ARTICLE IN PRESS

Cities xxx (xxxx) xxx

Contents lists available at ScienceDirect

Cities

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

Street pedestrianization in urban districts: Economic impacts in Spanish cities

Yuji Yoshimura [a,*], Yusuke Kumakoshi [a], Yichun Fan [b], Sebastiano Milardo [c], Hideki Koizumi [a], Paolo Santi [c,d], Juan Murillo Arias [e], Siqi Zheng [b], Carlo Ratti [c]

[a] Research Center for Advanced Science and Technology, the University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
[b] Department of Urban Studies and Planning, Center for Real Estate and Sustainable Urbanization Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[c] SENSEable City Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[d] Istituto di Informatica e Telematica, CNR, via Giuseppe Moruzzi, 1, 56124 Pisa, Italy
[e] BBVA (Data Strategy area), c/Sauceda 28, Madrid 28050, Spain

ARTICLE INFO

Keywords:
Pedestrianization
Urban morphology
Street network
Transaction data
Consumer city

ABSTRACT

This study analyzes the influence of pedestrianization of urban space on the revenues of surrounding retail stores. Pedestrianization refers to the conversion of street use from vehicles to a walkable environment. We compiled a unique transaction dataset containing the estimates of sales volumes for stores across Spain and combine it with data from Open Street Map to provide the history of land-use changes at the street-level. Based on these high-granular datasets, we apply a difference-in-differences empirical method to measure the economic impact of pedestrian intervention. The results show that stores located in pedestrian environments tend to record higher sales volumes than stores located in non-pedestrian environments. We further analyze the mechanisms underlying this revenue-boosting effect and find that a key factor is the store density of the pedestrianized place, while geographic location is insignificant. This finding suggests that there are no differentiation impacts on stores' revenue based on whether pedestrianization occurs in the city center or periphery. Store category also acts as an important moderator for revenue impacts, with positive effects observed mostly for the café or restaurant category. Our results provide suggestive evidence that people prefer a pedestrian-friendly environment to a vehicle-oriented one for non-tradable, local consumption activities. This research provides evidence-based policy implications for urban planners interested in making smart, sustainable cities.

1. Introduction

For several decades, cities have been promoting restricted access of vehicles to the city center and transforming these spaces for pedestrian use (Newman & Kenworthy, 1989a, 1989, 2015). Regarding the recovery of public spaces, scholars argue that the land-use segregation of traditional Euclidian zoning and modernism-based development destroys traditional urban spaces (García Espuche, 1999; Jacobs, 1961). It has also been observed that vehicle-dependent city planning significantly increases environmental pollution (i.e., air pollution, noise, and anthropogenic heat), which affects human health (for a review, see Nieuwenhuijsen, 2020) and contributes to global warming (Ewing & Cervero, 2010). The unexpected COVID-19 pandemic and consequent lockdowns require us to reconsider the urban planning model (for a review, see Honey-Roses et al., 2020). We have been reminded that streets are more than optimized vehicle-based traffic spaces and of the simple pleasure of the natural human activity of walking (Solnit, 2001; Wilkinson, 2016). Some cities (e.g., Milan, Paris) have allowed restaurants, cafés, and bars to use streets to promote social distancing. Indeed, some cities (e.g., New York, Barcelona) have announced the intention to maintain their anti-pollution and anti-congestion measures post-COVID-19. The concerted goal is to develop a more human-centered concept of public spaces and pedestrian policies (Gehl & Svarre, 2013; Mueller et al., 2020; Sadik-Khan & Solomonow, 2016).

* Corresponding author.
E-mail addresses: yyyoshimura@cd.t.u-tokyo.ac.jp (Y. Yoshimura), ykuma@cd.t.u-tokyo.ac.jp (Y. Kumakoshi), ycfan@mit.edu (Y. Fan), milardo@mit.edu (S. Milardo), hide@cd.t.u-tokyo.ac.jp (H. Koizumi), psanti@mit.edu (P. Santi), juanmurillo@eoi.es (J. Murillo Arias), sqzheng@mit.edu (S. Zheng), ratti@mit.edu (C. Ratti).

https://doi.org/10.1016/j.cities.2021.103468
Received 10 September 2020; Received in revised form 31 May 2021; Accepted 15 September 2021

Despite this growing trend, scant research has attempted to quantify the impact of pedestrianization on retail activities based on empirical data. Although small-scale retailers are assumed to suffer revenue loss due to reduced accessibility after vehicles are banned, previous studies on the topic are largely descriptive and qualitative (Gehl, 1987; see Gehrke & Clifton, 2019 for a review; Jacobs, 1961; Lynch, 1960; Whyte, 1980). Within the limited body of quantitative studies, the popularity of streets, potential rents for commercial real estate, or expected revenues are proxied by geographical proximity to the city center (Hillier, 1996; Hillier et al., 1993; Porta et al., 2009, 2012; Sevtsuk, 2014), not on actual transaction data at the micro level. Furthermore, there is no well-established analytical framework to separate pedestrianization policy from other confounding environmental factors.

This study attempts to fill some of these gaps by quantifying the economic impact of pedestrianization on retail stores at the neighborhood level. We hypothesize that pedestrian-friendly spaces improve shopping experience and thus attract more customers for retailers, resulting in increased sales volumes. By examining the mechanisms and rules behind spatial agglomeration in urban economics, walking activity has been found to increase random encounters of stores compared to purposeful purchase, thereby increasing the effects of spatial agglomeration, as anchoring stores can play a larger role. Most empirical studies use economic indicators at coarse geographical scale, such as at the city, regional, national, and international levels (Fujita & Thisse, 2013). Few studies combine spatial configurations and land-use policies at neighborhood/street resolution to provide intra-city planning knowledge, owing to the difficulty of obtaining fine-grained data on revenues generated by individual retail stores (for a review, see Yoshimura et al., 2020).

