Do Cities Matter in a Pandemic? Urban Infrastructure and COVID-19 Contagion

Denis Fernandes Alves [*1,2] and André Luis Squarize Chagas [†2]

[1]Federal University of Pernambuco, PPGECON-UFPE
[2]Department of Economics, FEA-USP

Abstract

This paper examines how spatial inequalities in urban infrastructure influence the risk of COVID-19 contagion. Using individual-level georeferenced data from the city of Recife, Brazil, we estimate the effect of access to healthcare, built-environment density (Floor Area Ratio -FAR), and socioeconomic vulnerability (ZEIS zones) on infection probability. To address endogeneity, we implement an IV-Probit model, using historical rail proximity, housing density in the early 2000s, and the proportion of ZEIS per area as instruments for healthcare access, FAR, and ZEIS zones, respectively. We found that greater distance to health services is associated with an increase of 0.616 in the risk of infection. Similarly, higher vertical density, as indicated by the FAR index, is linked to an increase of 0.457 in the risk of contagion. Finally, living in ZEIS zones raises the risk of infection by 0.612, reflecting the impact of greater socio-economic vulnerability and inadequate infrastructure in these areas - legally designated zones of social interest - even after controlling for individual, environmental, and political factors. Distance and the FAR factor have a more pronounced impact on individuals aged 20-30. The ZEIS variable has a stronger influence on individuals under 20, with its impact diminishing in older age groups, suggesting greater sensitivity among younger people to urban infrastructure characteristics. The effects are magnified among older adults, underscoring the intersection of spatial and health vulnerability. Additional analyses confirm that contagion risk varies by urban context and land use type. These findings suggest that urban form and service allocation play a causal role in shaping epidemic exposure. Public health strategies must account for spatial disparities in infrastructure and accessibility, particularly in low-income and peripheral neighborhoods. By linking urban economics and epidemiology, this study contributes to a growing literature on how city structure mediates health risk in unequal urban settings.
Keywords: COVID-19 contagion; Floor Area Ratio (FAR); Urban inequality; Health accessibility; Instrumental variables; Urban infrastructure; ZEIS zones.
JEL Code: C26, I18, R14, R23, R53.
Area 7: Urban issues and metropolises.

Resumo

Este artigo examina como as desigualdades espaciais na infraestrutura urbana influenciam o risco de contágio pela COVID-19. Utilizando dados georreferenciados em nível individual da cidade do Recife, Brasil, estimamos o efeito do acesso à saúde, da densidade do ambiente construído (Índice de Aproveitamento do Solo - FAR) e da vulnerabilidade socioeconômica (zonas ZEIS) sobre a probabilidade de infecção. Para lidar com a endogeneidade, implementamos um modelo IV-Probit, utilizando a proximidade histórica às

*Corresponding Author: denis_fernandes@ufpe.br
†achagas@usp.br

ferrovias, a densidade habitacional no início dos anos 2000 e a proporção de ZEIS por área como instrumentos para acesso à saúde, FAR e zonas ZEIS, respectivamente. Constatamos que uma maior distância aos serviços de saúde está associada a um aumento de 0,616 no risco de infecção. Da mesma forma, uma maior densidade vertical, conforme indicado pelo índice FAR, está ligada a um aumento de 0,457 no risco de contágio. Por fim, viver em zonas ZEIS eleva o risco de infecção em 0,612, refletindo o impacto da maior vulnerabilidade socioeconômica e da infraestrutura inadequada nessas áreas – zonas legalmente designadas como de interesse social – mesmo após o controle por fatores individuais, ambientais e políticos. A distância e o fator FAR têm impacto mais pronunciado em indivíduos com idade entre 20 e 30 anos. A variável ZEIS tem influência mais forte em indivíduos com menos de 20 anos, com seu impacto diminuindo nas faixas etárias mais velhas, sugerindo maior sensibilidade entre os mais jovens às características da infraestrutura urbana. Os efeitos são ampliados entre os idosos, destacando a interseção entre vulnerabilidade espacial e de saúde. Análises adicionais confirmam que o risco de contágio varia de acordo com o contexto urbano e o tipo de uso do solo. Esses achados sugerem que a forma urbana e a alocação de serviços desempenham um papel causal na configuração da exposição a epidemias. Estratégias de saúde pública devem levar em conta as disparidades espaciais na infraestrutura e acessibilidade, especialmente em bairros periféricos e de baixa renda. Ao conectar economia urbana e epidemiologia, este estudo contribui para a crescente literatura sobre como a estrutura das cidades medeia os riscos à saúde em contextos urbanos desiguais.
Palavras-chave: Contágio COVID-19; Floor Area Ratio (FAR); Desigualdades Urbanas; Acessibilidade à saúde; Variáveis instrumentais; Infraestrutura urbana; Zonas ZEIS.
JEL: C26, I18, R14, R23, R53.
Área 7: Questões urbanas e metrópoles.

1 Introduction

Cities are natural focal points for the spread of infectious diseases due to their high population density and patterns of human interaction. In developing countries, this vulnerability is often exacerbated by limited access to public services, residential crowding, and urban segregation. The COVID-19 pandemic has renewed interest in how urban structure shapes contagion dynamics, particularly in large cities in the Global South (Coven et al., 2023).

In Brazil, the virus spread rapidly through metropolitan regions such as São Paulo, Rio de Janeiro, Salvador, and Recife, where social and spatial inequalities are pronounced. While the health effects of the pandemic have been widely documented, less is known about how pre-existing urban infrastructure — especially distance to healthcare facilities and residential density — affected individual exposure to infection. This paper contributes to the literature by exploring the role of spatial access to health services and built-environment characteristics in explaining COVID-19 contagion.

We focus on two mechanisms. First, greater distance from health facilities may increase exposure risk by discouraging early treatment and increasing time spent in transit — often using crowded public transportation. Second, high residential density may facilitate transmission through shared private and public spaces. These hypotheses are examined using detailed individual-level data on COVID-19 testing in Recife, Brazil.

We estimate the causal effect of healthcare access, built environment intensity, and socially vulnerable zones (ZEIS) using an IV-Probit model, where distance to health services is instrumented with proximity to 19th-century railway lines, current verticalization patterns (Floor Area Ratio — FAR) are instrumented with historical residential density from the early

2000s, and ZEIS zones are instrumented by the proportion of ZEIS area per census tract, where the combination of high density, poor infrastructure, and poor access to care can exacerbate exposure risks.

