Regional Science and Urban Economics 108 (2024) 104037

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

The diversity we breath: Community diversity and gas leak management ☆

Felipe Jordán [a,*], Enrico Di Gregorio [b]

[a] *Instituto de Economía e Instituto para el Desarrollo Sustentable, Pontificia Universidad Católica de Chile, Chile*
[b] *Western Hemisphere Department, International Monetary Fund*

ARTICLE INFO

JEL classification:
J15
K32
O18
P18
Q35
Q53

Keywords:
Diversity
Collective action
Public goods
Natural gas leaks
Environment

ABSTRACT

Diversity might reduce the ability of small-scale communities to protect local resources through its adverse impact on collective action and individual attachment. Using geocoded data on more than 1500 Grade-3 gas leaks in 2016 across Boston and Cambridge, MA, we show that when a leak emerges in a narrowly-defined area with higher ethno-racial fractionalization it enjoys lower chances of end-year reparation. After accounting for socio-economic and infrastructural factors, our preferred estimate suggests that moving from the 10th to the 90th percentile of the fractionalization distribution is associated to a 5.7 percentage point decrease in the probability of reparation, compared with an already low average of 3.7%. This result is robust to different definitions of the community surrounding a leak and to the inclusion of competing measures of diversity. Consistent with our framework, social capital appears to be especially lower in communities fragmented across ethno-racial lines. In turn, lower social capital is negatively associated to leak reparation.

1. Introduction

To what extent does community diversity associate with economic and social performance, and how small can a community be for its heterogeneity to matter for its outcomes? Although fragmentation along ethnic, racial and other cultural lines has often been studied in terms of its impact on production and public goods across countries and subnational units (Easterly and Levine, 1997; Alesina et al., 1999), the literature has provided relatively few contributions at a level of disaggregation lower than that of the city (Hjort, 2014; Algan et al., 2016). However, our ability to assess which aspects of public good provision are impacted by diversity and the extent to which we can pin down the relevant mechanisms strongly depend on the level of detail of our analysis.

In this paper, we develop and test a conceptual framework that connects a community's diversity to a crucial aspect of public good provision, namely the capacity to manage and protect local resources. When members of a community as small as a neighborhood or a group of blocks differ substantially by their racial and ethnic background, we posit that their potential to sustain collective action and the likelihood that they will engage in collectively beneficial actions as individuals are reduced compared to more socially cohesive settings. Through these mechanisms, group fractionalization may interfere with the preservation of valuable local resources.

We explore the validity of our conjecture in a relatively understudied setting, that of natural gas leaks emerging along the pipelines that serve end-users in the Greater Boston area. In Massachusetts, more than 20% of the gas distribution infrastructure is made of leak-prone materials, such as cast and wrought iron or bare steel (PHMSA, 2017). This fragility results in *thousands* of gas leaks along the public streets of cities like Boston and Cambridge. In turn, substantial losses of methane raise the risk of fires and explosions, vegetation damage, air quality degradation and higher incidence of respiratory pathologies. As the public utilities tasked with the identification and management of fugitive emissions can behave as de facto natural monopolists in their designated area, they may effectively underperform in the reparation of reported leaks, and shift part of the cost incurred due to lost supply onto their customers (Hausman and Muehlenbachs, 2019). For gas leaks with relatively lower reported grade of priority in terms of people

☆ The authors would like to thank Alberto Alesina, Ryan Enos, as well as Zeyneb Magavi and Audrey Schulman from HEET and Mothers Out Front, for useful comments and support throughout the development of this project. The authors would also like to thank the Editor (Stephen Ross) for his thoughtful feedback and Javiera Muñoz for her outstanding research assistance. The views expressed in this paper are those of the authors and should not be attributed to the IMF, its management, or its Executive Board. Felipe Jordán acknowledges financial support from ANID, Chile, Fondecyt de Iniciación 11240513. A replication package is available at https://github.com/felipejordanc/the-diversity-we-breath.

* Correspondence to: Vicuña Mackenna 4860, Departamento de Economía, Macul, Región Metropolitana, Chile.
E-mail address: felipe.jordan@uc.cl (F. Jordán).

https://doi.org/10.1016/j.regsciurbeco.2024.104037
Received 10 April 2023; Received in revised form 27 May 2024; Accepted 6 June 2024
Available online 25 June 2024

and property's immediate safety, we argue, community-wide aspects related to local engagement in the protection of common resources may play a role in ensuring reparations. While a recent environmental justice literature pointed to a correlation between local demographics such as race and ethnicity and the incidence of gas leaks (Luna and Nicholas, 2022; Weller et al., 2022), we set out to address the role of the coexistence of several groups in a given place.

Our main data leverage georeferenced information on the position and end-year reparation status of almost 1600 Grade-3 leaks along public streets in Boston and Cambridge as reported by public utilities for 2016. We reconstruct the features of the community surrounding each leak by aggregating information on nearby blocks provided by the 2010 Census and the 2011–2015 American Community Survey available in the IPUMS NHGIS online database. Exploiting this novel leak-level data, we establish a negative and significant relation between a neighborhood's ethno-racial fractionalization and the probability that the leak is repaired by year-end, even after controlling for a wide set of socio-economic, geographic, and infrastructure controls. Our ability to flexibly define the size of the community most plausibly affected by a given leak helps us to address the Modifiable Areal Unit Problem inherent in studies dealing with predetermined administrative boundaries. At the same time, we show that fractionalization is a more fundamental predictor of leak reparation than several competing indicators of diversity, including those that reflect the spatial pattern of residential sorting and segregation along ethnic and racial lines. Our preferred estimate implies that, when moving from the 10th to the 90th percentile of the local fractionalization distribution, the probability that a Grade-3 leak is repaired drops by 5.7 percentage points, compared to a considerably low average of 3.7%.

We complete our analysis leveraging a detailed source of information on social capital and communal attachment in the city of Boston, namely the 2008 and 2010 waves of the Boston Neighborhood Survey. Conditional on the features of our leak-level communities, we find that ethno-racial fractionalization correlates negatively and strongly with several survey-based variables capturing local collective efficacy and the extent of social networks, although we find that residents of more fragmented areas report more involvement in local activities. We put these results into perspective by providing a qualitative account of our field visits to The Port in Cambridge and the southern part of South Boston, two areas scoring at the opposite extremes of the leak reparation and diversity distributions. We conclude by showing that higher levels of social capital—but not individual attachment—are positively related to gas leak reparations, suggesting that local diversity might enable better public good provision and resource management by strengthening collective coordination and communal networks.

This paper makes a series of novel contributions to the vast literature focused on (but not limited to) the effects of community diversity. First, we provide evidence that the widely studied negative relation between community diversity and public good provision, in the form of local resource protection and management, may hold at an even smaller scale than that of cities, counties, and regions. We show this by exploiting information collected up to the leak level within a relatively small, but economically and culturally relevant area of the United States—the cities of Cambridge and Boston.

Second, although our results do not allow for a purely causal interpretation, they control to a significant extent for individual and group self-selection into areas with different amenities. We argue in favor of this on two grounds: on the one hand, we expect that the leading drivers of residential choice (such as ease of commute, rent prices, etc.) are largely unaffected by the share of Grade-3 leak reparations, which are low across the board to begin with, as we discuss below; on the other hand, we believe that individuals would not sort *within* a given city neighborhood and across narrowly-defined leak-level areas in a way systematically related to both diversity and the reparations of gas leaks with low safety risk.

Third, we explicitly address the complexity of studying empirical phenomena across space. Besides accounting for spatial correlation in our standard errors, our flexible leak-level strategy reduces the risk that the estimates are entirely driven by the arbitrary definition of the community affected by a given leak. We do so by assessing the robustness of our baseline results to the choice of the size of the community surrounding each leak, and explore the role of alternative concepts of diversity leveraging the geographic nature of the data.

Lastly, we address a problem that is both policy relevant and severely understudied in the traditional economics literature, namely that of faulty gas infrastructure in U.S. urban settings. Methane is generally considered a relatively clean fossil fuel, but the benefit of a transition from more historically damaging sources of energy such as coal and oil may be significantly reduced by inappropriate management along each phase of the natural gas supply chain. We discuss the role of consumer communities in ensuring a smooth transition and resource waste control.

The rest of this paper is organized as follows. Section 2 presents our conceptual framework. Section 3 introduces our working assumptions and research hypotheses based on the problem of natural gas leaks in the U.S. and the Greater Boston area. After describing our data in Section 4, Section 5 studies the relationship between Grade-3 gas leak reparation and community diversity. Next, Section 6 discusses potential mechanisms, for which we provide both a quantitative and qualitative analysis. Section 7 concludes. The Appendix covers all additional details and results.

2. Conceptual background

A large literature in economics and political science has pointed out that diverse communities are often associated with a relatively low provision of public goods, both in the U.S. and elsewhere (Alesina and La Ferrara, 2005). Evidence that public spending on items such as roads and schools (Alesina et al., 1999) or tax collection (Hopkins, 2009) are lower in diverse societies appears to bear out the predictions in Olson (1965) and Ostrom (1990, 2000), by which special challenges arise in the pursuit of collective or collective-minded action. Be it a racial, ethnic, or cultural element that drives the heterogeneity, diverse countries, cities, and neighborhoods alike seem to be at a disadvantage in trying to make or sustain optimal choices.

More recently, a subset of this literature with greater focus on causal identification has both dramatically reduced the level of empirical observation, surveying communities as small as apartment blocks, and suggested that diversity might translate into a local-level inability to protect common resources. In principle, managing an existing resource should require less constructive capacity than that required to advocate and provide for an entirely new public good, such as a modern sewage system or a requalified public park.[1] Miguel and Gugerty (2005) first documented this gap in communal management skills by showing how rural communities in Western Kenya displayed poor water well maintenance when they featured higher fragmentation across historical ethnicity lines. Similarly, Algan et al. (2016) report higher levels of voluntary degradation and fewer repairs of common space items not directly affected by residents (e.g. the central heating system and overall façade conditions) in French public housing blocks with higher birthplace and nationality diversity. In addition, Alesina et al. (2019) show that the levels and changes in ethnic fractionalization across Indonesian districts are powerful predictors of deforestation dynamics since the 2000s.

[1] For example, the protection of an existing resource might entail less political bargaining to strike a balance amid social groups' competing claims. Requesting the provision of a new public good or increasing provision in a given area may involve politically challenging tradeoffs within constrained municipal, state, or federal budgets, exacerbating cross-group conflicts.

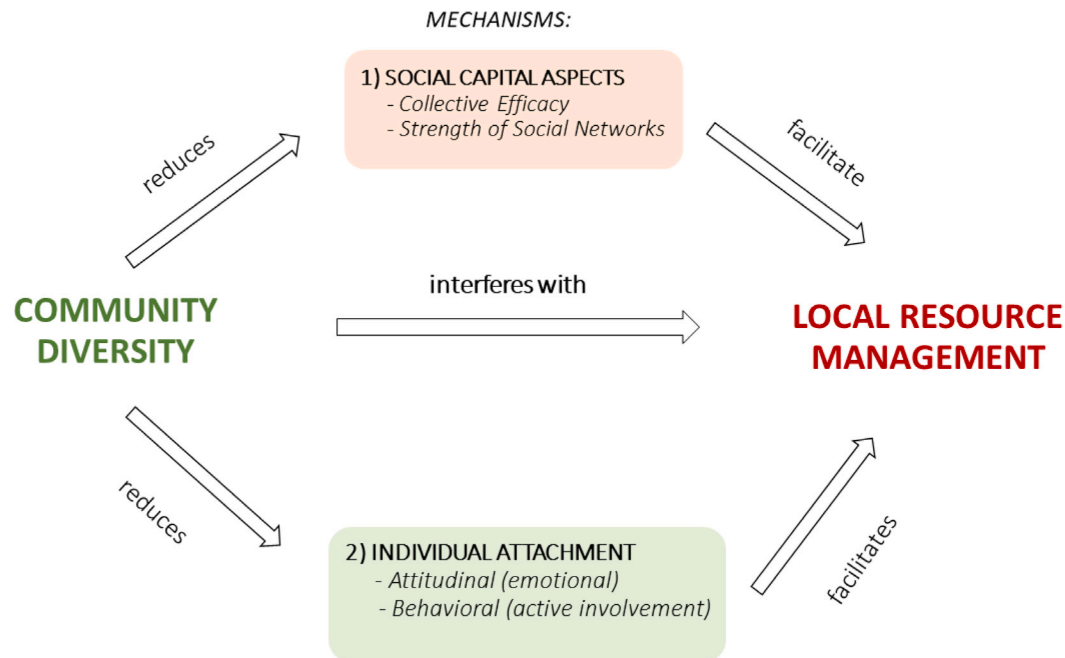

Fig. 1. Conceptual framework: diversity and resource management.

How does the adverse relationship between diversity and local resource management emerge? In the realm of public good provision, the traditional channels involve some combination of heterogeneous preferences across social groups, cooperation-inducing technologies, and strategy-selection processes (Habyarimana et al., 2007). In each case, fragmentation across groups would prevent members of a given community to agree on the optimal outcome to select, or to enforce a set of desirable actions. In turn, such failures would either stem from an inability to punish free-riders and non-cooperators from different groups (including elected officials), or from obstacles in coordinating for collective action.