To test these theoretical insights, we employ a novel card transaction dataset from a major Spanish bank to estimate the sales volumes for almost all cities in Spain with high spatial granularity. We also obtain the history of land-use changes from the Open Street Map (OSM) dataset. As retail store sales volumes are influenced by various factors, we adopt a difference-in-differences (DID) approach to infer the causal effect of street pedestrianization on the sales volume of stores in the neighborhood. DID models have been widely used to assess the effects of planning policies, such as light transit rail on the neighborhood socio-economic situations (Bardaka et al., 2018; Diao et al., 2017; Hurst & West, 2014), high-speed rail on the regional economy (Albalate et al., 2017; Wetwitoo & Kato, 2019), and urban redevelopment on safety (Branas et al., 2011) and on housing prices (Lee et al., 2017). This approach is sometimes combined with propensity score matching (PSM) (Caliendo & Kopeinig, 2008; Stuart, 2010; Stuart et al., 2014) to better deal with confounding factors. Despite the prolific literature applying these econometric approaches in the urban setting, ours is the first study to apply this quasi-experimental design to measure the economic impact of pedestrianization on retail stores.

2. Methodology

Our empirical strategy consists of three steps: First, we divided the entire city of each 14 Spanish cities into grids (128 m × 128 m) and calculated the increase in the pedestrian area inside each grid in the observation period (from November 1, 2010 to December 31, 2012). Second, we assigned grids in which the pedestrian areas were expanded during our observation period to the treatment group, and grids whose pedestrian areas remained stable to the control group. Third, we employed the propensity-score-matching-difference-in-differences (PSM-DID) method (Wang et al., 2019) to identify the causal effect. We removed geographically proximate grids of the treated units from the control group, as the stable unit treatment value assumption of the quasi-experimental design requires no inference (Rubin, 1980).

2.1. Data description

2.1.1. Transaction dataset

The transaction dataset consists of bankcard transactions provided by Banco Bilbao Vizcaya Argentaria (BBVA), one of Spain's largest banks, from November 1, 2010 to December 31, 2012. These include transactions performed by two groups of card users: the bank's direct customers, who hold a debit or credit card issued by BBVA, and other banks' customers, who performed transactions through any of the approximately 300,000 BBVA point-of-sale terminals for the whole country. The system registers customer transactions using debit or credit cards. The information includes a unique ID for each customer's credit or debit card that, for privacy reasons, was not associated with the real ID of the customer or the bank card, the shop where the customer made the transaction, the timestamp of the transaction, and the amount of money spent. The data were aggregated and anonymized in accordance with all Spanish privacy-protection laws and regulations.

The shops in the original dataset have been classified into seven general categories with 17 sub-categories based on the primary and secondary activities in the city (for all categories, see Table 1 in Yoshimura et al., 2018). Based on this re-disaggregated process, this study chose (1) stores that opened during the study period (successive store) and (2) the aggregated categories A and B (Fig. 1a). Fig. 1b presents the boxplot of those stores' sales volumes in each month, and shows flat sales for the entire period. Fig. 1c shows that stores satisfying condition (1) above produce 75.5% of the total sales volume on average of all stores, while representing just 34.5% of all registered stores in Spain. All the results validated the representativeness of the dataset for our analysis.

2.1.2. Study area

We chose 14 Spanish cities for our analysis. They were selected according to the following criteria. Although the transaction dataset covers the whole of Spain, small-scale cities have fewer point-of-sales terminals in the city, resulting in a smaller number of conducted transactions. To reduce measurement errors, we first identified the 50 most populated cities in Spain. Then, we attempted to collect pedestrian space information for these cities through OSM, using the method presented in the next section. Some of these 50 cities lacked high-quality data for pedestrian space, leaving 14 cities in our final sample. The number of shops located in each city and their basic statistics are in Table S1 in the Supplementary Materials.

2.1.3. Pedestrian space dataset

The OSM dataset contains the metadata for each street, providing the history of land-use changes. We collected all necessary information for the same time period corresponding to the transaction dataset through Ohsome API service (Grinberger et al., 2019). This platform was developed by the Big Spatial Data Analytics team at HeiGIT. It allowed us to access the full history of OSM easily. To retrieve the pedestrian area over time, we collected information on the boundaries of the urban areas of the 32 most populated Spanish cities using OSMnx (Boeing, 2017). Then, we divided them on the 128 m × 128 m grid. These grids, which are used as the input for the Ohsome API, returned the historical data of the pedestrian areas for each grid.

Fig. 2a presents the distribution of the average increment in the number of pedestrianized grids from the previous month for 14 cities. The figure shows an increase in the pedestrian grid, which was concentrated during May 2011 to December 2011. We also analyzed the geographical locations where pedestrianization occurred in each city and compared them for the 14 cities (Fig. 2b).

2.2. Characterization of grids: Link-betweenness centrality

Estimation of geographical accessibility requires a specific calculation and stratification. For this purpose, we employed the local link-

Category	Proportion of shops (%)	No. of classifications of shops	Activity	Description
A1	23.7	14 types (e.g., meat shop, fish shop)	Convenience goods	The ordinary needs of the neighborhood (e.g., vegetable shop)
A2	7.6	7 types (e.g., furniture, home decoration)	Household equipment	More oriented to specialized economic activities than A1
A3	23.3	7 types (e.g., jewelry shop, perfume shop)	Personal equipment	
A4	8.2	5 types (e.g., gas station, car dealer)	Sell, fix, and maintenance motor vehicleand fuel	Related to motor vehicles
A5	5.6	13 types (e.g., kiosk, souvenirs)	Others(goods)	Other types goods (e.g., kiosk)
A6	12.9	7 types (e.g., hairsalon, travel agency)	Others(services)	Other types goods (e.g., travel agency)
B1	16.1	5 types (e.g., restaurant, cafe)	Restaurant, cafe,pub	Restaurant,cafe,pub B3+B4
B2	2.2	1 type (e.g., hotels)	Hotel, B&B, hostel	Hotels
B3	13.1	1 type (e.g., restaurant)	Restaurant	Restaurant
B4	3.0	3 types (e.g., cafe, pub)	Cafe,pub	Cafe,pub
A	81.3	53 types	Retail Commerce	A1+A2+A3+A4+A5+A6
B	18.3	6 types	Hotel, B&B, hostel, Restaurant,cafe,pub	B2+B3+B4

(a)