Our results indicate that individuals living in denser areas, in ZEIS, and farther from healthcare services were significantly more likely to test positive for COVID-19. The odds ratio for ZEIS residents exceeds 6.5, even after controlling for individual and neighborhood characteristics. These findings suggest that spatial and socioeconomic inequalities play a central role in shaping vulnerability to epidemic shocks.

This research contributes to a growing literature on the urban determinants of contagion. In the U.S., denser central business districts experienced sharper initial declines in activity and persistent contagion risks (Rosenthal et al., 2021; Liu and Su, 2021). Effective population density has emerged as a strong predictor of outbreak severity (Desmet and Wacziarg, 2022), and housing market evidence from China indicates a shift in preferences toward low-density areas (Huang et al., 2023). On the supply side, pandemic-related strain on health systems disproportionately affected poorer populations across countries (Pujolar et al., 2022; Abel et al., 2024), with particularly acute effects in mental health services (Harrell et al., 2023).

Environmental conditions also influenced infection patterns, with temperature, humidity, and precipitation shown to affect viral spread (Paez et al., 2021; Carvalho et al., 2021; Cerqua and Letta, 2022). Social isolation policies interacted with urban form and socioeconomic status. In cities where enforcement was weaker, and informal employment more prevalent, adherence to distancing measures was lower, and infection rates higher (Li and Ma, 2022; Sheng et al., 2022).

This paper is positioned within this broader empirical effort by focusing on Recife, a city marked by high density, verticalization, and deep spatial inequality (Lima and Silveira Neto, 2019; Oliveira and Neto, 2015). Its heterogeneous urban fabric provides a valuable case to understand how legacy infrastructure and residential patterns influence health vulnerability.

2 Data and Strategy

Data Sources and Construction

We build a geocoded individual-level dataset linking COVID-19 test results to residential characteristics, infrastructure, and spatial covariates in the city of Recife, Brazil. The core dataset comes from the State Health Department of Pernambuco and includes all reported COVID-19 tests conducted in 2020, along with age, gender, comorbidities, and residential addresses.

Residential locations were geocoded to 2010 census tracts and linked to spatial data on healthcare facilities, built environment, and voting behavior. Daily meteorological data were matched to each individual's location and test date. Political variables reflect the proportion of votes for Bolsonaro and Haddad in the 2018 presidential election, at the precinct level.

Definition of Floor Area Ratio (FAR)

A key variable in our analysis is the Floor Area Ratio (FAR), which measures vertical density. It is constructed as:

$$FAR_i = \frac{arc_i + (arp_i \cdot n)}{arl_i} \tag{1}$$

where arc_i is the recorded constructed area, arp_i is the average residential unit size, n is the number of housing units in the building, and arl_i is the lot area. This ratio captures the intensity of land use at the property level and serves as a proxy for built-up density.

Descriptive Patterns

The final sample includes 158,189 individuals. Table 1 shows descriptive statistics by test result. Individuals who tested positive (26%) lived in areas with higher FAR (mean 2.26 vs. 1.56), were slightly older, and more likely to report comorbidities. They also lived, on average, closer to public health services. In addition, the table shows that the proportion of individuals in ZEIS is low, with a mean of 0.02 and considerable variation. The Positive group has a similar mean, while the Negative group has an even lower proportion.

Environmental conditions show moderate variation. Positive cases were more likely to occur on days with higher rainfall and wind speeds, though temperature and pressure were nearly identical across groups. Interestingly, population density was slightly lower in areas with more positive tests, suggesting potential interactions with mobility or testing behavior.

Table 1: Descriptive statistics

Variable	Total		Positive		Negative	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
CVD	0.26	0.43	-	-	-	-
Distance	0.88	0.50	0.79	0.51	0.91	0.49
FAR	1.71	1.66	2.26	1.81	1.56	1.58
ZEIS	0.02	0.13	0.02	0.15	0.01	0.12
Man	0.51	0.50	0.43	0.49	0.55	0.50
White	0.20	0.40	0.23	0.42	0.20	0.40
Age	39.63	15.71	41.97	15.93	38.80	15.54
Age^2	1,817.03	1,427.65	2,014.81	1,474.86	1,746.99	1,403.89
Comorbidities	0.08	0.27	0.10	0.30	0.08	0.27
Population density	4,502.89	2,755.15	3,919.08	2,935.08	4,676.69	2,674.79
Prop. neighborhood tested	0.01	0.00	0.01	0.00	0.01	0.00
Social isolation	0.39	0.04	0.40	0.05	0.39	0.03
Precipitation	3.39	5.92	4.14	7.14	3.12	5.39
Atmospheric pressure	1,012.88	1.68	1,012.95	1.60	1,012.86	1.71
Dew point temp.	21.18	1.16	21.22	1.16	21.16	1.16
Maximum temp.	30.35	1.56	30.29	1.50	30.38	1.58
Average temp.	26.14	1.40	26.06	1.33	26.16	1.42
Minimum temp.	22.88	1.63	22.81	1.50	22.91	1.68
Average relative humidity	75.22	5.15	75.80	6.03	75.01	4.78
Minimum relative humidity	55.45	6.00	56.04	7.01	55.24	5.59
Maximum wind gust	0.74	2.44	1.12	2.91	0.61	2.23
Average wind speed	0.15	0.53	0.23	0.64	0.12	0.49
Prop. of votes for Bolsonaro	0.09	0.17	0.10	0.18	0.08	0.17
Prop. of votes for Haddad	0.08	0.18	0.09	0.20	0.07	0.18
Null and blank votes	0.03	0.03	0.02	0.03	0.03	0.03
Total	**158,189**		**41,368**		**116,821**	

Source: Author's own elaboration.

Figures 1 and 2 illustrate spatial and temporal testing patterns. Central neighborhoods exhibited the highest test volumes and positivity rates. After September 2020, the number of tests expanded sharply, reflecting increased testing capacity and public health mobilization.