Fig. 1 extends this picture to the case of local resource management, introducing our conceptual framework. We highlight two mechanisms that could help translate community diversity into a reduced ability to preserve common resources at the local level: first, social capital, which we see as the key to organize and sustain collective action; second, individual attachment to the community, which takes up the role of preference heterogeneity in our setting.[2] As the basis of collective action, social capital can encompass a community's collective efficacy, or its ability to reach common goals through the enforcement of shared norms, and the social networks that constitute the fabric of that community. Individual attachment may play an additional role. Members of a community may be driven to undertake a collectively beneficial action (such as picking up trash from a public street, or protesting against the closing of a local school) not just as the result of peer pressure, but also due to an intrinsic motivation to preserve the quality of their neighborhood. The nature of this attachment can be attitudinal, or the emotional connection between a resident and their community of reference, or behavioral, reflecting of their active involvement in her community's activities (Greif, 2009).

By many accounts, local diversity may reduce the extent of both social capital and community attachment, with these two elements potentially reinforcing each other. Several theories and empirical results point to a negative relationship between local diversity and social cohesion or cooperation (Putnam, 2007; Tabellini, 2008; van der Meer and Tolsma, 2014), including seminal work on the United States documenting how residents of more racially and ancestry fragmented Metropolitan Statistical Areas display lower rates of membership in local organizations (Alesina and La Ferrara, 2000), and tend to trust others less (Alesina and La Ferrara, 2002). In some cases, the conspicuous presence of multiple racial, ethnic, or linguistic groups in one's locality has been found to erode emotional bonds between people and their neighborhoods (Rice and Steele, 2001; Greif, 2009; Górny and Toruńczyk-Ruiz, 2014; Laurence and Bentley, 2016; da Silva and Alvarez, 2019). This evidence is consistent with theories of group relationships such as the "threat hypothesis", whereby increasing social fragmentation would result in stiff competition for limited resources and other perceived drawbacks of cohabitation, including prejudice and hostility, mounting anomie, fear, social exclusion, and distrust. It is reasonable to expect that these obstacles to attachment can also be exacerbated by the lack of a shared ethnicity or racial identity, especially in urban environments where international migration contributes to these dimensions of diversity.[3]

While protecting local resources should be easier in communities with the capacity to engage in collective action and a strong network to leverage in the pursuit of a shared goal, translating attachment into collectively beneficial actions may be more complex—though encouraging evidence of this possibility exists. Wong (2010) explores at length how group closeness and someone's sense of community actually coincide with both preferences and behavior consistent with a willingness to proactively care for their locality. More specifically, Kyle et al. (2010) and Pradhananga and Davenport (2017) report how individuals scoring higher on measures of community attachment are also more likely to join in shared efforts to prevent damage from potential fires as well as to protect local water resources, respectively.

While a community might fail to reach satisfactory public good provision due to internal collective action problems, external forces can play a major role, too. An emerging literature on environmental justice

[2] Heterogeneity in preferences for a given public good can fundamentally affect group decisions about its provision both in absolute terms and relative to competing private goods. We expect preference heterogeneity to matter relatively less for resource protection: a functioning public good or resource should always be preferred to a damaged or wasted one under a relatively weak presumption of rationality.

[3] For a detailed review, see Górny and Toruńczyk-Ruiz (2014).

in the U.S. suggests that People of Color and lower-income residents are exposed to worse gas leak conditions on average (Luna and Nicholas, 2022; Weller et al., 2022). To the extent that historical discrimination and weak regulatory frameworks push non-White and immigrant communities to neighborhoods with faulty infrastructures, the simple correlation between gas leak status and demographic variables lends itself to multiple interpretations. As a result, our analysis will explicitly control for several proxies of local vulnerability such as average income and the share of Blacks, Latinos and households with limited English proficiency.

Though imperfect, our intent is to capture the role of diversity in the interaction of different groups residing in an area rather than as a mere composition of individuals whose features independently affect the outcome. We build upon the reasoning formalized in Vigdor (2002, 2004): after empirically controlling for the shares of all relevant groups in a population, suitable measures of diversity such as those based on the Herfindahl–Hirschman index have the potential to capture the residual cross-group affinity or distaste that might interfere with collective action. We perform our analysis with this goal in mind.

3. Gas leaks in the U.S., Massachusetts, and the Greater Boston area

We set out to test our conceptual framework in a novel setting, namely, that of natural gas leak reparations across Boston and Cambridge, MA. Although a relatively understudied problem, leaks along natural gas infrastructures endanger society's ability to reap the full benefit of a transition from coal and oil to cleaner energy sources (The Economist, 2016). Sizeable losses are reported at every stage of the gas supply chain, with collateral emissions starting at the extraction stage and continuing up to the point of delivery in consumers' buildings. Although estimates of lost and unaccounted for (LAUF) gas in the United States are still uncertain, and public agencies are often criticized for systematically underestimating the size of this phenomenon, independent studies suggest that downstream emissions may average to 2.5% of all methane (natural gas' main component) in the industry (Howarth, 2014).

Urban communities are especially vulnerable to the emergence of gas leaks in the dense infrastructure that runs beneath public streets and into homes and commercial buildings, since outdated pipes made of cast and wrought iron or unprotected steel give way to fissures and cracks from which methane escapes. If unchecked, emissions pose a variety of risks and costs to residents and passersby, from air pollution and vegetation damage to the higher rates that utilities charge to cover for lost gas and pipeline reparations. We survey these costs in Appendix A.

Despite the existence of multiple dimensions of risk, public agencies have been chiefly concerned with classifying leaks by their level of threat to physical safety based on their explosion potential. The U.S. Gas Piping and Technology Committee has provided the most widely adopted standard in terms of three grades[4]:

- *Grade-1* leaks represent an existing or probable hazard to persons or property, and require repair as immediately as possible and continuous action until the conditions are no longer hazardous.
- *Grade-2* leaks are recognized as non-hazardous to persons or property at the time of detection, but justify scheduled repair based on probable future hazard.
- *Grade-3* leaks, which are the focus of this project, are recognized as non-hazardous to persons or property at the time of detection and can be reasonably expected to remain non-hazardous.

Although this classification might suggest that lesser grade leaks do not require special or immediate attention, they still contribute to *all* other environmental, health, and economic risks and costs in Appendix A. Two considerations further complicate this picture. First, utilities' road-level methane readings are not necessarily correlated with the actual CH_4 flux emanating from a problematic site, which allows large-volume Grade-3 leaks to have substantial impact on the local environment. Second, lesser leaks may also ultimately evolve through corrosion, increased operating pressure, accumulation of the emitted gas in closed spaces, or its migration to nearby buildings. For these reasons, grades are not fixed over time and no leak can be regarded as "safely leaking" (Hendrick et al., 2016).

In this context, Massachusetts and the Greater Boston area represent a prominent study setting, not only in terms of the magnitude of the problem, but also in terms of local media coverage, academic attention, as well as in the political response it has generated. Given that pipeline mains across the state's distribution system in 2016 were still made of cast and wrought iron for 14.8% of their mileage, and by bare steel for 6.8% (PHMSA, 2017), it is perhaps unsurprising that technical studies have uncovered *thousands* of leaks in urban areas where the original infrastructure was laid down over the course of the 1800s (Phillips et al., 2013). As a result, about 60%–100% of methane emissions in the urban region of Boston are estimated to emerge from faulty distribution and end use of natural gas (McKain et al., 2015).

This has come at the expected cost for local communities and their residents. In the five years leading to 2017, PHMSA counted a total of 21 pipeline incidents in gas distribution, one fatality, 17 injuries, and \$4.7 million in related damages (PHMSA, 2017).[5] Although estimates for other types of disamenities are naturally harder to come by, costs in the form of lost commodity value and increased end-user rates may also be substantial. Places such as the Boston urban area alone might be losing ~15 billion standard cubic feet in emitted gas each given year, for a value of ~\$90 million (McKain et al., 2015), leading to claims of additional costs to ratepayers for \$39 million a year due to LAUF gas at large (Worcester Business Journal, 2017).

Many parts of the local community are involved in the management and resolution of the gas leak problem in Massachusetts. Public utilities such as National Grid, Eversource, and their local subsidiaries are the officially designated caretakers.[6] They are not only responsible for distributing gas to end-customers per se, but also to record and repair gas leaks appearing throughout their system, often in lockstep with the renewal of the most aging sections of pipeline mileage. To these ends, the involvement by the state administration has proved crucial in setting standards and reparation priorities. Since 2014, an amendment to the state's General Laws has required gas companies to report detailed information on each leak on an annual basis, including its location, grade, and the dates of its classification and reparation.[7]

[4] Although a national standard, the definitions below are readapted from the General Laws of Massachusetts, Chapter 164, Section 144, Subsection b.

[5] These statistics do not include end-user accidents, such as fires and explosions within buildings. The latter may also depend on the faulty state of local infrastructure, albeit in a private setting. In a tragic case, a gas leak ignited to cause the injury of 12 people in the spring of 2014 in Dorchester, a neighborhood in the Southern outskirts of Boston (The Boston Globe, 2014).

[6] Although the market for natural gas retail in Massachusetts was opened to competition in the late 1990s, enabling customers to source their gas from several retail agents and suppliers, designated utilities are still officially tied to given distribution areas, and retain responsibility in delivering gas to homes and businesses, fixing damaged pipes, responding to emergency calls, reading meters, and providing customer service (information retrieved at https://www.mass.gov/service-details/competitive-supply-for-natural-gas). The state of Massachusetts acts as price-regulator for investor-owned public utilities, which make up most of the supply side. On the other hand, municipalities control the remaining four companies, which exempt them in practice from the regulatory supervision of the state's Department of Public Utilities.

[7] The mentioned guidelines are enshrined in the General Laws of Massachusetts, Chapter 164, Section 144.

This complemented a series of provisions already in place, such as those establishing repair priorities for Grade-1 and Grade-2 leaks over those of Grade 3, especially in the case of upcoming projects on the public way.[8] In addition, municipal and state public safety officials can both request companies to disclose information on leaks, as well as ask for a reevaluation of a Grade-3 leak ahead of the routinely scheduled assessment time.

In line with our conceptual framework, civil society can play a significant role in addressing leaks. A host of advocacy groups, single-issue and not, have came to the fore to pressure local companies and authorities alike to intervene. In the Greater Boston area and Massachusetts more broadly, non-governmental and private organizations such as the Home Energy Efficiency Team (HEET), Mothers Out Front, Gas Safety Inc. and the Environmental Defense Fund, along with the technical work by scholars at local academic institutions, have all played a part in promoting debate and action around fugitive methane emissions. As a significant example, some of these groups partnered with the local public utilities to identify large-volume Grade-3 leaks and define, by the fall of 2017, a set of measures to prioritize their reparation (HEET, 2017b).

At the same time, exerting pressure on those responsible to ensure safety and prompt reparations is a task that individuals may undertake, too. Residents of areas affected by a gas leak can not only call the closest Fire Department through 911 upon smelling mercaptan or recognizing other signs of gas irroration (such as vegetation damage), but they can also contact the designated public utility directly or refer the issue to 311 where available. Making phone calls is by and large the course of action most widely suggested by official sources, including the websites of public utilities, advocacy groups, and the Commonwealth of Massachusetts itself. This does not necessarily have to happen only upon the discovery of a leak, although utilities partly rely on the public to report any outstanding issue with their infrastructure. Individuals, just as organized groups, can also exert pressure by checking in with the companies about the reparation status of nearby leaks, as well as contacting their local representatives at the municipality and state level to voice their concern for any disservice.

All in all, while the ability of a community to coordinate to report leaks, demand their reparation and hold officials accountable for it may constitute an essential channel towards the problem's resolution, individuals driven by strong attachment to the general wellbeing of their community may single-handedly act towards the same goals. We speculate that a pure self-interest motive may not be sufficient in our context: as we discuss below, our empirical analysis specifically focuses on gas leaks recorded along the public way. As such, it excludes all reparations which individuals would first and foremost request in defense of their private homes, rather than in protection of common resources.

3.1. Assumptions and research hypotheses

We now outline the set of assumptions that we derive from our setting as a guide to our empirical work.

First, we assume that the relevance of gas leaks in our context, in line with the reported costs and risks, should induce nearby residents to address the problem, either individually or as organized groups. Second, since gas leaks are substantially local in nature, at least in terms of their ease of detection, we posit that the closer an individual lives to a leak, the more she will be compelled to exert pressure on the assigned public utility or local authority. Third, we interpret the then-existing legal framework as requiring budget- and time-constrained utilities to prioritize reparation of Grade-1 and Grade-2 leaks regardless of local pressure.[9] As a result, residents' efforts aimed at having leaks repaired should focus on Grade-3 leaks. Fourth, we assume that calls to repair leaks within private buildings do not depend on community-wide characteristics to the same extent as those found in common spaces, given the immediate safety hazard that the former pose to those affected and their private property. We thus focus on the state of reparation of Grade-3 leaks along a neighborhood's public streets.

Next, we define the three leading hypotheses to test our conceptual framework in the data.

Hypothesis 1 (*Community diversity and gas leak reparation*)**.** Conditional on a community's socio-economic features and the state of its gas leak infrastructure, the chances of end-year reparation of a Grade-3 leak are negatively correlated with that community's diversity.

