to the official government lockdown to produce the IV estimates of the effect of lower share of home-stayers on cases and deaths from COVID-19. According to (Kapoor et al., 2020, p. 7), "a one percentage point increase in the number of people leaving home on the weekend before the shutdown causes case counts to rise by roughly 13 per 100,000, which translates to roughly one extra death per 100,000."

One may think of the estimates from Walker et al. (2020) as the upper bound, as they assume a permanent reduction in interpersonal contact and take into account the full counterfactual of an exponential growth. In contrast, the latter numbers from Kapoor et al. (2020) need to be viewed as the lower bound, as they study a temporary reduction in social distancing on one particular weekend and because they only take into account the data available at the time of writing of their article.

Russia First, we produce a back-of-the-envelope estimate of the potentially saved lives in Russia. We note that, according to our estimates in Columns 2 and 4 in Table 4, a one standard deviation increase of ethnic fractionalization (0.172) is associated with a $0.29 \times 0.172 \approx 0.05$ increase in the isolation index. We also note that the pre-first-case median in isolation index is 1.4, which means that a 0.05 increase in isolation index is associated with a 3.5% decline in social mobility.

For the upper-bound estimate based on the epidemiological literature, we assume that reduction in mobility is equivalent to reduction in interpersonal contact.[24] Furthermore, we assume that the estimates from Walker et al. (2020) can be applied linearly with the same ratio, i.e., that a 1% reduction in interpersonal contact is always associated with a 1.1% reduction in mortality rates. Then, under these assumptions, one finds that a 3.5% reduction in social contact is associated with

[24] In principle, this need not be the case, since one can move around and still adhere to strict social distancing rules.

a 3.85% reduction in death rates. For Europe and North America, a 3.85% reduction from 4 deaths per 1,000 population is 0.154 fewer deaths per 1,000 population (see Figure 4 in Walker et al., 2020). For Russia, this is equivalent to $0.154 \times 144,500 = 22,250$ fewer deaths.

For the lower-bound estimate, we assume that the a one standard deviation increase in isolation index in Russia is associated with a one standard deviation increase in the share of people staying at home in the US, i.e., that they are both measuring the same underlying factor. Under this assumption, a 0.05 increase in the isolation index as equivalent to $0.05 \times (6.63/0.85) = 0.39$ percentage point increase in the share of people staying home. Under the assumption that the US calculations in Kapoor et al. (2020) are perfectly applicable to Russia, this is equivalent to $0.39 \times 1450 = 565$ fewer COVID-19 deaths (out of around 3,000 at the time of writing this, May 21, 2020).[25]

United States In the United States, according to our estimates in Columns 2 and 4 in Table 7, a one standard deviation increase in ethnic fractionalization (0.252) is associated with a $0.252 \times 2.5 = 0.63$ percentage point increase in share of people staying home.

For the US, we start with the lower-bound estimate as it is straightforward to compute given that the estimates in Kapoor et al. (2020) rely on the same data from SafeGraph and the same variable of the share of people staying home. Under the assumption that the effect observed in Kapoor et al. (2020) is a LATE that is applicable to our "compliers" and that it is stable over time, a 0.63 p.p. increase in the share of people staying home is associated with 0.63 fewer deaths per 100,000 people. In the United States, it is equivalent to $0.63 \times 3,282 = 2,000$ fewer deaths (out of 94,000 at the moment of writing this, May 21, 2020).

For the upper-bound estimate, we rely on the same assumption as earlier. Since the pre-first-case median in share staying home is 22%, a 0.63 percentage point increase in share of people staying home is equivalent to a 2.8% increase in social distancing, or, as we assume, a 2.8% reduction in interpersonal contact. Using the same 1.1 ratio as above, a 2.8% reduction in social contact is associated with a 3% reduction in death rates. For Europe and North America, a 3% reduction from 4 deaths per 1,000 people is 0.12 fewer deaths per 1,000 people. Thus, for the United States, this is equivalent to $0.12 \times 328,200 \approx 39,400$, or roughly 40,000 fewer deaths.

10 Conclusion

This paper highlights the role of ethnic diversity in voluntary adherence to socially beneficial norms, such as self-isolation and social distancing during a pandemic. Using both Russian and

[25] Note that while the upper-bound estimates above take into account all potential future deaths from the disease, these lower-bound estimates are calculated assuming that no deaths would occur starting the day after the Kapoor et al. (2020)'s estimates were produced. This explains the wide range between the two estimates.

U.S. data, we show that people in more diverse places were more likely to restrict their mobility following the reports of the first local COVID-19 cases. While the Russian data allows us to establish a causal relation more cleanly than the data from the U.S. it is reassuring that our results are consistent for both countries. Theoretically, we argue that these results can be explained with a model where sick people self-isolate for altruistic reasons but do so less in more diverse societies due to out-group biases. At the same time, the decision of healthy individuals to self-isolate is determined by private benefits, so they are more likely to self-isolate in more diverse societies, where sick people are less likely to stay at home. As long as the majority considers themselves healthy, the second effect will dominate, and, on average, there will be more voluntary social distancing in more diverse societies. We document that this effect is observed at the beginning of the outbreak when most people believe they are healthy, especially if the threat of asymptomatic transmission is unknown or underestimated.