The map in Figure 1 illustrates the spatial distribution of 158,189 COVID-19 tests in Recife. Darker blue bars indicate higher testing volumes, concentrated in densely populated neighborhoods like Boa Viagem, Imbiribeira, and Várzea. Red circles represent positive results, with larger circles in the same areas signaling higher infection rates. In contrast, peripheral and northern regions show fewer tests and confirmed cases, likely reflecting underreporting. Overall, testing and positive cases are concentrated in central, high-density areas.

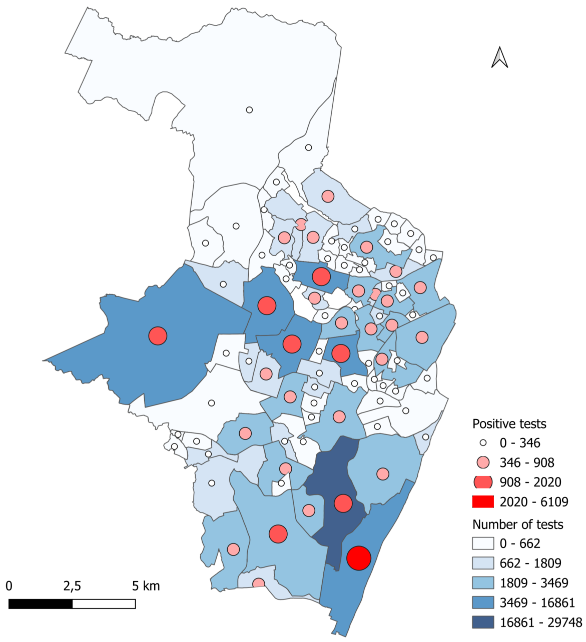


Figure 1: Positive COVID-19 tests by neighborhood

Source: Author's own elaboration.

The graph in Figure 2 shows the monthly progress of two COVID-19 tests carried out in Recife from March to December. There will be a gradual increase in tests in September, followed by a sharp increase in the end, when the total number of tests exceeds 40,000 and remains high in December. The number of positive tests increases more slowly in the first few months (same as lookdown measures proposed by the Prefecture of Recife, the Government of Pernambuco and the federal government between May and June), peaking in the summer, falling in November and again in December. The negative tests show a consistent upward trend, with a significant increase since October, reflecting a substantial expansion in testing efforts.

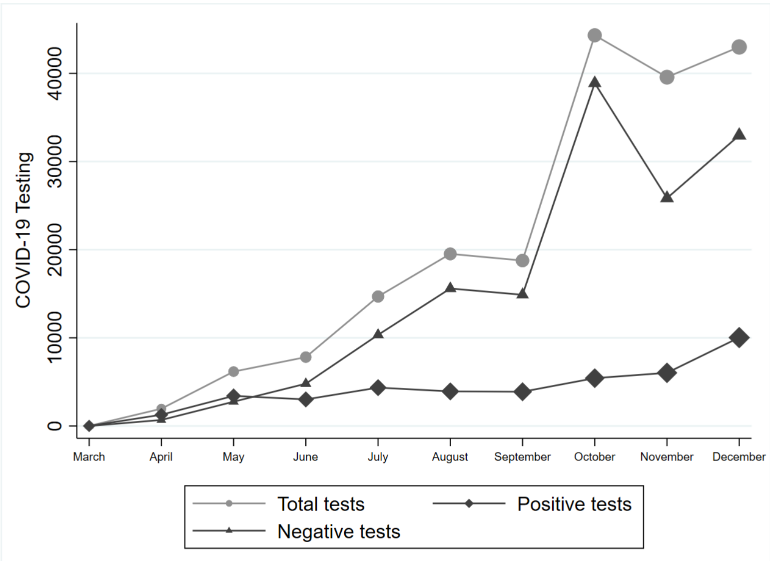


Figure 2: COVID-19 testing timeline, March–December 2020

Note: December data were released with delay in January. Only tests performed in December are included.
Source: Author's own elaboration.

These data indicate an intensification of COVID-19 detection strategies in the second half of the year, possibly linked to the strengthening of public health initiatives of the Recife prefecture and the increase in testing capacity in the city.

Econometric Specification

To estimate the probability of infection, we implement an instrumental variable Probit model (IV-Probit), where the outcome variable CVD_{isht} equals 1 if individual i in household h, located in spatial unit s, at time t, tested positive for COVID-19:

$$\begin{aligned} CVD_{isht} = {} & \beta_0 + \beta_1\, Dist_{is} + \beta_2\, FAR_i + \beta_3 ZEIS_i + Infra'_{isht}\beta_4 + X'_{ih}\beta_5 + Climate'_{it}\beta_6 \\ & + Politicalpos^{2018}_{ih}\beta_7 + \sigma_t + \delta_i + \phi_h + \varepsilon_{isht} \end{aligned} \tag{2}$$

$Dist_{is}$ denotes the distance to the nearest public healthcare facility. $ZEIS_i$ a binary indicator for residence in social interest zones (ZEIS). The vector $Infra_{isht}$ includes housing type and proximity to commercial areas. X_{ih} contains individual-level controls. We include fixed effects for epidemiological week (σ_t), for green and conservation areas (δ_i), and for neighborhood-level unobservables (ϕ_h). Standard errors are clustered at the sector level.

Identification Strategy

For identification, three IVs were used for the variables of interest: distance, FAR and ZEIS. For the first time IV is used in relation to a distance to the health unit, the old tracks of the imperial railway, built in Recife in the second half of the 19th century, were considered, which are no longer in operation (see Figure 3a). The railways were almost pioneering in the city and were intended for the transport of sugar and cotton production to the port of Recife Duarte et al. (2023). The Recife and São Francisco Railway was the second oldest railway established in Brazil, inaugurated in 1858 by the Great Western of Brazil Railway Co., linking Recife to Cabo, covering a distance of 31.5 km Duarte (2020). Later, other railways were built, greatly facilitating the connection between the interior and the coastline of the state Cardoso and de Albuquerque (2020).

The use of the distance to the old tracks of the imperial railway as an IV for the variable distance to health services is justified by its relevance and exogeneity. The tracks, although no longer in operation, were a historical and geographical landmark in the formation of Recife's urban infrastructure, influencing access to health services. This variable is relevant because it directly affects the proximity of individuals to health services, but it is exogenous, as its location was determined by historical factors and not by individual characteristics or decisions related to the search for medical care. Thus, the distance to the imperial tracks serves as a good instrument, as it is not correlated with the model's errors or with endogenous factors related to the choice of the service location.