In our setting, communities and their members must especially push to obtain Grade-3 leak reparations, since public utilities are first and foremost invested in fixing Grade-1 and Grade-2 leaks with relatively rigid priority plans. After controlling for the presence and density of leaks that require constant monitoring and immediate reparation, which simultaneously signal the state of local infrastructure and attract a company's attention to the area, we expect to find that more heterogeneous communities are less successful in protecting their common resources all else equal. Specifically, we consider the relevant local resources to be natural gas and a neighborhood's pipelines more directly, with nearby vegetation, air quality, residents' health, household budget, and property value possibly concurring to a broader concept of local resources.

Hypothesis 2 (*Intervening mechanisms*)**.** Community diversity is negatively correlated with local proxies for social capital and individuals' attachment to their neighborhood.

In other words, we expect suitably defined indicators of social capital and attachment, such as measures reflecting a community's collective efficacy, social networks, and its members' emotional attachment and involvement in local activities, to weaken as diversity rises at the community or neighborhood level. Finally, we conclude with

Hypothesis 3 (*The role of mechanisms for resource protection*)**.** In turn, local social capital and individual attachment are negatively correlated with the chances of Grade-3 leak reparation in a given community.

To the extent that higher levels of social capital and attachment are consistently correlated with a community's ability to rally around the successful management of local resources or the provision of public goods, we expect an increase in diversity to affect Grade-3 leak reparations by depressing a community's potential for collective action or undermining residents' feelings of and drive towards involvement with their neighborhood.

Our conceptual framework may only partially reflect the underlying dynamics of leak reparation. For example, unobservable pipeline features systematically related to local diversity patterns, unmeasured public utility practices, or confounding political pressures from state and municipal officials on utilities may produce a negative correlation as in Hypothesis 1 without directly involving diversity, social capital, or individual attachment. Below, we therefore attempt to assess the validity of our hypotheses.

4. Constructing the leak-level dataset

We rely on three sources of data to test our hypotheses: 2016 geolocated gas leaks across Boston and Cambridge, MA as reported

[8] Since November 2016, Grade-3 leaks with significant environmental impact have been added to the priority group in the case of public-way projects. Also, reparation of leaks within 50 ft of school premises remains a top priority independently of leak grade.

[9] We support this hypothesis based on the high (low) mean reparation rates for Grade-1 and Grade-2 (Grade-3) in our data (Table 1).

Table 1
Repaired and unrepaired leaks by grade and city.

Repaired	Boston			Cambridge			Total		
	Yes	No	% Yes	Yes	No	% Yes	Yes	No	% Yes
Grade-1	648	2	99.7	67	0	100	715	2	99.7
Grade-2	517	100	83.8	72	3	96	589	103	85.1
Grade 3	**57**	**1,273**	**4.3**	**2**	**256**	**0.8**	**59**	**1,529**	**3.7**

by the Home Energy Efficiency Team (HEET, 2017a); block and block-group level demographic and economic data from the 2010 U.S. Census and the 2011–2015 American Community Survey (NHGIS, 2017); and survey-based information on social capital and individual neighborhood attachment from the 2008–2010 Boston Neighborhood Survey (Dataverse, 2017). All data are used to characterize the relevant environment surrounding a given gas leak, the fundamental unit of observation of our analysis. Below, we describe each data source and outline our leak-level data construction process.

4.1. Gas leak data

We obtain the gas leaks database from the Home Energy Efficiency Team (HEET) web page. HEET is a grass-roots, non-governmental organization that promotes energy efficiency as a means to diminish greenhouse gas emissions, and has been deeply involved in assessing the prevalence and incidence of methane emissions from gas leaks throughout Massachusetts. Since 2014, utility companies operating in Massachusetts are required to report the location of repaired and unrepaired leaks to the state's Department of Public Utilities on a yearly basis, and HEET gathered and mapped this information. We focus on data pertaining to 2016 and the area of Boston and Cambridge since, at the time of writing this article, HEET records leak grades only for these municipality-year combinations.

In total, we access information about 3049 leaks, including the closest street address, grade, city, date of reparation for repaired leaks, and date of initial reporting for unrepaired leaks.[10] We remove 42 leaks with addresses that lack enough information to correctly georeference them.[11] The address of each remaining leak is geocoded using Google's Geocoding API. Lastly, we remove ten leaks that fall outside of Cambridge' and Boston's boundaries, leaving us with a working sample of 2997 leaks of all grades and 1588 Grade-3 leaks. The latter will constitute the unit of observation of our analysis.

Table 1 breaks down the number of repaired and unrepaired leaks by grade across Boston and Cambridge in 2016. As expected, the vast majority of Grade-1 and Grade-2 leaks undergo reparation by the end of the year, thus presumably with no need for significant pressure on the part of local residents. On the other hand, the relatively lower priority that resource-constrained utility companies can place on repairing Grade-3 leaks (repaired on average in 3.7% of cases) allows for the characteristics of local communities to play a role.

4.2. Demographic and economic data

To characterize the demographic and economic features of what we define to be a leak's relevant community, we draw data from the IPUMS National Historical Geographical Information System (NHGIS) (Manson et al., 2023). We collect available block-level data from the 2010 Census, while the 2011–2015 American Community Survey (ACS) provides us with information at the block-group level. Vector data with the polygons of Massachusetts blocks and block groups are similarly downloaded from this source. We further obtain an additional vector data with polygons of Massachusetts cities, used to clip the data to Boston and Cambridge, from the Government of Massachusetts online archive. Finally, vector data with polygons defining neighborhoods in Boston and Cambridge, used to create neighborhood-level fixed effects in our regression analysis, comes from the online archives of the cities of Cambridge and Boston.

We jointly use these geographic and socio-economic data to construct all controls in our regression analysis, as we further detail in Section 5. The 2010 Census provides us with the number of residents by race (White, Black, Native American, Asian, Native Hawaiian and Other Pacific Islander, Other, and Multiple) and Latino origin, the number of people in each age interval by gender, and the number of occupied and unoccupied housing units by ownership or renting status.

From the 2011–2015 American Community Survey we obtain the number of families whose income is above and below the poverty line, the number of people aged 25 or older at each one of six education achievement levels (some high school or less, high-school degree, associates degree, incomplete college, college degree, and post-graduate studies), the number of households speaking different types of languages (English Only, Spanish, Other Indo-European, Asian and Pacific Island, Other), the number of households that speak limited English for each language group besides English, and the block-group's per capita income.

4.3. Local social capital and individual attachment to the community

We hypothesize that the potential effect of diversity on local resource management operates through social capital and individual attachment to one's community. We shed light on these mechanisms leveraging data from the Boston Neighborhood Survey (henceforth, BNS). The BNS is a telephone-based survey project covering Boston adults conducted by the Injury Control Research Center at the Harvard T.H. Chan School of Public Health. We mostly exploit the aggregated version of the project's last two rounds carried out in 2008 and 2010.[12] The complete survey covers a variety of topics, from the perceptions of neighborhood security, to the extent of local networks and residents' participation in community activities.

Individual responses are geocoded by their 2000 Census block group of reference, so we can match our more recent data with responses averaged at the block-group level based on the 2010 Census block-group boundaries. After this aggregation phase, we select and work with a total of seven composite variables that the BNS provides to measure various aspects of our proposed mechanisms, such as collective efficacy, the strength of social networks, and the extent of emotional attachment or active engagement with one's own neighborhood. We describe these survey-based variables in more detail in Appendix C.

[10] We download data as KML files from the HEET web page and extract information using Python's *beautifulsoup* package.

[11] We discard 25 leaks that cannot be precisely geocoded because their addresses do not specify a road intersection or a street number, hence preventing their exact mapping; seven leeks with addresses that do not specify their road type (e.g. street, avenue, court) and are ambiguous because two or more roads have the same name in the same city; four leaks whose address displays a street number that lies outside the official numbering of the same street; four leaks listing seemingly non-existing intersection addresses; two addresses with seemingly non-existing street-town combinations; and one leak generically located at the Logan International Airport.

[12] The original data includes over 4000 interviews from three survey rounds starting in 2006. Unfortunately, there is no comparable data for the city of Cambridge, which we also describe in a qualitative exploration in Appendix D.

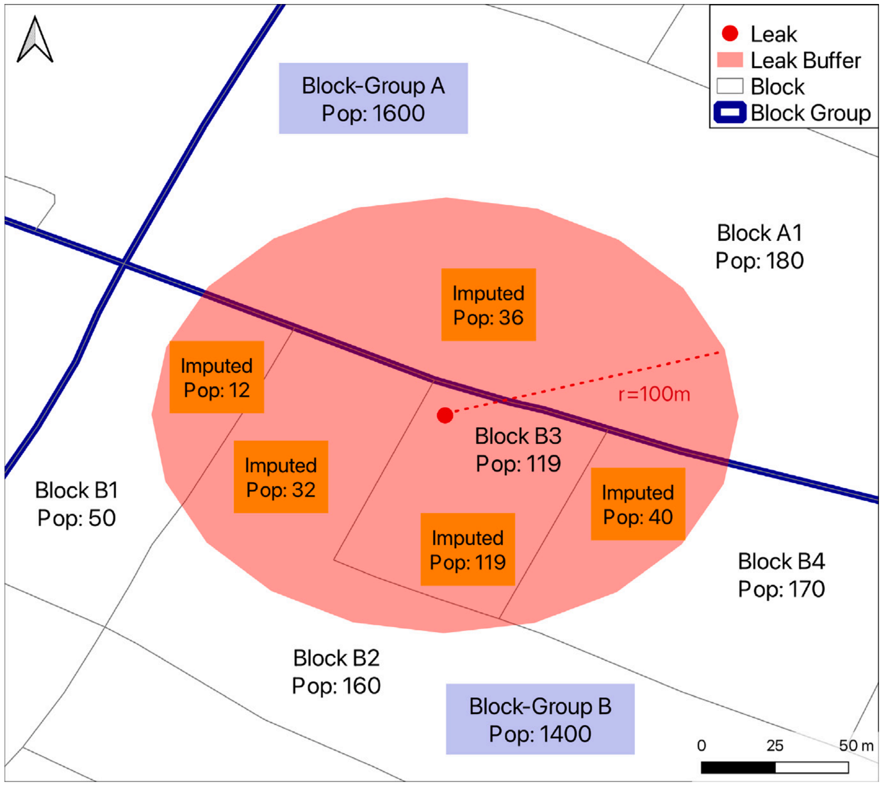

Fig. 2. *Illustration of leak-level data construction:* A radius of length r is set to determine a circular buffer surrounding each leak. The area overlap between the nearby Census blocks or block groups and the buffer helps to determine the weights with which population and other socio-demographic information from the U.S. Census and ACS is aggregated. This process provides us with a data-driven approximation of the characteristics of the community residing near each leak.

4.4. Aggregating data to the gas leak level

We wish to consider each gas leak as a separate observation and study how the features of the surrounding community relate to its end-of-year reparation status. This level of analysis is possible thanks to the structure of our data, which precisely places each leak on the map. Relative to most of the previous literature, we are thus able to go beyond an aggregate analysis at a fixed geography level, and allow each leak to be influenced by its idiosyncratic local conditions and dynamics. This approach constitutes the best available method with the information at hand to address the Modifiable Areal Unit Problem (Gehlke and Biehl, 1934; Openshaw, 1983).

To this end, we characterize a leak's relevant neighborhood as the circular area or buffer surrounding the leak for a given radius of length r. For any given r, we can use the available data to characterize the prevailing features of a leak's buffer (see Fig. 2). For variables available at the block level, the aggregation process consists of a linear combination of the values from the blocks that intersect with a leak's buffer, where the weights correspond to the fraction of the block's area that intersects that buffer.[13] Variables that are otherwise available at the block-group level (e.g. the number of college graduates) are aggregated using population weights. The total population of each block within the buffer is imputed using the process explained above for the block-level variables. The weight used for each block group that intersects with the buffer is equal to the population imputed to lie within the buffer over the total population of the block group. Finally, variables available at the block-group level representing average or median characteristics of the population (e.g. per capita income) are aggregated in a similar fashion, with the denominator now being the total population imputed for the buffer rather than the original block group's population. In this last case, therefore, the weights sum up to one.

5. Community diversity and gas leak reparation

We now set out to study how leak reparations relate to a community's diversity. In our analysis, our main dependent variable will be a leak's reparation status (an indicator equal to 1 when a leak is reported to have been repaired), while we choose our key regressor to be the level of ethno-racial fractionalization (henceforth, ERF) as defined in Alesina et al. (2003), a common measure of local diversity in the U.S. context. Given a neighborhood or buffer of radius r for each leak i and J groups partitioning the neighborhood's imputed population, we compute fractionalization as

$$ERF_i = 1 - \sum_{j=1}^{J} s_{ij}^2, \tag{5.1}$$

where s_{ij} is the share of population in leak i's neighborhood that belongs to group j. The index corresponds to the probability that two randomly drawn individuals belong to different groups. It ranges from 0, reflecting the case of a gas leak located in a perfectly homogeneous neighborhood, to 1, in the limit case where the leak is surrounded by an infinite number of equally sized groups. The ERF considers non-Latino members of each race in the Census data as a distinct group, with Latinos of any race forming the residual category.[14] Fig. 3 shows

[13] Suppose for example that a leak's buffer intersects blocks A and B. Block A has 100 residents and 70% of its area falls within the perimeter of the buffer. Block B has 50 residents and 10% of its area is part of the buffer. We thus estimate the population living in that leak's buffer to be $0.7 \cdot 100 + 0.1 \cdot 50 = 75$.