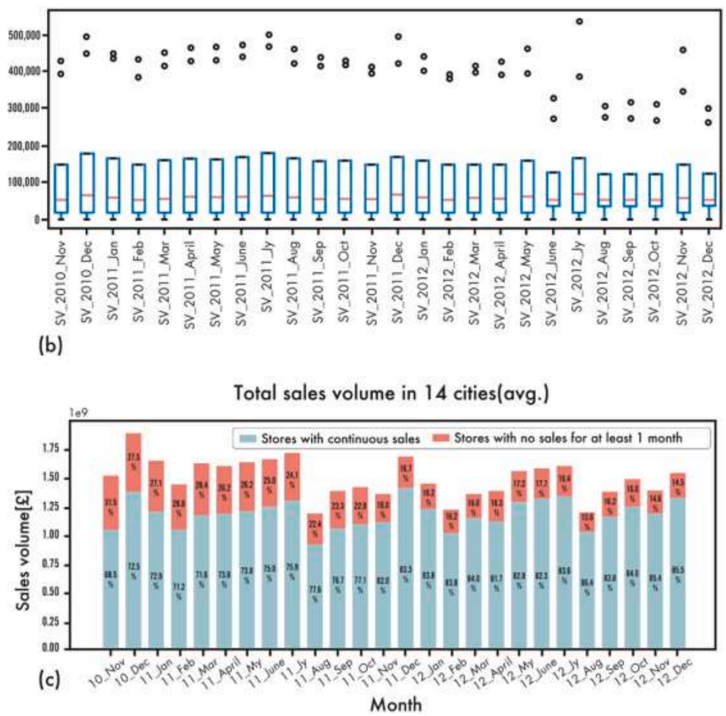


Fig. 1. (a) Basic properties for individual categories. The author of this paper made this table based on Fig. S2 in Yoshimura et al. (2020), using the employed dataset for this study. (b) Boxplot of stores' sales volumes in each month for all stores in Spain. (c) Distribution of average total sales volumes generated by two kinds of stores: stores with continuous sales and stores with no sales for at least one month during the observation period.

betweenness centrality indicator to classify the physical proximity of the grids (Yoshimura et al., 2020).

Link-betweenness centrality measures how frequently link *i* acts as the mediator on the shortest paths between the other pairs of links in the network. A link with a higher degree has a higher performance—that is, it is traversed in a relatively high number of shortest paths in the network. An intermediate link acts as a strategic control point. To compute betweenness, we considered only the shortest paths between streets up to a certain distance from the link of interest. We denoted link *i*'s betweenness by $C_i^{Br}$when we considered network distance *r*, defined in this study as the radius of the total length of the streets from the targeted link *i*, rather than the Euclidean distance from the link *i* ($r =$ 300 m, 400 m 500 m … 5000 m, increasing in units of 100 m). To determine C_i^{Br}, we computed the centrality indicator for each link, considering the street network from link *i*.

In this study, we divided all grids into three groups depending on the betweenness values. We labeled the first bin the "historical center cluster," the second bin the "artery cluster," and the last bin the "other cluster." Finally, we applied PSM-DID for each group.

2.3. Propensity score matching–difference-in-differences

PSM–DID reduces confounding factors by matching each treated unit with their comparable controls, comparing the same unit before and after to control for unobserved individual characteristics, and comparing treatment and control units to control for general time trends (Angrist & Pischke, 2008; Caliendo & Kopeinig, 2008).

The first step was to implement PSM to ensure balance between the treatment and control groups so that the parallel trend assumption could be satisfied. Compared to conventional matching algorithms, PSM enabled us to align the distribution of the multi-dimensional covariances matrix between the treatment and control groups by converting the feature vectors into a scalar quantity (Rosenbaum & Rubin, 1983). The propensity score was defined as the prediction probability of a specific individual being treated, given the individual's covariances:

$$f(X_i) = P(C_i = 1|X_i) \tag{1}$$

where $f(\bullet)$ is a propensity score function (logit function), C_i is a dummy variable representing the treatment condition of zone *i*, and X_i is a vector of control variables for zone *i*. The underlying assumption of PSM is that, conditional on the observable characteristics affecting treatment probability, the treatment status is nearly randomly allocated. We used nearest neighbor matching to select the treatment and control groups on a one-to-one basis based on their propensity scores.

Based on the matched treatment-control pairs, our second step was to implement DID estimation. Fig. 3a presents how standard DID is operationalized (Angrist & Pischke, 2008). The common trend assumption (CTA) of DID requires that the time trend of the treatment group be parallel to that of the control group if the intervention was not conducted (Abadie & Cattaneo, 2018). In this study, we defined the area of pedestrianization during May 1, 2011-July 31, 2011 as the treatment group. We excluded these three months of sales volumes from the regression analysis and tested the parallel trend using the following event study approach:

$$\mathrm{Iny}_{it} = \alpha_0 + \alpha_1 G_i + \sum_{i=-6, j\neq -1}^{16} \beta^j G_i \times m_{t,j} + X_i\gamma + \mu_m + \epsilon_{it} \tag{2}$$

where y_{it} is the total sales volume of grid *i* on date *t*; G_i is a treatment dummy that equals 1 if grid *i* is treated and 0 otherwise; and $m_{t,\,j}$ is a time dummy that equals 1 if date *t* is within month *j* (we recoded the treatment time period as 0). The dummy for $j = -1$ time period (i.e., April 1, 2011-April 30, 2011) was omitted in the regression to be used as the baseline. X_i is a rich vector of covariates for grid *i*. We also included month fixed effect μ_m to control for the place-invariant confounders specific to each month. ϵ_{it} is the robust standard error. Our coefficient of interest is β^j, which captures the average treatment effect of pedestrianization on sales volume in month *j*.

With the CTA hypothesis satisfied, we estimated the influence of pedestrianization on the total sales volumes as follows:

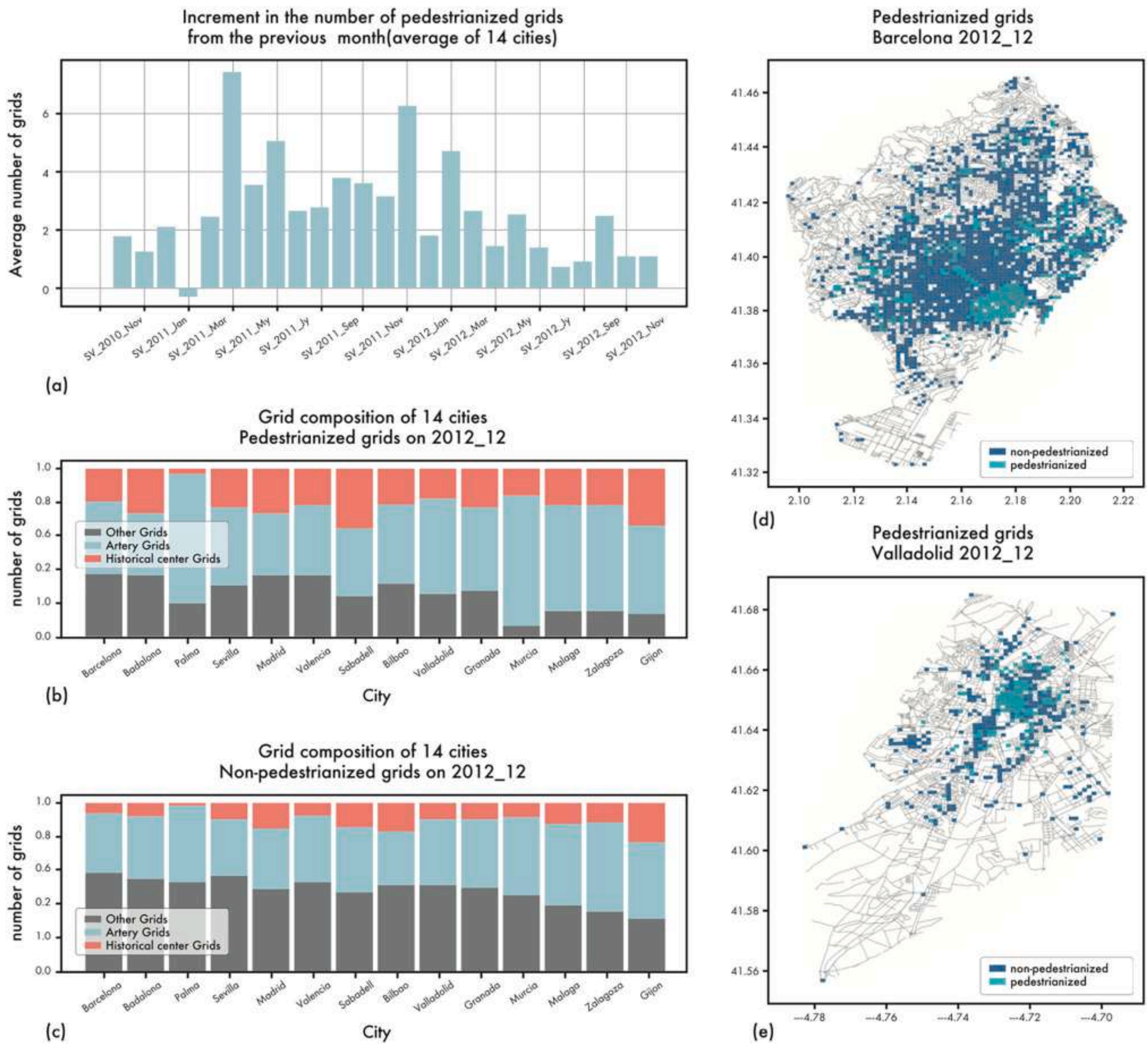


Fig. 2. (a) The distribution of the average of the increment in the number of pedestrianized grids from the previous month for 14 Spanish cities. (b) The geographical distribution of the pedestrian grids of 14 cities. (c) The geographical distribution of the non-pedestrian grids of 14 cities. (d) The visualization of the geolocation of the pedestrian and non-pedestrian grids for Barcelona, December 2012. (e) The visualization of the geolocation of the pedestrian and non-pedestrian grids for Valladolid, December 2012.

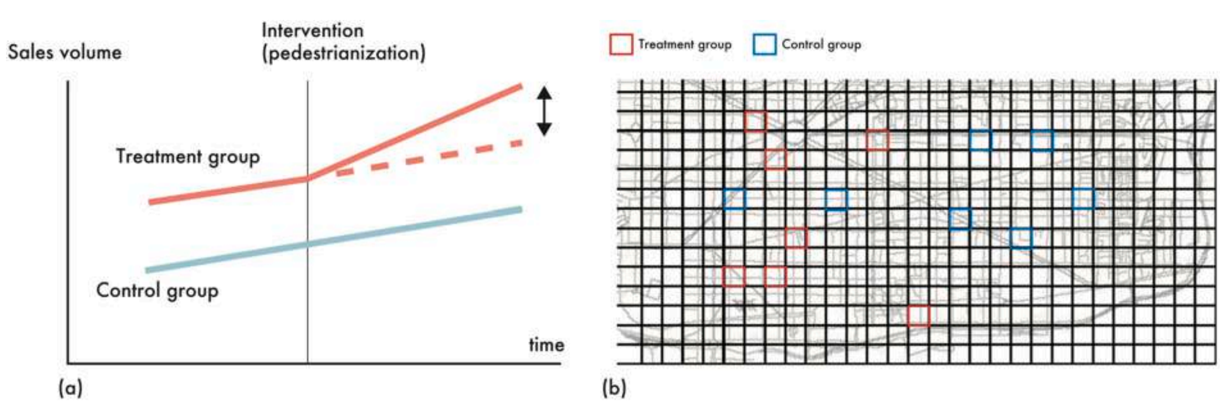


Fig. 3. (a) Schematic illustration of the economic trend and its parallel condition for DID application. The author of this paper made this illustration. (b) Flow chart for the subclassified PSM.

$$\alpha = \Delta\text{treat} - \Delta\text{control} \quad (3)$$

where $\Delta treat$ = (median of total sales volume of the treatment group after pedestrianization) – (median of total sales volume of the treatment group before pedestrianization), and $\Delta control$ = (median of total sales volume of the control group after pedestrianization) – (median of total sales volume of the control group before pedestrianization). Consequently, α provides the impact of the pedestrian policy. We used the median, instead of the average, to reduce the impacts of extreme values.

To ensure balance on key confounding variables, we employed stratification before PSM-DID based on store density, monthly sales volumes, geographic accessibility, and store category.

In this study, we hypothesized that four factors affect total sales volumes in the grid: (1) store density, (2) average monthly sales volumes of the pre-intervention period, (3) geographic accessibility, and (4) store category. Therefore, we stratified these four factors and chose the treatment and control grids within the same quartile of each group (Fig. 3c). Specifically, we first divided all grids into quartiles by their features. Second, we stratified the grids in the treatment group into quartiles, calculated the range of values in each quartile, and matched the control grids in the same range into the same stratum. This provided us with grids classified into 16 layers. Third, we selected all grids that were classified into the same layer as the treatment group, and we classified them into the control group. Finally, we chose the treatment and control groups from each layer of each feature and then conducted the DID analysis.