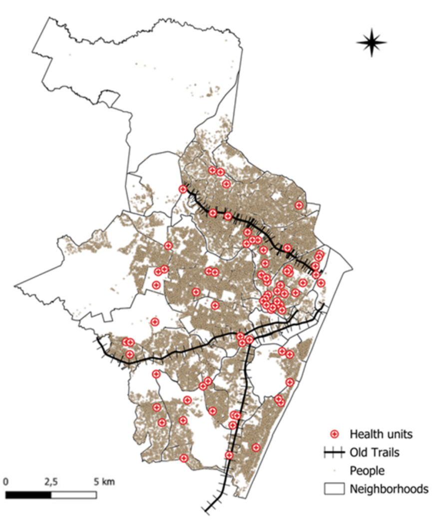
A potential concern regarding the use of distance to the historical imperial railway tracks as an instrument is the presence of indirect channels, whereby proximity to these tracks might correlate with contemporary urban characteristics—such as income, density, or mobility—that affect COVID-19 contagion. To address this, we incorporate a comprehensive set of controls at both the individual and neighborhood levels, including infrastructure indicators, political preferences, climate variables, and fixed effects for environmental and temporal shocks. Moreover, contemporary public health units were not located based on the historical tracks, but rather on more recent urban policies, suggesting that any legacy effect of the railway on the health infrastructure network is mediated through distance itself. Importantly, robustness tests using randomization of the instrument confirm that the observed effects are not spurious. Lastly, the original railway was designed primarily for the transportation of agricultural goods, not for serving residential or commercial zones, which reinforces the plausibility of the exogeneity assumption.

The second instrumental variable corresponds to the historical floor-area ratio (FAR) of the census tract—measured approximately two decades prior—which serves as an instrument for current residential construction density (see Figure 3b). The identification strategy relies on two main assumptions: relevance and exogeneity. First, the urban structure of the city exhibits substantial temporal persistence, such that areas with higher verticalization in the past are strongly predictive of current construction intensity. This persistence ensures instrument relevance, as confirmed by high first-stage F-statistics in our estimations.

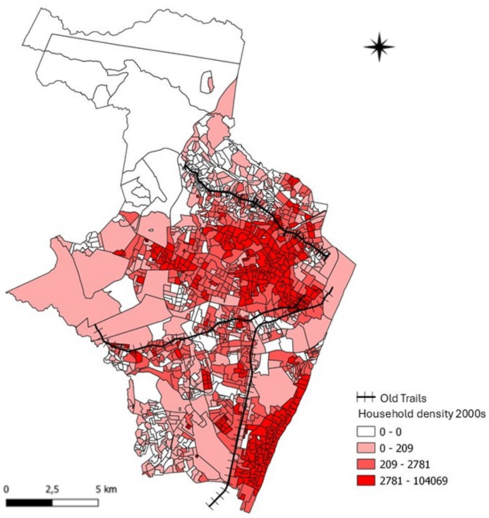
Second, we argue that the historical FAR is exogenous to the contemporary determinants of COVID-19 contagion. Although current construction intensity may be influenced by recent individual or market decisions, the FAR from 20 years ago reflects long-term urban planning and regulatory constraints, rather than short-term behavioral responses. Additionally, we incorporate a rich set of controls—demographic, socioeconomic, climatic, and political—as well as fixed effects at various levels, to mitigate potential omitted variable bias. We further show that once randomized, the instrument loses all explanatory power, reinforcing the validity of our exclusion restriction. This approach follows prior work, such as, which employs similar instruments in the Brazilian urban context.

The third instrumental variable corresponds to the proportion of Special Zones of Social Interest (ZEIS) within each census tract, which is used to instrument the binary indicator for whether an individual resides in a ZEIS area (see Figure 3c). This variable is strongly relevant: sectors with a higher concentration of ZEIS are more likely to contain individuals living in such areas, as ZEIS designation is spatially clustered by municipal housing policies.

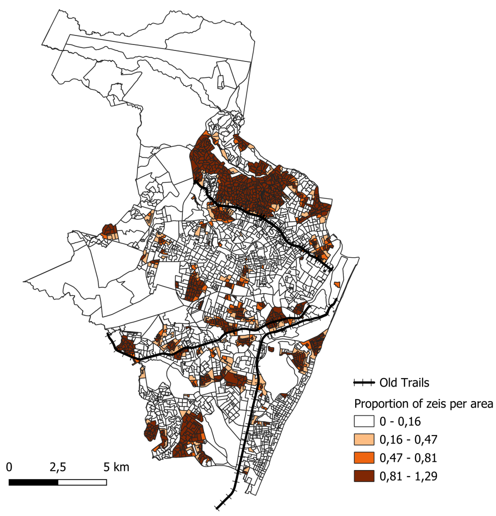
Regarding exogeneity, the ZEIS designation is determined by urban planning regulations and social housing policy, often enacted years before the pandemic. These decisions reflect long-term government priorities, not individual or behavioral characteristics that could be directly related to COVID-19 contagion. Importantly, we control for a wide range of contextual and individual variables—such as income, education, comorbidities, population density, climate, and political alignment—as well as fixed effects for spatial and temporal heterogeneity. These controls help mitigate any remaining correlation between the instrument and unobserved drivers of contagion. Finally, robustness checks involving randomization of the instrument confirm that its predictive power is not spurious, reinforcing its credibility for causal inference.



(a) Old railway tracks



(b) Household density (2000s)



(c) Proportion of zeis per area

Figure 3: Instruments used in the identification strategy

Source: Author's own elaboration.

We conduct a series of robustness checks, including alternative specifications with expanded controls, subsample analyses by age group, and tests using multiple distance and density measures. These include overidentification tests and assessments of instrument strength. All results are presented in Section 3, confirming the consistency and credibility of our identification strategy.