Our study has important implications for government policy. It highlights that not only the propensity of different groups of people (ethnic or social groups, or healthy as opposed to sick) to engage in prosocial behavior may be different, but that there may be important strategic effects. In the context of the pandemic, decisions of healthy and sick individuals to self-isolate are strategic substitutes. This means, for example, that in a homogeneous society with high levels of tolerance, extensive testing would allow people to learn that they are sick and self-isolate, thus enabling the rest to go out with little fear. In a heterogeneous society with low levels of tolerance, the same policy may allow people who learn that they are contagious to go out more because they have little to lose, with the exact opposite implications for the healthy population.

There are implications for optimal strategies on reopening the economy as well. As long as most people are not sick, we expect our results to hold even after the stay-at-home orders are lifted and the extrinsic motivation to stay at home becomes weaker. Naturally, expectations of voluntary observance of social distancing is likely to be one of the key elements of these strategies. As long as people observe social distancing even in the absence of restrictive government policies, the economy can be restarted even before pharmaceutical or technological solutions for the coronavirus problem are found.[26] These expectations, however, should depend, in particular, on local ethnic diversity, and therefore should reopening strategies. More generally, understanding the effects of government regulations in heterogeneous societies has practical importance beyond the pandemic, which makes it an interesting direction for future research.

[26]See, e.g. the blog post by John Cochrane for the discussion of these considerations https://johnhcochrane.blogspot.com/2020/05/dumb-reopening-might-just-work.html

References

AGER, P. AND M. BRUECKNER (2018): "Immigrants' Genes: Genetic Diversity and Economic Development in the United States," *Economic Inquiry*, 56, 1149–1164.

ALESINA, A., R. BAQIR, AND W. EASTERLY (1999): "Public goods and ethnic divisions," *The Quarterly journal of economics*, 114, 1243–1284.

ALESINA, A. AND E. L. FERRARA (2005): "Ethnic diversity and economic performance," *Journal of economic literature*, 43, 762–800.

ALESINA, A., J. HARNOSS, AND H. RAPOPORT (2016): "Birthplace diversity and economic prosperity," *Journal of Economic Growth*, 21, 101–138.

ALESINA, A. AND E. LA FERRARA (2000): "Participation in Heterogeneous Communities*," *The Quarterly Journal of Economics*, 115, 847–904.

ALGAN, Y., C. HÉMET, AND D. D. LAITIN (2016): "The Social Effects of Ethnic Diversity at the Local Level: A Natural Experiment with Exogenous Residential Allocation," *Journal of Political Economy*, 124, 696–733.

ALLCOTT, H., L. BOXELL, J. CONWAY, M. GENTZKOW, M. THALER, AND D. Y. YANG (2020): "Polarization and Public Health: Partisan Differences in Social Distancing during COVID-19," *Available at SSRN 3570274*.

ALTONJI, J. AND D. CARD (1991): "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in *Immigration, Trade, and the Labor Market*, National Bureau of Economic Research, Inc, 201–234.

ANANYEV, M., M. POYKER, AND Y. TIAN (2020): "The Safest Time to Fly: Pandemic Response in the Era of Fox News," *Working Paper*.

ANDERSEN, M. (2020): "Early evidence on social distancing in response to COVID-19 in the United States," *Available at SSRN 3569368*.

ARBATLI, C. E., Q. H. ASHRAF, O. GALOR, AND M. KLEMP (2020): "Diversity and conflict," *Econometrica*, 88, 727–797.

BARRIOS, J. M. AND Y. V. HOCHBERG (2020): "Risk perception through the lens of politics in the time of the COVID-19 pandemic," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

BORGONOVI, F. AND E. ANDRIEU (2020): "Bowling together by bowling alone: Social capital and Covid-19," *Working Paper*.

BRZEZINSKI, A., G. DEIANA, V. KECHT, AND D. VAN DIJCKE (2020): "The COVID-19 Pandemic: Government vs. Community Action Across the United States," *Working paper*.

BURSZTYN, L., G. EGOROV, R. ENIKOLOPOV, AND M. PETROVA (2019): "Social media and xenophobia: evidence from Russia," *NBER Working Paper*.

BURSZTYN, L., A. RAO, C. ROTH, AND D. YANAGIZAWA-DROTT (2020): "Misinformation During a Pandemic," *Working paper*.

CARD, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.

CHETTY, R., N. HENDREN, M. R. JONES, AND S. R. PORTER (2019): "Race and Economic Opportunity in the United States: an Intergenerational Perspective," *The Quarterly Journal of Economics*, 135, 711–783.