3. Results and discussion

This section provides the results of our analysis by the distribution of the stores' median sales volumes between the pedestrianized and non-pedestrian grids for 14 Spanish cities. Although the economic trend followed a similar pattern for both groups, the stores located in pedestrian grids generated higher sales volumes than those located in non-pedestrian grids (see Fig. 4a for all cities). In addition, we observed that the stores' revenues dropped drastically during the summer vacation period (i.e., August), while they significantly increased at the end of the year (i.e., December).

To verify whether such a difference is statistically significant, we applied Welch's *t*-test (Derrick et al., 2016) for the 14 cities. Fig. 4b presents the number of months by the score of Welch's *p*-value for the 14 cities, showing that during more than half of the months for 10 cities, *p*-values <0.05. This indicates that the stores located in pedestrian grids could generate higher revenues than those located in non-pedestrian grids. We also applied the Cohen's *d* indicator (Kelley & Preacher, 2012) for the 14 cities to evaluate the effect size of the difference in store sales volumes between the pedestrian and non-pedestrian grids. Fig. 4c shows that there is a tendency for small- to medium-sized cities to have a stronger effect than large cities.

Furthermore, we examined the distribution of *p*-values of Welch's *t*-test for each category for each month. Fig. 5a shows six cities that are statistically significant for category A, while there are only three cities in category B (see Fig. 5b). This indicates that the category-A stores located in the pedestrian grids in those six cities tend to produce higher sales volumes than the stores in non-pedestrian grids, and the difference is statistically significant.

These facts suggest that, from the longer-term point of view, the stores located in pedestrian grids tend to produce higher sales volumes than those located in non-pedestrian grids. Our outcome can be considered a result of land-use change after streets were converted to a pedestrian environment. We also discovered that the difference in stores' sales volumes is larger in small- to medium-sized cities than in larger cities. In addition, category-A stores located in pedestrian environments recorded higher sales volumes than did Category-A stores in non-pedestrian environments, and the effect size is larger in these cities than in large cities.

3.1. Preparation for difference-in-difference analysis: CTA and PSM

The previous results show correlations between change in land use (pedestrianization) and sales volume. It is possible that other confounding factors co-determine the possibility for a grid to be pedestrianized and to simultaneously generate higher sales volumes. To better understand the causal implications of pedestrianization, we proceeded to the DID analysis. First, we checked whether the dataset of each city satisfied the CTA and PSM. The cities that satisfied those conditions were included in the DID analysis for the next step.

To ensure that the selected grids shared similar characteristics, we examined whether the distribution of the selected factors was similar enough among the grids. (see "Methodology" for details).

Fig. 6a presents the randomly selected distribution of the monthly sales volumes for four different situations for different cities. The blue line is the estimated coefficient of the difference of the outcome between the treatment and control groups, and the blue dotted line is the confidence interval. We interpreted the result to mean that the distribution of the two groups is similar enough if the baseline (=o) is sandwiched by the confidence interval. In this way, we could check the distribution and confirmed that all situations from all cities satisfy the condition of CTA (see Fig. S5 in the Supplementary Materials for all cases).

The next step is to check PSM. Fig. S6 presents the randomly selected PSM for four cities (see "Methodology" for details). We tested the common support with a balance test (Bruhn & McKenzie, 2009). Fig. S6a and S6b show that those two units are similarly distributed, which satisfies the condition for inclusion in the DID analysis. Conversely, Fig. S6c and S6d do not show similar distribution, indicating these units do not satisfy the conditions for inclusion in the DID analysis. After checking the PSM for all cities, we discovered that Madrid, Barcelona, and Valencia satisfied the conditions, but the other cities did not. Consequently, we focused on these three cities for the DID analysis.

3.2. Stratification

We now advance to the stratification for the DID analysis, as the previous subsection uncovers an overall null effect, but it would

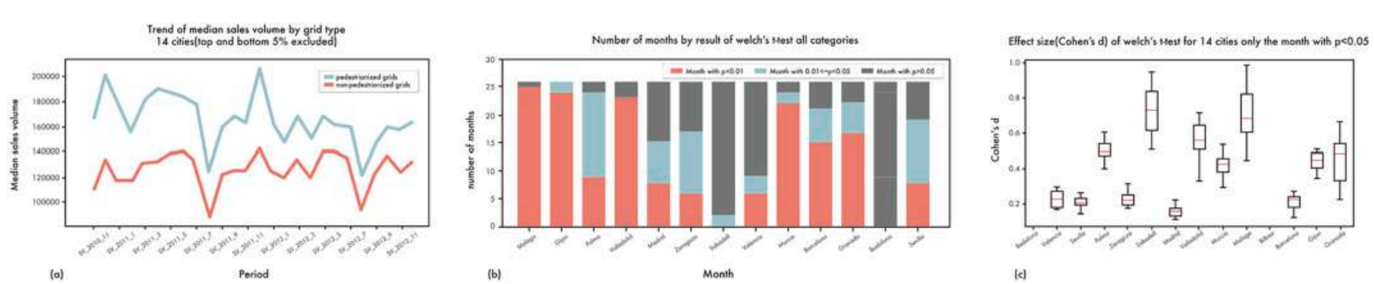


Fig. 4. (a) Revenues of stores located in the pedestrian and non-pedestrian grids. (b) Proportional distribution of *p*-values of Welch's *t*-test for stores' sales volumes in the pedestrian and non-pedestrian grids. (c) Boxplots of the effect size (Cohen's *d*) of Welch's *t*-test for 14 cities.