3 Results

3.1 Main Estimates: Urban Infrastructure and Contagion

This section presents the main empirical findings on the relationship between urban infrastructure, spatial inequality, and COVID-19 contagion. We begin by reporting baseline estimates from linear probability models, followed by instrumental variable (IV-Probit) results addressing potential endogeneity in access to healthcare and built-environment density. We then conduct robustness checks across multiple specifications and explore heterogeneity by age group. Finally, we extend the analysis by examining alternative spatial distance measures and the role of FAR across different land use categories.

Table 2 presents linear regression estimates[1] that investigate the association between urban infrastructure and the likelihood of COVID-19 infection. The coefficients of the key variables — distance to the nearest healthcare facility, Floor Area Ratio (FAR), and residence in Special Zones of Social Interest (Zeis) — are statistically significant but show differing signs: while FAR and Zeis have positive coefficients, indicating that higher building density and living in socially vulnerable areas are associated with greater contagion, the distance variable has a negative coefficient, suggesting that greater distance from health services is associated with a lower probability of infection.

This negative result for distance may reflect limitations of the linear model or local particularities, such as underreporting or reduced access to testing in more remote areas, which affect the measurement of contagion.

Additionally, population density, measured as population per square kilometer, also shows a positive correlation with infection. Social isolation is negatively associated with contagion, confirming its protective role. The proportion of individuals tested in each neighborhood is positively related to the likelihood of infection, possibly reflecting the severity of outbreaks or variations in testing coverage.

[1] The Estimations via instrumented linear regression model (2SLS) version can be found in the appendix (see table 2)

Table 2: Estimations via linear regression model

Contagion	1	2	3	4
Characteristics of urban infrastructure				
Distance	-0.070***	-0.069***	-0.057***	-0.058***
	(0.004)	(0.004)	(0.005)	(0.005)
Far	0.040***	0.040***	0.036***	0.036***
	(0.001)	(0.001)	(0.001)	(0.001)
Zeis	0.042***	0.038***	0.032**	0.032**
	(0.014)	(0.014)	(0.014)	(0.014)
Neighborhood infrastructure	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes
Climate	No	Yes	Yes	Yes
Political position	No	No	Yes	Yes
Instruments	No	No	No	No
Fixed Effects				
FE epidemiological week	No	No	No	Yes
FE green area	Yes	Yes	Yes	Yes
FE conservation area	Yes	Yes	Yes	Yes
FE tests by individuals	Yes	Yes	Yes	Yes
Tests of endogeneity				
Durbin (score) chi2(3)	250.78***	263.71***	198.695***	70.14***
Wu-Hausman F(3,63937)	83.89***	88.22***	66.3976***	23.39***
Observations	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

3.2 Instrumental Variable Estimates

To address potential endogeneity in Distance, FAR and ZEIS, we employ an IV-Probit model. Table 3 presents the main results, showing that both variables remain positive and statistically significant after instrumentation. This model is particularly appropriate since it estimates the impact of these variables on the probability of contagion, a dichotomous outcome, as supported by previous studies (Almagro et al., 2021; De Negri et al., 2021; Silva, 2023). Other methodologies applied in Brazil include spatial Poisson models (Brussi Filho, 2023). Distance to health services exhibits a robust effect on contagion, highlighting the role of spatial accessibility in pandemic exposure.

The IV-Probit results reveal a robust and significant effect of urban infrastructure on COVID-19 contagion risk. The distance to the nearest public health unit has a positive coefficient ranging from 0.290 in model 1 to 0.616 in model 4, indicating that greater distance from healthcare services increases the probability of infection, emphasizing the crucial role of spatial accessibility in pandemic exposure. Supporting evidence from (Brussi Filho, 2023) finds that every additional 10 minutes of commuting raises contagion risk by 3.88% among formal workers in Recife.

The FAR variable also shows a significant positive effect, with coefficients between 0.232 and 0.457, meaning that individuals living in areas with higher construction intensity—whether due to buildings or slums—face a higher probability of contagion. Moreover, the ZEIS variable, which identifies Special Zones of Social Interest characterized by poor infrastructure and social vulnerability, has a large and significant positive coefficient ranging from 0.439 to 0.612. This suggests that residents in ZEIS, reflecting precarious housing and higher population density, are substantially more exposed to infection risks in socially segregated urban contexts like Recife. In fact, estimates imply that at the sample median, people living in ZEIS areas were approximately 6.5 times more likely to be infected than those in non-ZEIS zones.

The model controls for additional factors such as neighborhood characteristics, individual traits, climate, and political ideologies. Exogeneity tests (Wald tests) validate the instrumental variables, reinforcing the robustness of these findings. Altogether, the results underscore how poor urban infrastructure and limited access to healthcare significantly increase vulnerability to COVID-19 contagion.

Table 3: IV Probit Estimates

Contagion	**1**	**2**	**3**	**4**
Characteristics of urban infrastructure				
Distance	0.290***	0.374***	0.680***	0.616***
	(0.041)	(0.049)	(0.083)	(0.097)
Far	0.232***	0.304***	0.469***	0.457***
	(0.013)	(0.020)	(0.050)	(0.043)
Zeis	0.439***	0.547***	0.611***	0.612***
	(0.051)	(0.063)	(0.087)	(0.079)
Neighborhood infrastructure	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes
Climate	No	Yes	Yes	Yes
Political position	No	No	Yes	Yes
Instruments	Yes	Yes	Yes	Yes
Fixed Effects				
FE epidemiological week	No	No	No	Yes
FE green area	Yes	Yes	Yes	Yes
FE conservation area	Yes	Yes	Yes	Yes
FE tests by individuals	Yes	Yes	Yes	Yes
Wald test of exogeneity	666,41***	702,96***	750,12***	779,21***
Observations	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

We also find that individuals living in detached houses face greater infection risk, likely due to their location in peripheral areas with lower infrastructure coverage. In contrast, those in multifamily residential buildings exhibit lower probabilities of contagion. Proximity to stores and hospitals is positively associated with infection, possibly due to foot traffic and local crowding.

3.3 Robustness Checks

Table 4 examines robustness across model specifications, incorporating different combinations of controls and fixed effects. The association between Distance and contagion remains consistently positive and significant. Coefficients for FAR vary in magnitude and sign across models, but remain broadly consistent with the hypothesis that greater built-up intensity is linked to higher contagion risk.