CHIOU, L. AND C. TUCKER (2020): "Social Distancing, Internet Access and Inequality," *NBER Working Paper*.

DOGANOGLU, T. AND E. OZDENOREN (2020): "Should I stay or should I go (out): The role of trust and norms in disease prevention during pandemics," *Working Paper*.

DURANTE, R., L. GUISO, AND G. GULINO (2020): "Civic Capital and Social Distancing: Evidence from Italians' Response to COVID-19," *VoxEU Column*.

EASTERLY, W. AND R. LEVINE (1997): "Africa's growth tragedy: policies and ethnic divisions," *The quarterly journal of economics*, 112, 1203–1250.

ELLICKSON, R. C. (1994): *Order without law*, Harvard University Press.

ENGLE, S., J. STROMME, AND A. ZHOU (2020): "Staying at Home: Mobility Effects of Covid-19," *Available at SSRN*.

FEHR, E. AND S. GÄCHTER (2000): "Fairness and retaliation: The economics of reciprocity," *Journal of economic perspectives*, 14, 159–181.

GOETTE, L., D. HUFFMAN, AND S. MEIER (2006): "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups," *American Economic Review*, 96, 212–216.

HJORT, J. (2014): "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 129, 1899–1946.

KAPOOR, R., H. RHO, K. SANGHA, S. BHAVYAA, S. AJAY, AND G. XU (2020): "God is in the Rain: The Impact of Rainfall-Induced Early Social Distancing on COVID-19 Outbreaks," *Working paper*.

KOROVKIN, V. AND A. MAKARIN (2019): "Conflict and Inter-Group Trade: Evidence from the 2014 Russia-Ukraine Crisis," *Available at SSRN 3397276*.

LEE, N. (2015): "Migrant and ethnic diversity, cities and innovation: Firm effects or city effects?" *Journal of Economic Geography*, 15, 769–796.

MIGUEL, E. AND M. K. GUGERTY (2005): "Ethnic diversity, social sanctions, and public goods in Kenya," *Journal of Public Economics*, 89, 2325 – 2368.

MIKHAILOVA, T. AND M. VALSECCHI (2020): "Internal migration and Covid-19 (in Russian)," in *Economic Policy in Times of Covid-19*, New Economic School, 26–33.

MONTALVO, J. G. AND M. REYNAL-QUEROL (2005): “Ethnic polarization, potential conflict, and civil wars,” *American economic review*, 95, 796–816.

——— (2020): “Ethnic diversity and growth: Revisiting the evidence,” *Review of Economics and Statistics, forthcoming*.

OSTROM, E. (1990): *Governing the commons: The evolution of institutions for collective action*, Cambridge university press.

OTTAVIANO, G. I. AND G. PERI (2006): “The economic value of cultural diversity: evidence from US cities,” *Journal of Economic geography*, 6, 9–44.

PAINTER, M. AND T. QIU (2020): “Political Beliefs affect Compliance with COVID-19 Social Distancing Orders,” *Available at SSRN 3569098*.

PERI, G. (2012): “The effect of immigration on productivity: Evidence from US states,” *Review of Economics and Statistics*, 94, 348–358.

PUTNAM, R. D. (2007): “E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture,” *Scandinavian Political Studies*, 30, 137–174.

RAIFMAN, J., K. NOCKA, D. JONES, J. BOR, S. LIPSON, J. JAY, AND P. CHAN (2020): “COVID-19 US state policy database.” *Available at: www.tinyurl.com/statepolicies*.

ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2013): “War Signals: A Theory of Trade, Trust, and Conflict,” *The Review of Economic Studies*, 80, 1114–1147.

ROSS, C. T. (2015): “A Multi-Level Bayesian Analysis of Racial Bias in Police Shootings at the County-Level in the United States,” *PLoS ONE*, 10(11).

SIMONOV, A., S. SACHER, J.-P. DUBÉ, AND S. BISWAS (2020): “The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the Covid-19 Pandemic,” *Working Paper*.

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.

VALSECCHI, M. (2020): “Internal migration and the Covid-19 virus,” *Working Paper*.

WALKER, P. G., C. WHITTAKER, O. WATSON, M. BAGUELIN, K. AINSLIE, S. BHATIA, S. BHATT, A. BOONYASIRI, O. BOYD, L. CATTARINO, ET AL. (2020): “The global impact of COVID-19 and strategies for mitigation and suppression,” *Imperial College London*.

WRIGHT, A. L., K. SONIN, J. DRISCOLL, AND J. WILSON (2020): “Poverty and Economic Dislocation Reduce Compliance with COVID-19 Shelter-in-Place Protocols,” *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

XIAO, K. (2020): “Saving Lives versus Saving Livelihoods: Can Big Data Technology Solve the Pandemic Dilemma?” *Working paper*.