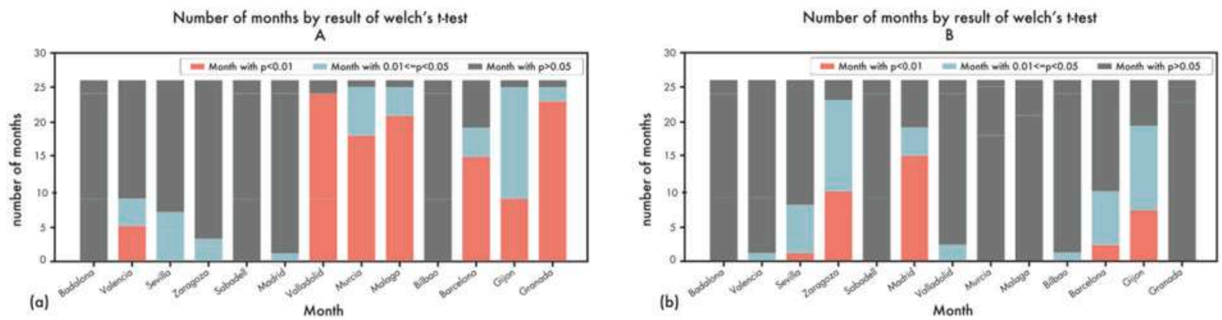


Fig. 5. (a) Number of months by the result of Welch's *t*-test for category A. (b) Number of months by the result of Welch's t-test for category B.

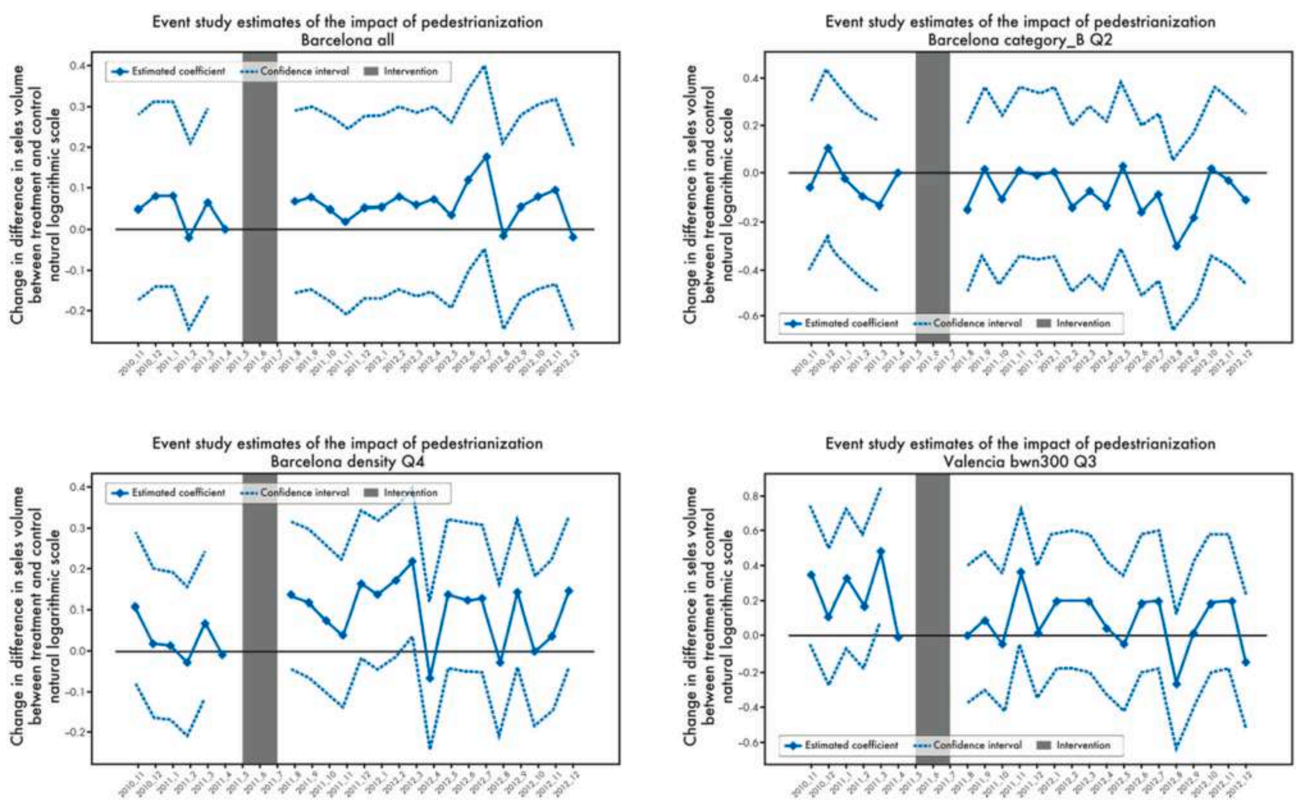


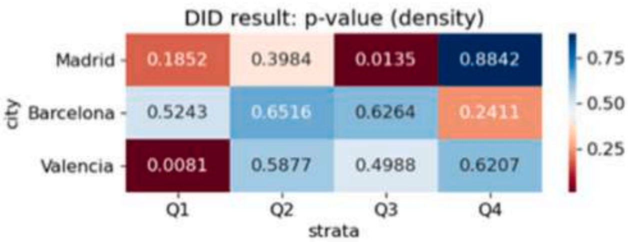
Fig. 6. (a) Distribution of monthly sales volume in Barcelona. (b) Distribution of monthly sales volume of store density (Q4) in Barcelona. (c) Distribution of monthly sales volume of category B (Q2) in Barcelona. (d) Distribution of monthly sales volume betweenness 300 m (Q3) in Valencia.

probably mask the effect of the underlying heterogeneity. Thus, we focused on some factors and implemented DID again for the subgroup. This subsection presents the results. The four factors considered most influential were further classified into quantiles depending on their scores. These factors are: (1) density of stores; (2) sales volumes before pedestrianization; (3) betweenness 300 m; (4) betweenness 5000 m; (5) geographic characteristics (three categories); (6) store category A; and (7) store category B. Consequently, we developed 27 models (6 variables $\times$ 4 + 1 variable $\times$ 3 stratifications), applying them to the cases of Madrid, Barcelona, and Valencia. This method enabled us to examine the influence of each variable in more detail. We present only the density of stores and category B, which show statistically significant results. Other results are presented in Figs. S7–S14 in the Supplementary Materials.

3.2.1. *Density of stores*

We observe in Fig. 7 that Q3 in Madrid and Q1 in Valencia have *p*-values <0.05, which indicates statistical significance. However, the coefficients of the interaction terms have negative signs for both cases. This indicates that pedestrianization causes, if anything, a decrease in stores' sales volumes. In the case of Madrid, the higher-store-density environment across the city tends to generate lower sales volumes by pedestrianization than before the intervention. In the case of Valencia, the lower-density environment tends to generate lower sales volumes by pedestrianization than before the intervention.