Table 4: Randomization of Distance, FAR and ZEIS

Contagion	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Characteristics of urban infrastructure								
Distance	9.349	7.002	16.466	16.985	0.001	0.007	0.022	0.001
	(31.17)	(20.53)	(65.33)	(66.78)	(0.137)	(0.335)	(0.027)	(0.076)
Far	-0.790	-0.523	1.578	1.639	-0.002	-0.002	0.001	-0.002
	(4.355)	(3.443)	(10.09)	(10.30)	(0.052)	(0.142)	(0.004)	(0.052)
Zeis	-0.181	-0.130	-0.195	-0.197	0.041	0.036	0.025	0.024
	(0.930)	(0.717)	(1.274)	(1.314)	(0.005)	(0.775)	(0.174)	(0.248)
Neighborhood infrastructure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Climate	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Political position	No	No	Yes	Yes	No	No	Yes	Yes
Instruments	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects								
FE epidemiological week	No	No	No	Yes	No	No	No	Yes
FE green area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE conservation area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE tests by individuals	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

These results are robust to the inclusion of meteorological controls, political orientation, and week-specific fixed effects. The large sample size and stability across specifications strengthen the internal validity of our main findings.

3.4 Heterogeneity by Age Group

The results presented in Table 5 regarding heterogeneity by age cohorts reveal how urban infrastructure characteristics influence the probability of COVID-19 contagion across different age groups. The variable distance to the health unit has a significant effect only for the 20-30 years age group, with a positive coefficient of 0.743, indicating that for this age group, greater distance from healthcare services is associated with a higher probability of contagion. For the other age groups, the distance variable does not show a significant impact. The variable Far also shows a significant effect only for the 20-30 years group (coefficient of 0.460), suggesting that for this group, living in areas with higher crowding and density increases the likelihood of contagion. For the other groups, the effect is not significant.

Regarding the variable Special Zones of Social Interest, a positive and significant impact is observed across almost all age groups, with the effect being stronger in younger cohorts. The coefficient for the 20-30 years group is 0.719, while for those under 20 years old, it is 0.558. For the 40-60 years, the effect remains positive but with smaller coefficients (0.191), respectively). This suggests that living in ZEIS is associated with a higher probability of COVID-19 contagion across all age groups, but the impact is more pronounced in younger individuals.

For younger groups, such as those under 20 and 20-30 years old, the probability of contagion is more strongly influenced by the presence of ZEIS. This may be linked to higher crowding, poor housing conditions, and limited access to healthcare, which are more common in

ZEIS areas and particularly affect young people who have greater mobility and, consequently, greater exposure to the virus. In contrast, for older groups (40-60) the effect is still positive but more modest. This could reflect lifestyle differences, where older individuals are less exposed due to behaviors like social isolation, while younger people are more exposed through work, social activities, and urban mobility.

Table 5: Age cohorts

Contagion	**<20years**	**20-30**	**30-40**	**40-60**	**>60years**
Characteristics of urban infrastructure					
Distance	0.269	0.743***	0.227	0.068	0.401**
	(0.333)	(0.135)	(0.183)	(0.179)	(0.863)
Far	0.020	0.460***	0.004	0.054	0.423**
	(0.093)	(0.076)	(0.053)	(0.060)	(0.543)
Zeis	0.558**	0.719***	0.044	0.191*	0.429*
	(0.257)	(0.173)	(0.106)	(0.105)	(0.039)
Neighborhood infrastructure	Yes	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes	Yes
Climate	No	Yes	Yes	Yes	Yes
Political position	No	No	Yes	Yes	Yes
Instruments	Yes	Yes	Yes	Yes	Yes
Fixed Effects					
FE epidemiological week	Yes	Yes	Yes	Yes	Yes
FE green area	Yes	Yes	Yes	Yes	Yes
FE conservation area	Yes	Yes	Yes	Yes	Yes
FE tests by individuals	Yes	Yes	Yes	Yes	Yes
Observations	**7,723**	**45,948**	**28,244**	**42,920**	**17,598**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

The estimated coefficients on Distance and FAR are larger in this subsample, indicating that proximity to healthcare and vertical density are especially important for older adults. The effect of residing in ZEIS is also pronounced, reinforcing that elderly individuals in peripheral and socially vulnerable areas face increased risk of exposure due to limited access to infrastructure and healthcare services. While ZEIS areas increase vulnerability across all age groups, the compounded effect of remoteness, poor living conditions, and reduced mobility makes older residents particularly susceptible. Social isolation maintains its protective effect, while the elevated risk in detached homes persists. These results underscore the need for targeted infrastructure and public health interventions for at-risk populations, especially older adults in underserved urban zones.

3.5 Additional Distance and FAR Effects

To explore spatial heterogeneity in more detail, we estimate models with alternative distance measures. Table **??** presents results using distances to subways, parks, beaches, the CBD, and other urban features. In both 2SLS and IV-Probit specifications, longer distances to parks and beaches are associated with higher contagion risk, whereas greater distance from subways and the airport is negatively associated with infection, suggesting a complex relationship between local amenities, density, and exposure. These patterns are especially pronounced in socially vulnerable areas such as ZEIS, where proximity to recreational spaces and central urban zones tends to heighten risk due to structural overcrowding, high mobility, and limited access to healthcare. In such contexts, even beneficial urban features can act as amplifiers of contagion, underscoring how spatial inequality and inadequate infrastructure shape the geography of exposure during a pandemic.