[14] For a discussion on the choice of groups defining this index in the U.S. context, see Rushton (2008).

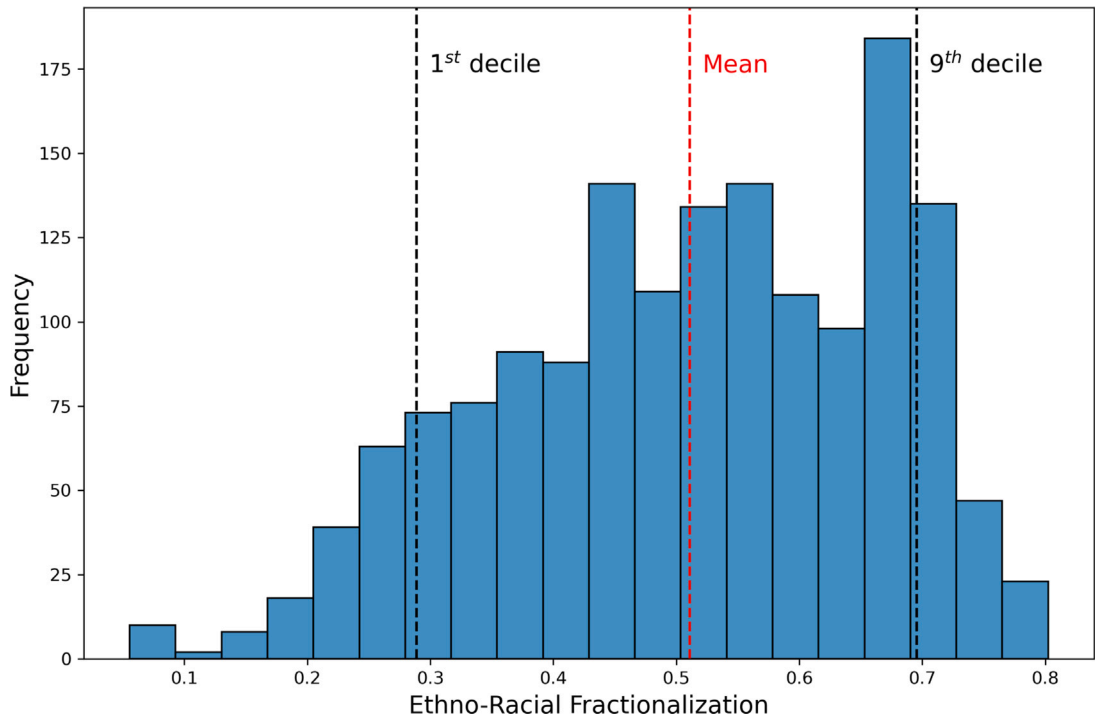

Fig. 3. *Distribution of ethno-racial fractionalization in Grade-3 leak neighborhoods for Boston and Cambridge:* Each leak's neighborhood is defined as a circular buffer area with baseline radius length $r = 500$ m.

the distribution of the ERF across the Grade-3 leaks in Boston and Cambridge in buffer neighborhoods of radius $r = 500$ m. In these data, the ERF almost spans its entire theoretical support, spanning from 0.06 (almost completely homogeneous) to 0.8 (equivalent to a neighborhood equally split among five groups). The distribution is slightly skewed to the left, as might be expected in relatively diverse urban centers with significant group mixing. The ERF mean of 0.5 is equivalent to having the average leak located in a buffer inhabited by two equally sized groups.

Fig. 4 presents a choropleth map of the quintiles of the ERF defined at the block level, together with the location of repaired (red circles) and unrepaired (black dots) leaks. A first look suggests that repaired leaks tend to lie in relatively more homogeneous areas, in line with our first hypothesis. Indeed, the pairwise correlation between a leak's reparation status and ERF measured in the leak's buffer area is negative (−0.06) and statistically significant at the 5% level. The following statistical analysis puts more structure on this observation.

5.1. Estimation strategy

We estimate the following linear probability model (LPM) for each Grade-3 leak i reported in one of 29 city neighborhoods j (see Table A.3 and Fig. A.1) by the utility companies in 2016:

$$Repaired_{ij} = \alpha_j + ERF_{ij}\beta + X'_{ij}\gamma + u_{ij}, \quad (5.2)$$

where $Repaired_{ij}$ is one if the leak was repaired and zero otherwise, α_j are neighborhood fixed effects, ERF_{ij} is the ethno-racial fractionalization index, X_{ij} is a vector including additional controls, and u_{ij} is a zero-mean disturbance. For inference, we use Conley standard errors with a triangular kernel and a 2 km bandwidth to account for spatial autocorrelation.

Our preferred specifications control both for city neighborhood fixed effects and the population shares of the groups that make up the ERF. As a result, our identifying variation comes from the non-linearity underpinning the ERF index in Eq. (5.1), that is, the non-linear relationship between ERF and the individual group shares. This is ensured once we account for any group-specific feature or propensity that can affect leak reparations by controlling for the group shares linearly.

The most comprehensive specifications also control for several socio-demographic features of a leak's neighborhood or buffer, such as its population density (thousand inhabitants per square kilometer), the share of female residents, the fraction of occupied housing units, and the share of the population that reports speaking limited English. We also control for the residents' average age computed as the weighted average of the midpoints of the age intervals provided by the Census, where the weights correspond to the population share in each interval.[15]

Economic controls include the fraction of families below the poverty line, the share of occupied houses that are rented, the natural logarithm of per capita income, and the fraction of residents 25 or older at different levels of educational achievement. Appendix Table A.4 provides the summary statistics. While the study area is small, the data reveal substantial heterogeneity across all leaks' neighborhoods.

We further account for the state of a leak's surrounding infrastructure by including two dedicated variables. First, we control for the local leak density as the number of Grade-1 leaks within a given Grade-3 leak's buffer divided by the buffer area in square kilometers. This measure proxies for the local quality of gas pipelines, since older and more deteriorated pipelines are more prone to leak, addressing the potential correlation between infrastructure quality and a community's socio-economic conditions. The fact that Grade-1 leaks are nearly certainly repaired within year end ensures that our leak density variable is not a so-called bad control, as would be the case with the density of Grade-3 leaks which would reflect both the quality of the local pipelines and the accumulation of unrepaired leaks over time. Second, we control for

[15] For the top interval (85 or above) we used 89.44, which corresponds to the weighted average of expected people alive above 85 assuming a stationary distribution of the population across ages based on the most updated actuarial life tables of the U.S population provided by the Social Security Administration (Social Security Administration, 2017).

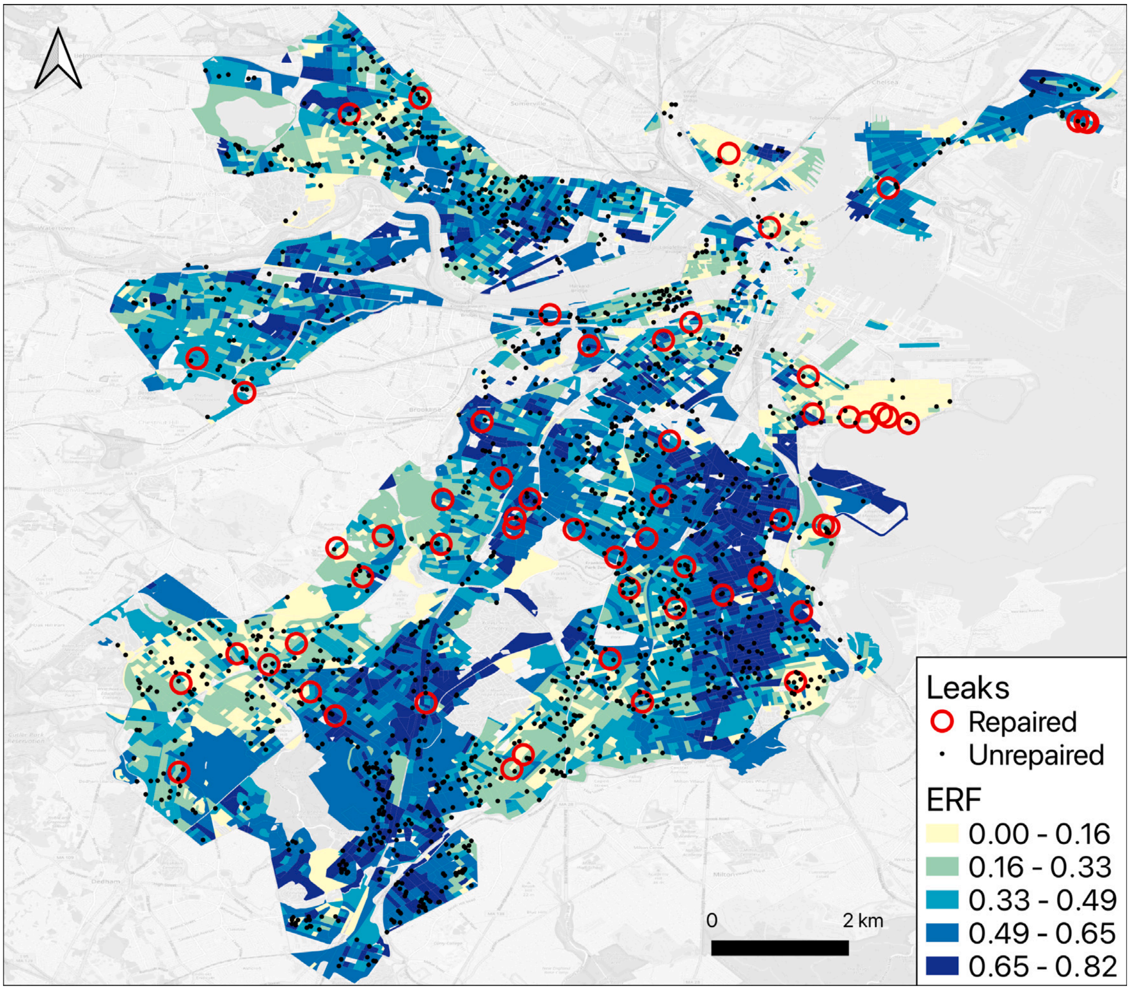

Fig. 4. *Block-level ethno-racial fractionalization and Grade-3 leaks by reparation status:* Quintiles of the ERF for each block in Boston and Cambridge are represented in a color palette from light yellow (less fractionalization) to dark blue (more fractionalization). Black dots represent recorded but unrepaired Grade-3 leaks, while red circles represent Grade-3 leaks reported as repaired by the end of 2016.

the natural logarithm of the distance of a given Grade-3 leak to the closest Grade-1 leak. This second variable could proxy for infrastructure quality at a finer geographical scale than the former, and possibly capture the potential economies of scale in the leak reparation process at the local level: once a Grade-1 leak is recorded and most likely repaired, it might increase the chances that nearby leaks are addressed, too.

Overall, our rich controlling strategy reduces the possibility that any estimated association between reparations and fractionalization is a mere reflection of common local demographic and economic factors, including an area's ethno-racial makeup or the overall quality of its gas distribution infrastructure, as well as other unobserved city neighborhood characteristics. For instance, exploiting within-city neighborhood variation should reduce the scope for residential sorting based on ethnic considerations and gas leak management to drive the results.

5.2. Main results

Table 2 shows the results of the estimation of Eq. (5.2), where we set each leak's buffer radius $r = 500$ m to construct the relevant controls.[16] Column 1 controls for the shares of all ethno-racial groups. This first estimate suggests that the chances that a leak is repaired drops by 18.4 percentage points in a hypothetical highly fractionalized area relative to a fully homogeneous one. The absolute size of the coefficient is only slightly reduced when introducing city neighborhood fixed effects in Column 2, suggesting that resident sorting across neighborhoods is not a fundamental driver of the relationship between ERF and reparations, at least after accounting for the relative presence of each ethno-racial group. Similarly, controlling for the quality of the local infrastructure (Column 3) and for a broad selection of socio-economic features (Column 4) has a relatively limited impact. In our preferred and most stringent specification, the estimate equals −13.9 percentage points, implying that moving from an area at the 10th percentile of the ERF distribution ($ERF = 0.29$) to one at the 90th percentile ($ERF = 0.7$) is correlated with a 5.7 percentage point reduction in the probability of leak reparation. The size of this association is sizable when compared to the mean share of repaired Grade-3 leaks in Table 1 (3.7%).

5.3. Sensitivity to community size

How large is the relevant community surrounding each leak in the analysis? The appropriate buffer radius should approximate the average distance to a leak after which people generally cease to care about the leak's presence. Although we have no accurate way to determine that distance, most of a leak's immediately felt effects are substantially local in nature (including its odor, implied explosion potential, and vegetation damage), so that people would mostly care for leaks that

[16] Table A.5 shows that the corresponding marginal effects from a probit estimation are similar to those in Table 2.

Table 2
Diversity and leak reparation: LPM estimates.