Conversely, betweenness 300 m and 5000 m do not show statistically significant results (see the Supplementary Materials). This suggests that geographic location is not a significant factor for pedestrianization. Neither the central locations in terms of the betweenness indicator nor the locations that are clustered near an artery, historic center, or other environment show *p*-values <0.05. This suggests that if we expect an economic impact, the location of pedestrianization is not necessarily considered from the geographic viewpoint, but from that of other factors.



Barcelona, Store density				
Variable	Q1	Q2	Q3	Q4
Intercept	10.243 (0.164)***	11.201 (0.234)***	13.05 (0.285)***	12.272 (0.123)***
G	0.333 (0.108)***	0.372 (0.108)***	-0.25 (0.099)**	0.038 (0.077)
T	0.109 (0.193)	0.157 (0.193)	0.131 (0.169)	0.158 (0.137)
G∘T	**-0.078 (0.123)**	**-0.056 (0.123)**	**0.052 (0.107)**	**0.102 (0.087)**
Adj.R2	0.202	0.244	0.199	0.387
N	966	1150	874	690

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Madrid, Store density				
Variable	Q1	Q2	Q3	Q4
Intercept	11.418 (0.228)***	-9.888 (1.173)***	10.633 (0.247)***	11.446 (0.233)***
G	-0.071 (0.122)	3.042 (0.206)***	1.361 (0.099)***	0.374 (0.087)***
T	0.202 (0.275)	-0.405 (0.241)*	0.49 (0.179)***	0.158 (0.182)
G∘T	**0.206 (0.156)**	**-0.116 (0.136)**	**-0.252 (0.101)****	**-0.015 (0.103)**
Adj.R2	0.161	0.828	0.812	0.66
N	782	184	322	92

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Valencia, Store density				
Variable	Q1	Q2	Q3	Q4
Intercept	11.006 (0.218)***	9.744 (0.196)***	1.469 (0.764)*	11.065 (0.324)***
G	-0.056 (0.137)	0.409 (0.1)***	0.071 (0.109)	0.149 (0.152)
T	0.407 (0.241)*	0.159 (0.176)	0.116 (0.163)	0.117 (0.14)
G∘T	**-0.407 (0.153)*****	**-0.061 (0.112)**	**-0.07 (0.104)**	**0.044 (0.089)**
Adj.R2	0.383	0.341	0.748	0.627
N	736	920	368	322

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Fig. 7. Distribution of the *p*-value of difference-in-differences in terms of store density.

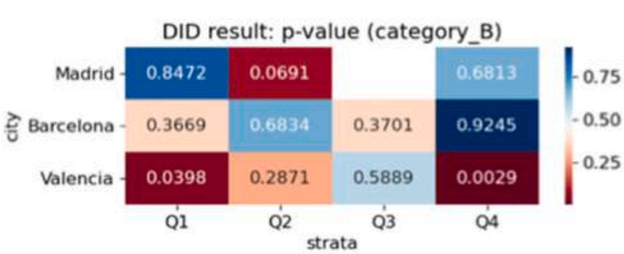
3.2.2. *Category B*

Conversely, Fig. 8 presents the *p*-value of DID for the category-B stores. In Valencia, Q4 and Q1 show *p*-values <0.05, indicating that this factor is statistically significant. Q4 is positive and Q1 is negative, meaning that the higher-density environment of store category A in Valencia caused an increase in sales volumes by pedestrian intervention, while the lower-density environment caused sales volumes to decrease.

We speculated that a higher density of cafés and restaurants attracts more people to the pedestrian environment than to the non-pedestrian environment. People may prefer to have coffee, lunch, or dinner in the pedestrian environment than in the non-pedestrian environment. In addition, this phenomenon occurs in the higher-density environment rather than in the lower-density environment. This result suggests, on the one hand, that a lively environment created by the crowdedness of restaurants and cafés attracts more people, resulting in increased sales volumes. On the other hand, the result also suggests that in the lower-density locations, pedestrian intervention decreases revenues for restaurants and cafés in Valencia. Crowdedness could be important for stores to attract people, and thus, to increase their revenues.

3.3. *Summary*

All of these results revealed that store density is one of the significant moderators affecting the treatment effect of pedestrianization on sales volume. In addition, we discovered that store category is important for pedestrianization. For example, considering category B in Valencia, the higher-density environment caused an increase in store revenue by pedestrian intervention, while neither the higher- nor lower-density environment of store category A showed a *p*-value <0.05. These results indicate that density is one of the deterministic factors for the impacts of pedestrianization on stores' sales volumes, but it largely depends on the store category.



Barcelona, Store category B				
Variable	Q1	Q2	Q3	Q4
Intercept	10.369 (0.136)***	10.868 (0.168)***	8.284 (0.65)***	12.0 (0.138)***
G	0.154 (0.097)	0.176 (0.105)*	-1.513 (0.197)***	0.015 (0.101)
T	0.059 (0.176)	0.213 (0.185)	0.18 (0.168)	0.367 (0.138)***
G∘T	**-0.101 (0.112)**	**-0.048 (0.118)**	**0.096 (0.107)**	**0.008 (0.088)**
Adj.R2	0.241	0.283	0.775	0.804
N	1426	1380	322	414

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Madrid, Store category B				
Variable	Q1	Q2	Q3	Q4
Intercept	10.288 (0.225)***	10.275 (0.236)***	-	11.641 (0.267)***
G	0.669 (0.127)***	-0.844 (0.113)***	-	0.902 (0.164)***
T	-0.38 (0.287)	0.291 (0.229)	-	0.136 (0.224)
G∘T	**0.031 (0.162)**	**0.237 (0.13)***	**-**	**-0.052 (0.127)**
Adj.R2	0.322	0.457	-	0.765
N	552	506	-	230

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Valencia, Store category B				
Variable	Q1	Q2	Q3	Q4
Intercept	10.795 (0.167)***	10.37 (0.193)***	11.162 (0.219)***	10.803 (0.136)***
G	-0.152 (0.119)	0.807 (0.125)***	-0.482 (0.155)***	-0.374 (0.117)***
T	0.217 (0.215)	0.242 (0.212)	0.225 (0.217)	-0.021 (0.134)
G∘T	**-0.282 (0.137)****	**-0.144 (0.135)**	**0.075 (0.138)**	**0.255 (0.085)*****
Adj.R2	0.181	0.374	0.534	0.771
N	1058	598	414	368

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Note: All models include control variables. Standard diviations are inside parentheses.
Dependent variable is ln(Sales Volume)

Fig. 8. Distribution of *p*-value of difference-in-differences in terms of store category B.