Table 6: Other distances

Contagion	**Subway**	**Parks**	**CBD**	**Beach**	**Avenues**	**Capibaribe**	**Airport**	**Border**
Panel A - 2SLS								
Distance's	-0.098***	2.028***	0.021***	0.042***	0.370	0.031***	-0.028***	0.001***
	(-0.052)	(0.038)	(0.053)	(0.051)	(0.039)	(0.053)	(-0.052)	(0.052)
FAR	0.026	0.457***	0.048	0.168**	-0.986	0.031	-0.035	0.055
	(0.043)	(0.033)	(0.019)	(0.023)	(-0.039)	(0.056)	(-0.061)	(0.097)
ZEIS	0.249**	0.537**	0.258***	0.354***	-0.589	0.221**	0.205**	0.244**
	(0.030)	(0.029)	(0.033)	(0.037)	(-0.015)	(0.029)	(0.027)	(0.031)
Panel B - IVProbit								
Distance's	-0.152***	2.709***	0.113***	0.173***	0.349***	0.171***	-0.142***	0.001***
	(-0.012)	(0.083)	(0.010)	(0.091)	(0.011)	(0.010)	(-0.010)	(0.010)
FAR	0.013	0.438***	0.354**	0.291***	-0.959***	0.178	-0.174	0.306*
	(0.084)	(0.020)	(0.023)	(0.059)	(-0.095)	(0.011)	(-0.011)	(0.020)
ZEIS	0.309***	0.577***	0.488***	0.465***	-0.625***	0.307***	0.271**	0.323***
	(0.037)	(0.076)	(0.048)	(0.075)	(-0.054)	(0.037)	(0.027)	(0.044)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

We then examine the effect of FAR by land use type (Table 7), confirming that higher FAR in ZEIS, residential, and commercial zones is significantly associated with increased risk of COVID-19 contagion. The weighted FAR, calculated using the sample median, also demonstrates a strong positive relationship between building density and transmission across diverse urban settings. These findings align with earlier results showing that ZEIS and slum areas, characterized by overcrowding and inadequate infrastructure, experience the greatest vulnerability, as reflected in the highest coefficients in both 2SLS and IV-Probit models. Residential and commercial densities similarly contribute to heightened contagion risk, though to a somewhat lesser degree. Overall, the evidence underscores how the intensity of the built environment, combined with socio-economic disadvantages—especially in ZEIS zones—amplifies infection risks, highlighting the critical need for targeted public health and urban planning interventions in these high-density, underserved areas.

Table 7: Other FAR

Contagion	Weighted	ZEIS/Slum	Residential	Commercial	Others
Panel A - 2SLS					
FAR's	0.264***	0.388***	0.263***	0.175***	0.188**
	(0.013)	(0.055)	(0.069)	(0.037)	(0.010)
Distance	Yes	Yes	Yes	Yes	Yes
ZEIS	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes
All Fixed Effects	Yes	Yes	Yes	Yes	Yes
Panel B - IVProbit					
FAR's	0.373***	0.871***	0.413***	0.849***	0.534**
	(0.038)	(0.026)	(0.049)	(0.078)	(0.027)
Distance	Yes	Yes	Yes	Yes	Yes
ZEIS	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes
All Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	**158,189**	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.

Together, these findings confirm that spatial variation in urban form—particularly residential density and access to health infrastructure—plays a central role in shaping contagion dynamics. The results reinforce the importance of incorporating spatial and socioeconomic heterogeneity into urban planning and public health strategies.

4 Conclusion

This study aimed to investigate the relationship between COVID-19 contagion and various urban and social characteristics, focusing on access to healthcare units, residential infrastructure density, and conditions in low-income housing areas (ZEIS) in Recife, Brazil. To achieve this, we employed instrumental variable models, specifically the Two-Stage Least Squares (2SLS) and IV-Probit models, to analyze the effects of factors such as proximity to healthcare services, FAR and ZEIS areas on the likelihood of contagion. The analysis controlled for individual and contextual factors through fixed effects, ensuring robust results that accounted for urban characteristics and infrastructure variability.

The results show that urban and residential characteristics significantly influence the risk of COVID-19 contagion. In particular, the proximity to healthcare units, density of residential infrastructure (measured by FAR), and living in ZEIS areas were found to have a positive and significant relationship with the increased likelihood of contagion. Our findings suggest that longer distances from healthcare services increase the risk of contagion, with each additional unit of distance from healthcare increasing the risk by 0.616. Furthermore, the FAR index was associated with a higher risk of infection, with a coefficient of 0.457, indicating that more densely populated areas with poor infrastructure are more prone to virus transmission. Additionally, living in ZEIS areas increased the risk of contagion by 0.612, highlighting the role of socio-economic vulnerability and inadequate infrastructure in these neighborhoods. The study also found that younger age groups, particularly those between 20-30 years old, are more sensitive to the impact of these urban characteristics, showing a greater risk of contagion in these environments.

One of the central contributions of this study is to provide causal evidence that urban inequality—manifested through spatial segregation, infrastructure deficits, and exclusion from healthcare—shaped who was most exposed to COVID-19. The pandemic did not affect urban populations randomly. Individuals residing in peripheral, overcrowded, and underserved areas—particularly in legally defined Zones of Social Interest (ZEIS)—were significantly more likely to be infected. In this sense, the very structure of the city produced vulnerability. Spatial exclusion was not merely correlated with contagion risk; it was a determinant. These findings suggest that future public health responses must go beyond individual behavior or medical capacity. They must confront the deep-rooted spatial inequalities that systematically place certain populations at greater risk during health crises.

Although this study is based on the case of Recife, where urban segregation and infrastructure deficits are pronounced, the generalizability of the findings to other cities—particularly those with different spatial, institutional, or demographic structures—should be approached with caution. Future research could expand this framework to other urban contexts and incorporate a broader set of socio-political and cultural factors. Comparative studies across both developed and developing countries—using spatial econometric models—could further illuminate how the built environment and social vulnerability interact in shaping epidemic dynamics.

References

Abel, Z. D., Roope, L. S., Duch, R., and Clarke, P. M. (2024). Access to healthcare services during the covid-19 pandemic: a cross-sectional analysis of income and user-access across 16 economically diverse countries. *BMC Public Health*, 24(1):2678. pages 3

Almagro, M., Coven, J., Gupta, A., and Orane-Hutchinson, A. (2021). Disparities in covid-19 risk exposure: Evidence from geolocation data. *NYU Stern School of Business Forthcoming*. pages 11

Brussi Filho, P. L. E. (2023). Trajeto casa trabalho: uma análise do impacto do communting na pandemia de covid-19. pages 11

Cardoso, A. L. R. and de Albuquerque, M. Z. A. (2020). Patrimônio ferroviário e urbanizaçao em pernambuco, brasil. *PatryTer*, 3(6):66–80. pages 7

Carvalho, J., Campos, P., Carrito, M., Moura, C., Quinta-Gomes, A., Tavares, I., and Nobre, P. (2021). the relationship between covid-19 confinement, psychological adjustment, and sexual functioning, in a sample of portuguese men and women. *The Journal of Sexual Medicine*, 18(7):1191–1197. pages 3