Outcome:	Repaired			
	(1)	(2)	(3)	(4)
Ethno-Racial Frac.	−0.184***	−0.165*	−0.161*	−0.139*
	(0.059)	(0.087)	(0.088)	(0.083)
Controls:				
Group shares	Yes	Yes	Yes	Yes
City neigh. FEs		Yes	Yes	Yes
Infrastructure			Yes	Yes
Econ. & demo.				Yes
N	1,588	1,588	1,588	1,588

Notes: The outcome is a binary indicator equal to 1 when a leak is recorded as repaired by the end of 2016, and 0 otherwise. Ethno-racial fractionalization (ERF) is computed as in Eq. (5.1) for each leak's buffer with radius $r = 500$ m, and has values between 0 and 1. Group shares refer to the population share of ethno-racial groups in each leak's buffer area, with Latinos of all races as a distinct group. Controls are built setting each leak's buffer radius $r = 500$ m. Infrastructure controls include the Grade-1 leak density and a leak's log distance to closest Grade-1 leak. Demographic and economic controls include the share of occupied housing units, population density, the share of females in the population, average resident age, the share of residents reporting only limited English proficiency, the share of households falling below the poverty line, the share of renters, the log of per capita income, and the share of the population with different levels of educational attainment. Conley standard errors (triangular kernel, 2 km cutoff) are shown below point estimates. ***, **, * reflect 1, 5, and 10 percent significance, respectively.

are relatively close to them. To assess the importance of these considerations we estimate the specification in Table 2, Column 4 by resetting the leak buffer radius r between 50 and 1000 m at each 50 meter step. This allows us to redefine the contours of the relevant community surrounding each leak, each time capturing a slightly different set of control features. Fig. 5 plots the evolution of the estimated coefficient for ERF at progressively larger sizes of the buffer area, along with 90% confidence intervals. Along the studied range, estimates are always negative and peak in absolute size at $r = 700$ m, after which they lose significance. This suggests that diversity is a significant determinant of gas leak reparations once we control for the features of the closest surrounding community, while it is harder to carve out a role for fractionalization for a specific leak when a much broader area is considered.

5.4. Alternative measures of diversity

While the relations among ethno-racial groups might be of special importance in the U.S. context, and fractionalization is a common choice to capture these in the applied literature, other cleavages and aspects of diversity might be equally important. Appendix B lists several group-based diversity measures that can help us address this point, including linguistic fractionalization and indices of dissimilarity, isolation, clustering, and concentration in space. Table 3 studies the role of each of these in two ways. First, it evaluates the correlation between each alternative index and leak reparation. Second, it provides a horse race between ERF and each alternative to gauge the robustness of our baseline estimate.

Panel *a* reveals that none of these measures is significantly associated with gas leak reparations and that coefficient signs point in different direction. In turn, panel *b* reveals that both the size and significance of the ERF coefficient estimate mostly survive the inclusion of other indicators, despite the extent of the correlation between ERF and several of these (see Table A.2). Overall, the evidence suggests that, at least in our setting, the interaction across ethno-racial groups is especially tied to a community ability's to solve a collective action problem.

6. Mechanisms

We now turn to unpack the relationship between community diversity and Grade-3 leak reparation in terms of its possible channels. As noted in Section 2, we expect more diverse communities to face relatively larger obstacles in sustaining collective action and achieving shared goals. This could happen due to challenges in forming strong social ties across the underlying (e.g. ethnic or racial) groups and the associated reduction in the residents' emotional attachment to the neighborhood, which can lower their willingness to engage in community-minded activities.

6.1. Diversity, social capital, and attachment

We study these aspects in the case of Boston. Using survey-based data from the 2008 and 2010 BNS waves, we explore the correlations between ERF and seven standardized variables discussed in Appendix C, ideally capturing specific aspects of local social capital and individual attachment in the city.[17]

Panel *a* in Table 4 replicates the analysis in Table 2, Column 4, taking one BNS variable at a time to serve as outcome. Even in communities as small as defined by our baseline leak buffers ($r = 500$ m), the results suggest that ERF is negatively related to several dimensions of social capital, and significantly so in the case of social control (Column 1) and the size of social networks (Column 3). For example, all else equal, a leak emerging in a hypothetical fully fractionalized rather than homogeneous area would be met with residents with networks 2.2 standard deviations smaller. On the other hand, a more mixed picture emerges for individual attachment, with emotional attachment only weakly related to ERF (Column 6) and direct involvement in neighborhood activities significantly higher in more fractionalized areas (Column 7).

We put flesh on our analysis with a qualitative account from two exemplary neighborhoods scoring at the opposite ends of the gas leak and diversity distributions in Greater Boston: South Boston and The Port in Cambridge. Appendix D presents our observations. While the historically Irish-based community in South Boston boasts relatively high levels of Grade-3 leak reparations along with a number of fought-for public goods and signs of collective ethos, the largely diverse Port neighborhood received no reparations in 2016 amidst attempts to forge inter-ethnic trust and a modern multi-cultural identity.

6.2. Social capital as a channel

After confirming that social capital—unlike attachment—falls as fractionalization increases, we assess whether the chances of leak reparations comove consistently with the proposed channel variables. In Table 4, Panel *b*, we regress reparation at the leak level on each BNS variable according to the specification in Table 2, Column 4. All else equal, Columns 1 through 5 imply that one more standard deviation of social capital is associated with 1.6–2.6 percentage points higher chances of Grade-3 leak reparation. On the other hand, attachment variables are negatively but only weakly correlated with gas leak fixes.

Taken together, we take the estimates from both panels to suggest that social capital—as opposed to individual attachment—can play a consistent role in protecting local resources in connection to diversity. For example, conditional on our controls, as ERF rises from 0 to 1 and social control at the community level falls by 1.2 standard deviations,

[17] We aggregate BNS data at the leak-level following the same procedure used to aggregate Census data, but using the 2000 Census geographies since this is the level at which the 2008–2010 BNS data are georeferenced. Out of the 1330 Grade-3 leaks in Boston, 1329 are within 500 m of at least one 2000 Census block group with available BNS observations. These leaks provide the basis for the ensuing analysis.

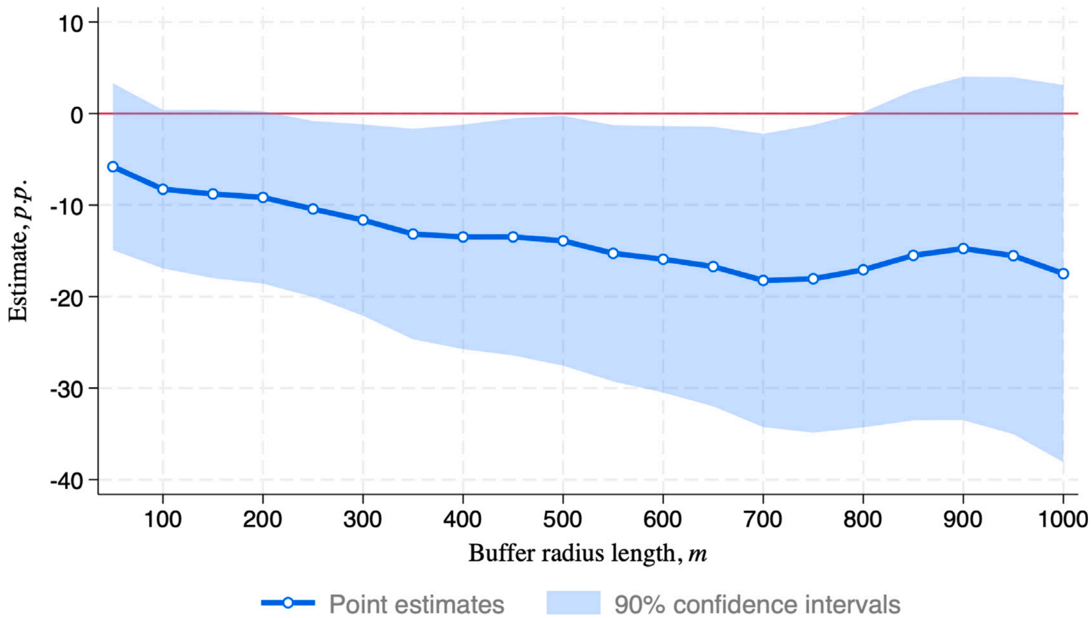

Fig. 5. *Diversity and Grade-3 leak reparation: coefficients by buffer size:* Estimates for the regression coefficient on ethno-racial fractionalization at various lengths of the buffer radius set to construct control variables around each Grade-3 leak. Underlying regressions include all controls from the model in Table 1, Column 4. White dots represent point estimates, while the blue band reflects 90% confidence intervals based on Conley standard errors (2 km cutoff, triangular kernel). In the vertical and horizontal axis titles, *p.p.* stands for percentage points, and *m* stands for meters, respectively.

Table 3
Linear probability model: alternative diversity indices.

Outcome:	Repaired				
Alternative index	Linguistic Frac. (1)	Dissimilarity (2)	Isolation (3)	Clustering (4)	Concentration (5)
Panel a: Alternative index only					
Alternative index	0.110 (0.290)	−0.064 (0.079)	0.075 (0.143)	0.011 (0.037)	−0.005 (0.021)
Panel b: Alternative index & ERF					
Alternative index	0.375 (0.360)	−0.043 (0.075)	0.034 (0.160)	0.037 (0.036)	−0.009 (0.021)
Ethno-Racial Frac.	−0.205* (0.118)	−0.132* (0.080)	−0.134 (0.093)	−0.172** (0.080)	−0.142* (0.080)
Correlation	0.76	−0.07	0.20	0.34	−0.03
N	1,588	1,588	1,588	1,588	1,588

Notes: Indices are described in Appendix B. For all regressions, the outcome is the binary indicator of a leak's reparation status. Panel *a* substitutes ERF with the alternative index written at the top of the column. Panel *b* includes both ERF and the alternative specified at the top of the column. Specifications include all controls from Table 2, Column 4. In Column 1, we further control for the population shares of the linguistic groups underlying the linguistic fractionalization index. Conley standard errors (triangular kernel, 2 km cutoff) are shown below point estimates. ***, **, * reflect 1, 5, and 10 percent significance, respectively. At the bottom of the table, "Correlation" reports the pairwise correlation between ERF and each alternative index.

the chances of Grade-3 leak reparations drop by approximately 3.1 percentage points—a substantial change compared to the 3.7% year-end reparation mean.

These results call for further exploration of the channels linking a community's fragmentation to its public good quality. First, future research can more exhaustively investigate the role of the public utilities in determining Grade-3 reparation schedules on top of Grade-1 and Grade-2 leak interventions, paying particular attention to their procedures when responding to organized groups and the community's demand for leak reparation.

Second, local politics may also provide an indirect and complex channel. Career-concerned local councilors could be sensitive to her constituents' opinions and expect higher electoral punishment from homogeneous communities upon underperformance (Alesina et al., 2019), since residents in the same precinct may also share political preferences. This would leave a role for opinion and preference commonality, while our BNS exercise should otherwise account for the organizational aspect of political feedback provision to underperforming representatives. Future work should shed light on this potential mechanism.

7. Conclusion

Natural gas leaks are a strikingly common feature of U.S. urban environments, but their distribution may not be homogeneous within cities. Looking at the cities of Boston and Cambridge in 2016, we show that Grade-3 leaks have a substantially higher chance of survival in communities observably more diverse along ethnic and racial lines, even after controlling for plausible confounders such as local infrastructure quality, income, education, the size of several demographic groups, and the presence of households with limited English proficiency. Ethno-racial fractionalization is robustly associated to lower reparation chances when we consider households within up to 750 m away from a gas leak. These correlations are generally robust to

Table 4
Social capital and individual attachment as mechanisms.

	Social capital					Ind. attachment	
	Collective efficacy		Social networks				
	Soc. Con. (1)	Soc. Coh. (2)	Soc. Net. (3)	IG Cl. (4)	Rec. Ex. (5)	Miss (6)	Inv. (7)
Panel a: Outcome is the BNS variable in the column header							
ERF	−1.192***	0.500	−2.217***	−0.851	−0.822	0.365	1.912**
	(0.448)	(0.518)	(0.538)	(0.731)	(0.675)	(0.626)	(0.829)
Panel b: Outcome is Repaired. Regressor is the BNS variable in the column header							
BNS var	0.026***	0.020**	0.016**	0.019*	0.016**	−0.003	−0.007
	(0.010)	(0.009)	(0.008)	(0.010)	(0.008)	(0.008)	(0.007)
N	1,329	1,329	1,329	1,329	1,329	1,329	1,329

Notes: Panel *a* regresses each standardized BNS variable on ethno-racial fractionalization (ERF). Panel *b* regresses a leak's reparation status on a given standardized BNS variable measuring social control, social cohesion, social networks, intergenerational closure, reciprocal exchange, the extent to which respondents would miss their neighborhoods upon moving, and their neighborhood involvement, respectively. BNS variables are described in Appendix C. Regressions include the controls from Table 2, Column 4. Conley standard errors (triangular kernel, 2 km cutoff) are shown below point estimates. ***, **, * reflect 1, 5, and 10 percent significance, respectively.

addressing the spatial structure of our data, including using city neighborhood fixed effects to account for residential sorting and varying the definition of the relevant community surrounding a leak.

Pinning down all intervening mechanisms might prove more elusive. These mechanisms might ultimately reflect both long-term discrimination practices across communities as well as the difficulty in coordinating for the public good within them. An emerging environmental justice literature suggests that historically unfair racial relations in the U.S. have marginalized non-White communities to areas with worse environmental conditions (Luna and Nicholas, 2022; Weller et al., 2022). We approach our study context through the lenses of a vast literature in the social sciences linking diversity to collective action challenges within communities in and outside of the U.S. Yet, while we observe that more diverse areas report lower levels of social capital and that in turn social capital is positively related to gas leak reparation, external factors such as local politics or resource-constrained utility companies might also play a relevant role behind the associations that we observe.

Interestingly, alternative measures of diversity such as linguistic fractionalization or indicators of group sorting across space do not seem to have as much predictive power for gas leak reparations. We take this as a general warning that, at least in our setting, the mere coexistence of different ethnic and racial groups in a relatively limited urban area can fundamentally affect that community's ability to promote collective well-being and protect local resources.

CRediT authorship contribution statement

Felipe Jordán: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Enrico Di Gregorio:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Risks and costs associated to natural gas leaks

We provide a brief list of the main risks and costs that natural gas leaks emerging from the distribution pipelines can pose to urban communities. These risks and costs include but are not limited to:

a. **Fire and explosions**: given the right conditions, gas accumulating in closed spaces and subject to ignition sources can escalate into full blown fires and outright explosions. The Pipeline and Hazardous Materials Safety Administration, a U.S. Department of Transportation's agency, records a yearly average of 121 pipeline related incidents over the 2007–2016 period, which led to an average of 10 fatalities, 51 injuries, and approximately $34 million in property damages during the same timespan (PHMSA, 2017).
b. **Vegetation damage**: being lighter than air, gas leaking from urban centers' underground pipelines migrates through paths of least resistance, and interferes with the roots of any local vegetation in its way. Natural gas components alter the soil mixture, resulting in oxygen deprivation and root asphyxiation of trees, shrubs and lawn grass alike, whose chlorosis and necrosis develop faster with the volume of the involved leak (Hoeks, 1972; Davis, 1977; Fraedrich, 2009). In this sense, even emissions which are not deemed directly dangerous to people and property may still result in visible damage to the affected area.
c. **Greenhouse gas (GHG) emissions**: methane (CH_4), by and large the main component of natural gas, is an extremely potent GHG, with each molecule trapping 86 times as much heat as one of carbon dioxide over a 20-year period (IPCC, 2013). According to the Environmental Protection Agency, in 2015 natural gas systems provided the second largest anthropogenic source of methane emissions (EPA, 2017).
d. **Health hazard**: by creating ground-level ozone, as well as releasing potent volatile organic chemicals, gas leaks contribute to smog formation in urban areas. The associated methane emissions can thus exacerbate the incidence of asthma and other respiratory conditions, playing a role in premature mortality through reduced air quality and increased pollution (West et al., 2006).
e. **Nauseating odor**: public utilities add mercaptan to the otherwise colorless and odorless methane before it reaches urban areas to facilitate detection when the gas escapes its pipelines. While this grants nearby residents the ability to spot a leak and promptly act to forestall excessive concentration and possible ignition, even a small quantity of mercaptan is recognizable for its rotten eggs odor due to its sulfurous content. According to Columbia Gas, a utility company operating in Massachusetts, the smell is almost unbearable in its concentrated form.

Table A.1
Selected diversity indices: overview.

Index	Description	Formula
Ling. Fractionalization	Fractionalization across linguistic groups.	$1 - \sum_{m=1}^{M} \pi_m^2$
Dissimilarity	Average deviation from evenness distribution across blocks.	$\sum_{m=1}^{M} \sum_{j=1}^{J(i)} \frac{t_j}{2TI} \lvert \pi_{jm} - \pi_m \rvert$
Isolation	Average extent to which minority groups interact with themselves.	$\sum_{m \neq w}^{M} \pi_m \sum_{j=1}^{J(i)} q_{jm} \pi_{jm}$
Concentration	Average ratio between area inhabited by minority groups and the majority group, relative to maximum possible ratio.	$\sum_{m \neq w}^{M} \pi_m \left(\frac{1 - \frac{\sum_{j=1}^{J(i)} q_{jm} a_j}{\sum_{j=1}^{J(i)} q_{jw} a_j}}{1 - \frac{\sum_{j=1}^{n_1} t_j a_j}{\sum_{j=J(i)-n_2}^{J(i)} t_j a_j}} \right)$
Spatial Proximity	Average extent individuals live close to other individuals from their own group, relative to average proximity to all individuals.	$\frac{\sum_{m}^{M} \pi_m \sum_{j=1}^{J(i)} \sum_{k=1}^{J(i)} q_{jm} q_{km} \exp(-d_{jk})}{\sum_{j=1}^{J(i)} \sum_{k=1}^{J(i)} q_j q_k \exp(-d_{jk})}$

Notes: Each index is calculated for each leak i neighborhood, using the imputed population for each one of the M groups (linguistic or ethno-racial) indexed by m, in each one of the $J(i)$ blocks that intersects the leak's buffer, indexed by j. The subscript w stands for the largest ethno-racial group in the area, non-latino whites. π_{jm} and π_m are the fraction of m-group individuals in block j and leak i neighborhood, respectively. I is Simpson's Interaction Index, $I = \sum_{m=1}^{M} \pi_m (1 - \pi_m)$. t_j and T are the total population of block j and leak i neighborhood, respectively. q_{jm} is the share of m-group individuals that live in block j. a_j is the area of block j, n_1 is defined as the rank of the block where the commutative distribution of population equals the minority population of group m in leak i's neighborhood (ordering blocks from smallest to largest), and n_2 is the same for the majority population.

f. **Reparation- and lost supply-related costs**: managing leaks, as much as financing the cost of lost supply, befalls the public utility companies which dominate the U.S. retail landscape. However, given that most such companies are either local natural monopolies due to high fixed infrastructure costs and economies of density, or price-regulated entities (generally by state administrations), they can often shift this burden onto end-user customers (Hausman and Muehlenbachs, 2019). Although regulators have in some cases put caps to the increase in service rates associated to recouping these costs, allowing for a leak in one's community may in practice amount to the implicit acceptance of a utility's wasteful practices, and the resulting cost that derive from them.

Appendix B. Alternative diversity indices

Ethno-racial fractionalization (ERF) captures but one of several dimensions of community diversity identified in the literature. If any alternative dimension correlates both with ERF and leak reparation, our estimates of the baseline correlation between ERF and leak reparation would be biased. We assess the relevance of this concern by studying five additional indices of diversity commonly used in the literature (Table A.1). Our main goal is to test whether our main estimates remain stable when controlling for these relevant dimensions of diversity. At the same time, this effort allows us to place our preferred measure of diversity in a broader conceptual and empirical context and to see if measures reflecting different aspects of diversity display similar correlations with gas leak repairs. As a companion to the analysis in Section 5.4, this appendix discusses each index and reports all pairwise correlations in Table A.2.

B.1. Discussion

We start by introducing linguistic fractionalization, a straightforward counterpart to fractionalization along ethnic and racial lines. This index is meant to capture a relevant cleavage along which diversity could hamper collective action, as language differences often translate into consolidated cultural barriers and higher costs of coordination (Alesina et al., 2003; Glennerster et al., 2013; Sturm and De Haan, 2015). Linguistic fractionalization is defined as ERF (Eq. (5.1)), but it leverages the presence of speakers of different languages in lieu of ethno-racial groups. In the U.S. context, the type of language that residents speak at home is recorded and used to compute group shares. In our data, these groups cover speakers of English, Spanish, other Indo-European languages, Asian and Pacific Island languages, and all other languages as a residual category.

Besides for the level of cross-group fragmentation, the way in which groups are distributed across space might also affect a community's ability to address collective action problem. The literature has highlighted the links of residential segregation with disproportional policing of minorities (Desmond and Valdez, 2013), conflict in contested boundaries (Legewie and Schaeffer, 2016), erosion of trust (Ziller and Spörlein, 2020), among other findings that support a negative relationship between segregation and social capital. Massey and Denton (1988) classified twenty indices of residential segregation commonly used in the literature into five categories reflecting five orthogonal dimensions: unevenness or dissimilarity, isolation, concentration, clustering, and centralization. We calculate an index for each of the first four dimension, for each leak neighborhood i, using census blocks $b_{ij} \in J(i)$ that intersect with the buffer that defines leak i's neighborhood. We exclude centralization since it is meant to captures an area's distance to the city center, a concept which would not work well with our already narrowly-defined leak buffers (with radius $r = 500$ m at baseline).

Dissimilarity or unevenness measures the extent to which different groups are distributed unevenly across blocks within a leak's neighborhood. It is minimized when the share of different groups in every block equals the share of groups in the leak's neighborhood as a whole, and maximized when each block is inhabited by only one group. We follow the recommendation in Massey and Denton (1988) and select the Dissimilarity Index as our preferred measure of unevenness. This is calculated as the weighted average of the absolute deviation of a group's share of the block's population from the same group's share of the population in the leak's buffer, relative to the maximum possible deviation. For the calculation, we use the multigroup dissimilarity formula presented in Table 2 of Reardon and Firebaugh (2002),

$$DIS_i = \sum_{m=1}^{M} \sum_{j=1}^{J(i)} \frac{t_j}{2TI} \lvert \pi_{jm} - \pi_m \rvert, \tag{B.1}$$

where M in the total number of groups, π_{jm} the share of group m in block j's population, π_m the population share of group m in leak i's neighborhood, t_j the total population of block j, T the total population in leak i's neighborhood, and I is Simpson's Interaction Index, $I = \sum_{m=1}^{M} \pi_m (1 - \pi_m)$ (Lieberson, 1969; White, 1986).

The index of isolation for a given group m in the buffer of leak i measures the population share of group m that a randomly drawn individual from m residing in leak i's buffer shares his or her block with. Massey and Denton (1988) recommend using this index (or its converse, exposure) on the back of its wide use in the literature and its orthogonality to the other dimensions. To accommodate our setting with multiple minorities (non-White ethno-racial groups), we calculate the overall isolation index for each leak i as a weighted average of the isolation indices computed for each minority group. That is,

$$ISO_i = \sum_{m \neq \text{White}}^{M} \pi_m \sum_{j=1}^{J(i)} q_{jm} \pi_{jm}, \tag{B.2}$$

where q_{jm} is the share of groups' m members that live in block j.

Concentration refers to the extent to which a minority group is concentrated in space. It captures the fact that areas with the same fraction of a minority population will seem more segregated to an observer if the minority is concentrated in a smaller share of the area. After surveying the indices used in the literature, Massey and Denton (1988) provide their own index, which we adapt for our own purpose. The index measures the area inhabited by a minority group m relative to the majority group, as a fraction of the maximum possible ratio—i.e. when, for given block populations and total majority and minority populations, the minority and majority groups are distributed such that their areas are minimized and maximized, respectively:

$$CON_{im} = \frac{1 - \frac{\sum_{j=1}^{J(i)} q_{jm} a_j}{\sum_{j=1}^{J(i)} q_{jw} a_j}}{1 - \frac{\sum_{j=1}^{n_1} t_j a_j}{\sum_{j=J(i)-n_2}^{J(i)} t_j a_j}}, \tag{B.3}$$

where the subscript w stands for White (majority group), a_j is the area of block j, n_1 is defined as the rank of the block where the commutative distribution of population equals the minority population in leak i's neighborhood (ordering blocks from smallest to largest), and n_2 is the same for the majority population.[18] Blocks are ordered from the smallest to the largest. Then, we calculate a leak's neighborhood concentration index as the weighted average of the concentration indices of all minority groups:

$$CON_i = \sum_{m \neq \text{White}}^{M} \pi_m CON_{im}. \tag{B.4}$$

Clustering encapsulates the extent to which a group's members live close to one another. It differs from concentration, since the same level of concentration is consistent with homogeneous communities from different groups in a checkerboard pattern (low clustering) or the same communities living in large continuous clusters (high clustering). Massey and Denton (1988) recommend the Spatial Proximity Index (SP) as a measure of clustering, as it strikes a balance between potentially being sufficiently uncorrelated with other dimensions of segregation and having a well-behaved distribution. For group m in a leak i's neighborhood, the average proximity is calculated as:

$$P_{im} = \sum_{j=1}^{J(i)} \sum_{k=1}^{J(i)} q_{jm} q_{km} \exp(-d_{jk}), \tag{B.5}$$

where $-d_{jk}$ is the Euclidean distance between blocks j and k. Then, the SP is defined as the weighted average of all groups' proximity indices relative to the proximity of the total population. That is:

$$SP_i = \frac{\sum_m^M \pi_m P_{im}}{P_i}, \tag{B.6}$$

where P_i is defined analogously to (B.5), but using the whole population of each block group.

Below, Table A.2 reports the pairwise correlations across our new alternative indicators of diversity at the leak buffer level, along with their correlation with ERF. In most cases, all alternative indices correlate with each other and are significantly correlated with ERF (with the exception of concentration), and linguistic fractionalization and clustering show the largest positive correlations. This suggests that we should extend our baseline analysis to including these alternative indices, a task that we take on in Section 5.4.

[18] Equality is not possible in practice, as the number of blocks that intersect a leak is finite and small in practice. Thus, we select the rank of the first block whose cumulative distribution is larger than the groups share. This leads to a small number of Grade-3 leaks (10, less than 1% of the sample) with a value smaller than −1, the theoretical minimum of the Concentration index. We set the value of those leaks' concentration index equal to −1.