Cerqua, A. and Letta, M. (2022). Local inequalities of the covid-19 crisis. *Regional science and urban economics*, 92:103752. pages 3

Coven, J., Gupta, A., and Yao, I. (2023). Jue insight: Urban flight seeded the covid-19 pandemic across the united states. *Journal of urban economics*, 133:103489. pages 2

De Negri, F., Galiezz, R., Miranda, P., Koeller, P., Zucoloto, G., Costa, J., Farias, C., Travassos, G., and Medronho, R. (2021). Socioeconomic factors and the probability of death by covid-19 in brazil. *Journal of Public Health*, 43(3):493–498. pages 11

Desmet, K. and Wacziarg, R. (2022). Jue insight: Understanding spatial variation in covid-19 across the united states. *Journal of urban economics*, 127:103332. pages 3

Duarte, L. B. (2020). Acessibilidade ao emprego e resultados no mercado de trabalho. Master's thesis, Universidade Federal de Pernambuco. pages 7

Duarte, L. B., Neto, R. d. M. S., and da Silva, D. F. C. (2023). The influence of job accessibility on individual labor income: Evidence for the city of recife, brazil. *Journal of Transport Geography*, 112:103684. pages 7

Harrell, B., Fumarco, L., Button, P., Schwegman, D. J., and Denwood, K. (2023). The impact of covid-19 on access to mental health care services. In *AEA Papers and Proceedings*, volume 113, pages 420–422. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. pages 3

Huang, N., Pang, J., and Yang, Y. (2023). Jue insight: Covid-19 and household preference for urban density in china. *Journal of Urban Economics*, 133:103487. pages 3

Li, B. and Ma, L. (2022). Jue insight: Migration, transportation infrastructure, and the spatial transmission of covid-19 in china. *Journal of urban economics*, 127:103351. pages 3

Lima, R. C. d. A. and Silveira Neto, R. d. M. (2019). Zoning ordinances and the housing market in developing countries: Evidence from brazilian. *Journal of Housing Economics*, 46:101653. pages 3

Liu, S. and Su, Y. (2021). The impact of the covid-19 pandemic on the demand for density: Evidence from the us housing market. *Economics letters*, 207:110–150. pages 3

Oliveira, T. G. and Neto, R. d. M. S. (2015). Segregação residencial na cidade do recife: um estudo da sua configuração. *Revista Brasileira de Estudos Regionais e Urbanos*, 9(1):71–92. pages 3

Paez, A., Lopez, F. A., Menezes, T., Cavalcanti, R., and Pitta, M. G. d. R. (2021). A spatio-temporal analysis of the environmental correlates of covid-19 incidence in spain. *Geographical analysis*, 53(3):397–421. pages 3

Pujolar, G., Oliver-Anglès, A., Vargas, I., and Vázquez, M.-L. (2022). Changes in access to health services during the covid-19 pandemic: a scoping review. *International journal of environmental research and public health*, 19(3):1749. pages 3

Rosenthal, S. S., Strange, W. C., and Urrego, J. A. (2021). Jue insight: Are city centers losing their appeal? commercial real estate, urban spatial structure, and covid-19. *Journal of Urban Economics*, pages 103–381. pages 3

Sheng, J., Malani, A., Goel, A., and Botla, P. (2022). Jue insights: Does mobility explain why slums were hit harder by covid-19 in mumbai, india? *Journal of Urban Economics*, 127:103357. pages 3

Silva, M. T. L. d. (2023). Efeito das firmas sobre o contágio da covid-19 em recife. Master's thesis, Universidade Federal de Pernambuco - UFPE. pages 11

A Estimations via instrumented linear regression model (2SLS)

The results of the linear regression model with instrumental variables (2SLS) indicate that the three key variables have positive and statistically significant coefficients across all specifications. The distance to the nearest healthcare facility is positively associated with

COVID-19 contagion, suggesting that greater distances increase the risk of infection. The Floor Area Ratio (FAR), which represents building density, also shows a positive effect, indicating that areas with higher verticalization and crowding have a greater likelihood of contagion. Additionally, residing in Special Zones of Social Interest (Zeis), characterized by socioeconomic vulnerability and poor infrastructure, is associated with a significant increase in infection risk. The first-stage F-tests show high and significant values, confirming the strength and relevance of the instruments used to address endogeneity. Endogeneity tests (Durbin and Wu-Hausman) indicate that the explanatory variables are indeed endogenous, justifying the use of the instrumental variables model for more consistent and reliable estimates. Thus, the results reinforce the importance of urban infrastructure characteristics and social conditions in the dynamics of COVID-19 contagion.

Table 8: Estimations via instrumented linear regression model (2SLS)

Contagion	**1**	**2**	**3**	**4**
Characteristics of urban infrastructure				
Distance	0.281***	0.379***	0.598***	0.610***
	(0.047)	(0.055)	(0.101)	(0.103)
Far	0.216***	0.300***	0.436***	0.446***
	(0.058)	(0.022)	(0.044)	(0.046)
Zeis	0.476***	0.605***	0.649***	0.663***
	(0.058)	(0.071)	(0.090)	(0.092)
Neighborhood infrastructure	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes
Climate	No	Yes	Yes	Yes
Political position	No	No	Yes	Yes
Instruments	Yes	Yes	Yes	Yes
Fixed Effects				
FE epidemiological week	No	No	No	Yes
FE green area	Yes	Yes	Yes	Yes
FE conservation area	Yes	Yes	Yes	Yes
FE tests by individuals	Yes	Yes	Yes	Yes
Test F first stage				
Old trails	1675.11***	1353.26***	1695.38***	1654.06***
Household density 2000	2049.91***	1714.52***	1787.20***	1744.10***
Proportion of zeis by census sector	674.95***	468.08***	480.58***	467.26***
Tests of endogeneity				
Durbin (score) chi2(3)	250.78***	263.71***	198.695***	70.14***
Wu-Hausman F(3,63937)	83.89***	88.22***	66.3976***	23.39***
Observations	**158,189**	**158,189**	**158,189**	**158,189**

Notes: (***) statistically significant at least 1%, (**) statistically significant at least 5% and (*) statistically significant at least 10%.
Source: Author's own elaboration.