Appendix C. Additional details on data sources

We draw from the publicly available data from the Boston Neighborhood Survey (henceforth, BNS) to measure the extent of social capital and individual attachment across the city's neighborhoods as suggested by the Survey documentation. We describe all measures below.

To capture the extent of a neighborhood's social capital, we measure two related concepts as follows:

- For Collective Efficacy, defined in the BNS as a neighborhood's capacity to accomplish shared goals, we rely on the two underlying variables that the Survey documentation suggests to measure this concept, namely[19]:
  (1) **Social Control**, or the perceived ability of a neighborhood to enforce shared norms, as the mean of five original survey items.
  (2) **Social Cohesion**, or the strength of positive social relationships between people in the neighborhood, as the mean of five survey items.
- For Social Networks, we rely on three variables which capture different aspects of the size and strength of such relations:
  (1) **Social Networks**, or the size of the social networks of neighborhood residents, as the mean of two survey items.
  (2) **Intergenerational Closure**, or the extent to which unrelated adults and children in the neighborhood have positive relationships, as the mean of five survey items.
  (3) **Reciprocal Exchange**, or the perceived degree to which neighbors interact with one another in the neighborhood, as the mean of five survey items.

We then turn to aspects of individual attachment to one's own neighborhood. To reflect this concept, we draw from two distinct variables, ideally gauging the emotional and attitudinal or behavioral aspects of attachment, respectively:

(1) **Emotional Attachment**: the average response to the 2008 survey question "If you had to move away, would you miss this neighborhood?". Answers are given on a 4-point scale ranging from "Not at all" to "A lot". Although one could interpret this question in a utilitarian way, as residents could feel attached to their neighborhood purely due its amenities or material benefits, we believe it to be sufficient to also capture the emotional dimension of attachment.
(2) **Neighborhood Involvement**, or the participation and awareness with respect to neighborhood activities, as the sum of four survey items.

As we work with the sample of 2016 Grade-3 leaks, we standardize each of these measures so that they are centered around their mean computed across all 1588 leaks' buffer areas and one unit reflects one standard deviation.

This provides us with seven variables uniquely reflecting the social capital and individual attachment conditions across Boston's neighborhoods as recorded before gas leak reparation data. We use this information to provide suggestive evidence on the channels linking diversity to a community's relative success in managing their resources.

[19] For more details, we refer the reader to the BNS Documentation (Dataverse, 2017).

Table A.2
Pairwise correlation across all diversity indices.

Index	ER Frac.	Ling. Frac	Dissimilarity	Isolation	Clustering	Concentration
ER Frac.	1					
Ling. Frac	0.76**	1				
Dissimilarity	−0.07**	−0.23**	1			
Isolation	0.20**	0.47**	0.05*	1		
Clustering	0.34**	0.13**	0.75**	0.08**	1	
Concentration	−0.03	−0.03	−0.05*	0.06*	−0.18**	1

Notes: **, * reflect correlations that are statistically significant at the 1 and 5 percent level, respectively.

Appendix D. Qualitative evidence: visiting two exemplifying communities

Our empirical analysis has explored the correlation between community diversity and leak reparation, further showing that diversity is also at least negatively related to social capital proxies (though positively to attachment proxies), that is, one of our proposed mechanisms. In this appendix, we add to the statistics by providing a more qualitative account of the type of neighborhoods that our data portray.

We set out to identify a number of blocks across two neighborhoods within the area of study that would maximize distance in terms of diversity and leak reparations, and proceeded to visit them in the course of November 2017. The stated goal was to qualitatively record any tangible evidence of the main mechanisms we describe in our conceptual framework, in order to ensure their plausibility. Our demographic and gas leak sources helped us select The Port in Cambridge ($ERF = 0.63$, Grade-3 leak reparation share = 0), and the historical neighborhood of South Boston ($ERF = 0.23$, Grade-3 leak reparation share = 0.21).[20]

As shown in Fig. 4, the map compiled by HEET in 2016 provides a rather grim picture for the state of gas leaks in the municipality of Cambridge, with several Grade-3 leaks concentrated in vast sections in the center and the west of the city. The Port is as a case in point, a densely populated neighborhood which roughly extends between the MBTA T-stops of Central and Kendall Square just north of Massachusetts Avenue. The Port has perhaps one of the most interesting history of urban development in Middlesex county, going from being one of the early hubs of the candy industry in the United States since the 1800s, to hosting a sizable share of Cambridge public housing and eventually becoming a beacon of cultural diversity, housing relatively large African American and Latino communities at the doorstep of MIT labs and diverse businesses.

This diversity is reflected throughout the block-groups we visited, which roughly stretch from Broadway to Main Street, and include a minor intersection of the latter with Windsor Street where, among other instances, the data report the astounding number of 7 unrepaired leaks in 2016. A multiplicity of churches of different denominations in a relatively limited area along with the presence of Puerto Rican flags signal the possibility that the community is not just fragmented along racial lines. Places such as Clement Morgan Park commemorate the residents' early contribution to the cause of the Civil Rights Movement, including Clement Morgan being the first black Cambridge City Councilor and a co-founder of the Niagara Movement in 1905 (a predecessor of the National Association for the Advancement of Colored People).

Translating this diversity into social capital and local attachment is perhaps a fairly recent idea in the context of a neighborhood that was still defining its name by the late 2015, when residents officially ditched the former, more impersonal label of Area 4. While at the time of writing the local Cambridge Art Center showcases murals celebrating the community's "variety of languages, cultures, races and skin color", and the sense of group closeness it engenders, the situation might have been quite different in the previous decades[21]: a poignant article featured in 1961 on the Harvard Crimson, a university student newspaper, belittled those living in the nearby public housing projects, known as Washington Elms, as "a community that might have been the model for Emile Durkheim's frightening concept of anomie" (The Harvard Crimson, 1961).

On the other hand, the blocks we visited in South Boston seemed to display the type of correlates one would expect from a fairly high leak reparation community. To better observe the characteristics of areas most proximate to our main sequence of 2016 repaired leaks in this neighborhood, we started from the intersection of Dorchester and E 8th Street, eventually reaching M Street Beach just off Boston's Old Harbor. The corresponding southern stretch of what Bostonians popularly refer to as "Southie" combines at least two distinctive urbanistic elements that may relate to our mechanisms: on the one hand, the orderly and dense rows of houses, which partly slope down from Telegraph Hill in a setting akin to that of the typical Tuscan town, often celebrated in the classical contributions on social capital (Putnam et al., 1994; Guiso et al., 2016). On the other hand, the recurrent presence of history, both through monuments celebrating national events and memorials devoted to local achievements and members of the community.

This mix of national and local historical elements potentially blurs the distinction between the specificities of the neighborhood's community and the original events that brought it into existence. The key moment that transformed the high point of South Boston (known as Dorchester Heights) into a signature place of early American history, is the so-called Evacuation Day, when British troops were forced out of Boston on March 17, 1776 following the fortification of the Heights by the Continental Army. The event is commemorated nowadays by a county-wide holiday and a national historic site showcasing plaques commemorating both the Royal Army's retreat and the subsequent development of the local community.

The panels around the main monument provide evidence of the tight connection that may arise between community homogeneity, public good provision, and local resource management. According to these sources, in the second half of the 1800s, local calls for clean water and an open park space first and for a local high school next were met by the city. A fundamental role was played by the dense Irish working class which many current residents can trace their origins back to.

This strong Irish presence, which contribute to the low fractionalization score for the area, still stands out as one of the distinctive elements of South Boston, even as this morphs into an assimilated American identity (the world-famous Saint Patrick's Parade that Southie hosts

[20] The two neighborhoods also differ in terms of linguistic fractionalization measures, with The Port scoring 0.51 and South Boston scoring 0.28. This suggests that differences in their ability to coordinate across resident groups might also reflect their different stages in the assimilation of foreign migrants, who are more likely to report speaking a language different than English at home.

[21] The quote is one of many along these lines written on the then recent-looking murals that the Cambridge Art Center has sponsored to host along its entry alley.

Fig. A.1. Boston and Cambridge city neighborhoods.

every year is held in close connection with Evacuation Day). The neighborhood is dotted with American and Irish flags alike, and Irish symbology such as the shamrock adorns the many alleys informally dedicated to fallen members of the U.S. Army and Catholic priests. This matches the visual impression of a predominantly White neighborhood that does not yet display the level of wealth more typical of the American suburbia historically targeted by affluent Whites' flights. Still, the abundance of public goods such as the well-kept public beaches and the community centers standing where the old bath houses used to be might rationalize the strength of the correlations found in Section 6 irrespective of local affluence.

We take the accounts from our visits to contribute to a familiar picture: on average, more homogeneous communities tend to live in the proximity of well-managed resources and public goods, and display their attachment to the prevailing local culture with more vigor, as shown in the past and present of South Boston. Where diversity is high, on the other hand, the coexistence of multiple groups can greatly stimulate creativity, but may not readily translate into the extent of cross-group coordination needed to adequately manage the full set of available local resources.

Appendix E. Additional tables and figures

See Fig. A.1 and Tables A.3–A.6.

Table A.3
Boston and Cambridge city neighborhoods: Grade-3 leak count.

Boston		Cambridge	
Neighborhood	Grade-3 leaks	Neighborhood	Grade-3 leaks
Allston/Brighton	77	Agassiz	15
Back Bay/Beacon Hill	61	Area 2/MIT	7
Central	25	Cambridge Highlands	5
Charlestown	17	Cambridgeport	15
East Boston	36	East Cambridge	10
Fenway/Kenmore	38	Mid-Cambridge	47
Hyde Park	229	Neighborhood Nine	44
Jamaica Plain	67	North Cambridge	23
Mattapan	110	Riverside	13
North Dorchester	56	Strawberry Hill	10
Roslindale	108	The Port	23
Roxbury	150	Wellington-Harrington	14
South Boston	33	West Cambridge	32
South Dorchester	178		
South End	41		
West Roxbury	104		
Total	**1,330**	**Total**	**258**

Table A.4
Descriptive statistics for Grade-3 leak buffer areas.

Variable	Mean	Std. Dev	Min	p25	p50	p75	Max
Repaired (indicator)	0.037	0.189	0.000	0.000	0.000	0.000	1.000
Ethno-Racial Fractionalization	0.511	0.155	0.056	0.399	0.523	0.651	0.802
Frac. White	0.464	0.296	0.008	0.173	0.521	0.728	0.972
Frac. Black	0.267	0.259	0.002	0.048	0.160	0.458	0.849
Frac. Native American	0.002	0.001	0.000	0.001	0.002	0.003	0.007
Frac. Asian	0.074	0.074	0.001	0.018	0.046	0.112	0.515
Frac. Pacific Islander	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Frac. Other	0.014	0.027	0.000	0.004	0.006	0.013	0.260
Frac. Multiple	0.026	0.011	0.006	0.018	0.024	0.031	0.084
Frac. Latino	0.152	0.107	0.008	0.063	0.122	0.223	0.655
Grade-1 Leak Density	6.561	5.037	0.000	2.548	5.096	8.917	33.121
Log Closest Grade-1 Leak	4.992	1.038	0.000	4.575	5.150	5.627	6.989
Frac. Occupied	0.929	0.030	0.747	0.916	0.934	0.949	1.000
Population Density	6,429	3,475	293	3,640	6,070	8,333	20,058
Average Age	36.4	3.8	20.9	34.1	36.4	38.7	62.6
Frac. Female	0.527	0.030	0.227	0.513	0.530	0.545	0.706
Poverty (% HH below FPL)	0.139	0.097	0.000	0.058	0.119	0.209	0.514
Frac. Renters	0.582	0.186	0.047	0.462	0.617	0.708	1.000
Log. Income	9.089	0.503	7.895	8.725	9.015	9.353	11.203
Frac. Limited English	0.093	0.065	0.000	0.039	0.087	0.128	0.410
Frac. High School	0.218	0.109	0.020	0.119	0.236	0.308	0.439
Frac. Some College	0.146	0.063	0.016	0.090	0.146	0.199	0.309
Frac. Associates	0.049	0.028	0.005	0.026	0.046	0.067	0.162
Frac. Bachelors	0.233	0.100	0.051	0.155	0.231	0.316	0.521
Frac. Graduates	0.228	0.169	0.007	0.078	0.168	0.364	0.675
N (Grade-3 leaks)	1,588						

Notes: The unit of observation is a Grade-3 leak's buffer area or neighborhood defined by a radius $r = 500$ m. For each leak, we follow the methodology outlined in Section 4.4 to aggregate all underlying information from the U.S. Census, ACS, or the HEET database which collects the leak details reported by utility companies to the Commonwealth of Massachusetts. Leak reparation status is a binary indicator, while variables reported as fraction and the level of poverty are constrained in the 0-1 range. As a result, the summary statistics for these variables can be interpreted as percentages.

Table A.5
Diversity and leak reparation: probit marginal effects.

Outcome:	Repaired			
	(1)	(2)	(3)	(4)
Ethno-Racial Frac.	−0.172**	−0.144	−0.122	−0.138
	(0.071)	(0.099)	(0.107)	(0.117)
Controls:				
Group shares	Yes	Yes	Yes	Yes
City neigh. FEs		Yes	Yes	Yes
Infrastructure			Yes	Yes
Econ. & demo.				Yes
N	1,588	1,099	1,099	1,099

Notes: Estimates reflect the marginal effect of ethno-racial fractionalization on the probability that a Grade-3 leak is repaired. In each column, controls are as described for Table 2. Standard errors clustered at the city-neighborhood level in parenthesis are shown below the point estimates. **, * reflect 5, and 10 percent significance, respectively.

Table A.6
Diversity and leak reparation: complete LPM estimates.

Outcome:	Repaired			
	(1)	(2)	(3)	(4)
Ethno-Racial Fractionalization	−0.184***	−0.165*	−0.161*	−0.139*
	(0.059)	(0.087)	(0.088)	(0.083)
Group shares:				
Frac. Black	−0.044	−0.096	−0.101	−0.086*
	(0.031)	(0.064)	(0.063)	(0.050)
Frac. Latino	0.197*	0.084	0.079	0.155
	(0.101)	(0.155)	(0.157)	(0.182)
Frac. Native	7.043*	4.640	5.205	6.465
	(4.021)	(4.818)	(4.747)	(5.141)
Frac. Asian	0.205*	0.182	0.180	0.198
	(0.122)	(0.159)	(0.157)	(0.172)
Frac. Hawaiian	9.579	5.380	6.564	2.040
	(20.918)	(22.347)	(22.589)	(24.511)
Frac. Multiple	−1.195**	−0.521	−0.545	−0.627
	(0.591)	(0.781)	(0.774)	(0.881)
Frac. Other	0.612**	0.261	0.241	0.308
	(0.281)	(0.328)	(0.325)	(0.310)
Infrastructure:				
Grade-1 Leak Density (leaks/km^2)			0.001	0.001
			(0.002)	(0.002)
Log. Closest G1-Leak			−0.008	−0.009
			(0.006)	(0.006)
Demographic & economic:				
Frac. Occupied				0.210
				(0.205)
Pop. Density (Thousand/km^2)				0.002
				(0.003)
Average Age				−0.000
				(0.003)
Frac. Female				0.381
				(0.247)
Frac. Below Poverty				−0.067
				(0.151)
Frac. Renters				−0.066
				(0.048)
Log. Per Capita Income				0.026
				(0.032)
Frac. Limited English				0.212
				(0.150)
Frac. High School				0.255
				(0.178)
Frac. Some College				0.397*
				(0.215)
Frac. Associate Degree				0.076
				(0.340)
Frac. Bachelor Degree				0.177
				(0.170)
Frac. Graduate Degree				0.229
				(0.209)
City neighborhood FEs	No	Yes	Yes	Yes
N	1,588	1,588	1,588	1,588

Notes: Whites are the omitted ethno-racial category in all regressions. Conley standard errors (triangular kernel, 2 km cutoff) are shown below point estimates. ***, **, * reflect 1, 5, and 10 percent significance, respectively.

References

Alesina, A., Baqir, R., Easterly, W., 1999. Public goods and ethnic divisions. Q. J. Econ. 114 (4), 1243–1284.

Alesina, A., Devleescauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. J. Econ. Growth 8, 155–194.

Alesina, A., Gennaioli, C., Lovo, S., 2019. Public goods and ethnic diversity: Evidence from deforestation in Indonesia. Economica 86 (341), 32–66.

Alesina, A., La Ferrara, E., 2000. Participation in heterogeneous communities. Q. J. Econ. 115 (3), 847–904.

Alesina, A., La Ferrara, E., 2002. Who trusts others? J. Public Econ. 85 (2), 207–234.

Alesina, A., La Ferrara, E., 2005. Ethnic diversity and economic performance. J. Econ. Lit. 43 (3), 762–800.

Algan, Y., Hémet, C., Laitin, D., 2016. The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. J. Polit. Econ. 124 (3), 696–733.

Boston Neighborhood Survey (BNS) Dataverse, 2017. Boston neighborhood survey responses. Last Accessed 20 November 2017 at https://dataverse.harvard.edu/dataverse/BostonNeighborhoodSurvey.

Davis, S., 1977. The effect of natural gas on trees and other vegetation. J. Arboric. 3 (8), 153–154.

Desmond, M., Valdez, N., 2013. Unpolicing the urban poor: Consequences of third-party policing for inner-city women. Am. Sociol. Rev. 78 (1), 117–141.

Easterly, W., Levine, R., 1997. Africa's growth tragedy: Policies and ethnic divisions. Q. J. Econ. 112, 1203–1250.

EPA, 2017. Environmental protection agency: Inventory of U.S. greenhouse gas emissions and sinks: 1990–2015. Last Accessed 12 December 2017 at https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2015.

Fraedrich, B., 2009. Gas injury to trees: Identification and treatment. Bartlett Tree Research Laboratories Technical Report.

Gehlke, C.E., Biehl, K., 1934. Certain effects of grouping upon the size of the correlation coefficient in census tract material. J. Amer. Statist. Assoc. 29, 169–170.

Glennerster, R., Miguel, E., Rothenberg, A.D., 2013. Collective action in diverse Sierra Leone communities. Econ. J. 123 (568), 285–316.

Górny, A., Toruńczyk-Ruiz, S., 2014. Neighbourhood attachment in ethnically diverse areas: The role of interethnic ties. Urban Stud. 51 (5), 1000–1018.

Greif, M., 2009. Neighborhood attachment in the multiethnic Metropolis. City Community 8 (1), 27–45.

Guiso, L., Sapienza, P., Zingales, M., 2016. Long-term persistence. J. Eur. Econom. Assoc. 14 (6), 1401–1436.

Habyarimana, J., Humphreys, M., Posner, D., Weinstein, J., 2007. Why does ethnic diversity undermine public goods provision? An experimental approach. Am. Political Sci. Rev. 101 (4), 709–725.

Hausman, C., Muehlenbachs, L., 2019. Price regulation and environmental externalities: Evidence from methane leaks. J. Assoc. Environ. Resour. Econ. 6 (1), 73–109.

HEET, 2017a. Home energy efficiency team: City maps of gas leaks. Last Accessed 14 November 2017 at https://www.heetma.org/squeaky-leak/natural-gs-leaks-maps/.

HEET, 2017b. Home energy efficiency team: Shared action plan. Last Accessed on 8 December 2017 at https://www.heetma.org/fixbiggasleaks/shared-action-plan/.

Hendrick, M., Ackley, R., Sanaie-Movahed, B., Tang, X., Phillips, N., 2016. Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. Environ. Pollut. 213, 710–716.

Hjort, J., 2014. Ethnic division and production in firms. Q. J. Econ. 129 (4), 1899–1946.

Hoeks, J., 1972. Effect of leaking natural gas on soil and vegetation in urban areas. Agricultural Research Reports, Wageningen (NL).

Hopkins, D., 2009. The diversity discount: When increasing ethnic and racial diversity prevents tax increases. J. Politics 71 (1), 160–177.

Howarth, R., 2014. A bridge to nowhere: Methane emissions and the greenhouse gas footprint of natural gas. Energy Sci. Eng. 2 (2), 47–60.

IPCC, 2013. Climate change 2013: The physical science basis. Cambridge University Press, Cambridge.

Kyle, G., Theodori, G., Absher, J., Jun, J., 2010. The influence of home and community attachment on firewise behavior. Soc. Nat. Resourc. 23 (11), 1075–1092.

Laurence, J., Bentley, L., 2016. Does ethnic diversity have a negative effect on attitudes towards the community? A longitudinal analysis of the causal claims within the ethnic diversity and social cohesion debate. Eur. Sociol. Rev. 32 (1), 54–67.

Legewie, J., Schaeffer, M., 2016. Contested boundaries: Explaining where ethnoracial diversity provokes neighborhood conflict. Am. J. Sociol. 122 (1), 125–161.

Lieberson, S., 1969. Measuring population diversity. Am. Sociol. Rev. 34 (6), 850–862.

Luna, M., Nicholas, D., 2022. An environmental justice analysis of distribution-level natural gas leaks in Massachusetts, USA. Energy Policy 162, 112778.

Manson, S., Schroeder, J., Riper, D.V., Knowles, K., Kugler, T., Roberts, F., Ruggles, S., 2023. IPUMS National Historical Geographic Information System: Version 18.0 [dataset]. IPUMS, Minneapolis, MN, http://dx.doi.org/10.18128/D050.V18.0.

Massey, D.S., Denton, N.A., 1988. The dimensions of residential segregation*. Soc. Forces 67 (2), 281–315.

McKain, K., Down, A., Raciti, S., Budney, J., Hutyra, L., Floerchinger, C., Herndon, S., Nehrkorn, T., Zahniser, M., Jackson, R., Phillips, N., Wofsy, S., 2015. Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. Proc. Natl. Acad. Sci. USA 112 (7), 1941–1946.

Miguel, E., Gugerty, M.K., 2005. Ethnic diversity, social sanctions, and public goods in Kenya. J. Public Econ. 89 (12), 2325–2368.

NHGIS, 2017. National historical geographical information system: U.S. geographic summary data and boundery files. Last Accessed 17 November 2017 via IPUMS at https://www.nhgis.org/.

Olson, M., 1965. The logic of collective action. Harvard University Press, Cambridge.

Openshaw, S., 1983. Concepts and techniques in modern geography. Geo Books, Norwich.

Ostrom, E., 1990. Governing the commons: The evolution of institutions for collective action. Harvard University Press, Cambridge.

Ostrom, E., 2000. Collective action and the evolution of social norms. J. Econ. Perspect. 14 (3), 137–158.

Phillips, N., Ackley, R., Crosson, E., Down, A., Hutyra, L., Brondfield, M., Karr, J., Zhao, K., Jackson, R., 2013. Mapping urban pipeline leaks: Methane leaks across Boston. Environ. Pollut. 173, 1–4.

PHMSA, 2017. Pipeline and hazardous materials safety administration: Data and statistics overview. Available at https://www.phmsa.dot.gov/data-and-statistics/pipeline/data-and-statistics-overview. (Last Accessed on 5 December 2017).

Pradhananga, A., Davenport, M., 2017. Community attachment, beliefs and residents' civic engagement in stormwater management. Landsc. Urban Plan. 168, 1–8.

Putnam, R., 2007. E pluribus unum: Diversity and community in the twenty-first century. The 2006 johan skytte prize lecture. Scand. Political Stud. 30 (2), 137–174.

Putnam, R., Leonardi, R., Nanetti, R., 1994. Making democracy work: Civic traditions in modern Italy. Princeton University Press, Princeton.

Reardon, S.F., Firebaugh, G., 2002. Measures of multigroup segregation. Sociol. Methodol. 32, 33–67.

Rice, T., Steele, B., 2001. White ethnic diversity and community attachment in small Iowa towns. Soc. Sci. Q. 82 (2), 397–407.

Rushton, M., 2008. A note on the use and misuse of the racial diversity index. Policy Stud. J. 36 (3), 445–459.

da Silva, N., Alvarez, I., 2019. Clicks and cliques: Exploring the soul of the community. Comput. Stat. 34 (4), 1537–1563.

Social Security Administration, 2017. Actuarial life table. Last Accessed 30 November 2017 at https://www.ssa.gov/oact/STATS/table4c6.html.

Sturm, J.-E., De Haan, J., 2015. Income inequality, capitalism, and ethno-linguistic fractionalization. Amer. Econ. Rev. 105 (5), 593–597.

Tabellini, G., 2008. The scope of cooperation: Values and incentives. Q. J. Econ. 123 (3), 905–950.

The Boston Globe, 2014. Twelve injured as house explodes in dorchester. Published on April 17, 2014.

The Economist, 2016. A dirty little secret - methane leaks. Published on July 23, 2016.

The Harvard Crimson, 1961. Washington elms - brass tacks. Published on May 31, 1961.

van der Meer, T., Tolsma, J., 2014. Ethnic diversity and its effects on social cohesion. Annu. Rev. Sociol. 40, 459–478.

Vigdor, J., 2002. Interpreting ethnic fragmantation effects. J. Econ. Lit. 75 (2), 271–276.

Vigdor, J.L., 2004. Community composition and collective action: Analyzing initial mail response to the 2000 census. Rev. Econ. Stat. 86 (1), 303–312.

Weller, Z., Im, S., Palacios, V., Stuchiner, E., von Fischer, J., 2022. Environmental injustices of leaks from urban natural gas distribution systems: Patterns among and within 13 U.S. metro areas. Environ. Sci. Technol. 56 (12), 8599–8609.

West, J., Fiore, A., Horowitz, L., Mauzerall, D., 2006. Global health benefits of mitigating ozone pollution with methane emission controls. Proc. Natl. Acad. Sci. USA 103 (11), 3988–3993.

White, M.J., 1986. Segregation and diversity measures in population distribution. Popul. Index 52 (2), 198–221.

Wong, C., 2010. Boundaries of Obligation: Geographic, Racial, and National Communities. Cambridge University Press, Cambridge (UK).

Worcester Business Journal, 2017. Natural gas leaks cost Massachusetts $39m annually. published on October 2, 2017.

Ziller, C., Spörlein, C., 2020. Residential segregation and social trust of immigrants and natives: Evidence from the Netherlands. Front. Sociol. 5, 45.