We speculated that people tend to buy ordinary articles (category A) near their houses, and this tendency does not change if the environment is converted into pedestrian space. On the other hand, people may prefer pedestrian environments to have coffee, lunch, or dinner (category B) over non-pedestrian environments. In addition, we speculated that the agglomeration of stores creates an attractive environment, resulting in an increase in the number of customers. The overall result is increased sales volumes.

Finally, we found that geographic location, when controlling for density and store composition, has little effect on the economic impacts of pedestrianization. Neither betweenness 300 m nor 5000 m show *p*-values <0.05 (see "Methodology" and (Yoshimura et al., 2020) for the betweenness 300 m and 5000 m in more detail). These results indicate that spatial accessibility in terms of network centrality is not an important factor in triggering stores' higher sales volumes by pedestrianization, and it does not create an economic impact for the stores.

4. Conclusion

This study provides evidence-based policy implications for urban planners interested in developing attractive, resilient, and sustainable cities. Understanding the potential moderators for the economic impacts of pedestrianization is especially relevant for post-COVID-19 urban planning models. Due to the pandemic, many cities plan to expediate land-use changes to pedestrian-oriented space, but urban planners usually do not have a consistent evidence-based policy, even methodology, to choose the location, resulting in seemingly random choices over the city. The empirical evidence provided in this study on the role played by store density and store types can guide planners in prioritizing the pedestrianization of urban areas where more experience-based stores are concentrated. It is important to observe that store density, while apparently important to have a significant positive economic impact on pedestrianized areas, is not necessarily at odd with public health considerations in case of pandemic events, which tend to discourage the planning of dense urban environments. We believe that in this context it would be key to provide people with the opportunity of connecting with each other, while adequately controlling their face-to-face contact occasion even under the densified environment. Such planning and policies can achieve a vibrant and lively environment created by the store density while preventing the pandemic of the infection.

In addition, our results are helpful for policymakers to explain land-use change policy to retailers located along streets to be pedestrianized. Although pedestrianization can improve many aspects of a city, such a policy is not always welcome. For example, some retailers may wonder whether land-use change to pedestrian use would have a positive or negative impact on their revenues, which is a significant factor for their businesses. This critical information is a novel result of our study.

Although the proposed method and results provide clear value and novel perspectives that extend the existing research, there are several limitations of this study. First, we focus on a limited number of Spanish cities for the DID analysis, since the intrinsic differences between pedestrianized and non-pedestrianized areas in most cities make it difficult to develop valid counterfactual scenarios. Alternative quasi-experimental strategies or randomized controlled trials are needed to investigate a larger number of cities to extend the validity of our results. Second, we require a longer observation period to measure the mid- to long-term impacts of pedestrianization. Although we have already identified some patterns of economic impacts in the short term (i.e., two years), the major impacts of pedestrianization could emerge after a longer period of implementation. Finally, pedestrianization policy has broader social impacts than only boosting retail revenue. For example, the literature has found positive impacts of pedestrian spaces on people's psychological frame of mind (Solnit, 2001). Pedestrianization also helps reduce negative environmental effects, such as air pollution or urban noise (Nieuwenhuijsen, 2020). Considering these previous findings, we argue that urban space should be pedestrian oriented for environmental and psychological well-being, even if the economic impact is limited and context specific.

Pedestrian streets could transform the shopping experience from goal oriented (i.e., drive to destination, buy the planned items, and leave) to experience oriented (i.e., walk around, take a coffee break, and chat with friends). The latter would increase urban vibrancy and facilitate the urban agglomeration effect by increasing social interaction. In summary, pedestrianization does more than provide obvious environmental benefits. Our study differs from previous ones as it documents other potential social impacts, which is of particular value when considering the post-COVID-19 lifestyle evolution. The causal impacts of land-use planning on the local economy at the sub-city level have been largely understudied and should receive increased attention in future research.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Declaration of competing interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgements

We would like to thank the Banco Bilbao Vizcaya Argentaria (BBVA) for providing the data set for this study.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.cities.2021.103468.

References

Abadie, A., & Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics, 10*(1), 465–503.

Albalate, D., Campos, J., & Jiménez, J. L. (2017). Tourism and high speed rail in Spain: Does the AVE increase local visitors? *Annals of Tourism Research, 65*, 71–82. https://doi.org/10.1016/j.annals.2017.05.004

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics. An empiricist's companion.* Princeton university press.

Bardaka, E., Delgado, M. S., & Florax, R. J. G. M. (2018). Causal identification of transit-induced gentrification and spatial spillover effects: The case of the Denver light rail. *Journal of Transport Geography, 71*, 15–31. https://doi.org/10.1016/j.jtrangeo.2018.06.025

Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems, 65*, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Branas, C. C., Cheney, R. A., MacDonald, J. M., Tam, V. W., Jackson, T. D., & Ten Have, T. R. (2011). A difference-in-differences analysis of health, safety, and greening vacant urban space. *American Journal of Epidemiology, 174*(11), 1296–1306. https://doi.org/10.1093/aje/kwr273

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics, 1*(4), 200–232.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Derrick, B., Toher, D., & White, P. (2016). Why Welch's test is type I error robust. *Quantitative Methods for Psychology, 12*(1), 30–38.

Diao, M., Leonard, D., & Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics, 67*(May), 64–77. https://doi.org/10.1016/j.regsciurbeco.2017.08.006

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association, 76*(3). https://doi.org/10.1080/01944361003766766

Fujita, M., & Thisse, J. F. (2013). *Economics of agglomeration*. Cambridge: Cambridge University Press. https://doi.org/10.1006/jjie.1996.0021

García Espuche, A. (1999). *The reconquest of Europe. Urban public space, 1980–1999.* Barcelona: CCCB.

Gehl, J. (1987). *Life between buildings: Using public space*. Washington, DC: Island Press.
Gehl, J., & Svarre, B. (2013). *How to study public life*. Washington, DC: Island Press—Center for Resource Economics. https://doi.org/10.5822/978-1-61091-525-0
Gehrke, S. R., & Clifton, K. J. (2019). An activity-related land use mix construct and its connection to pedestrian travel. *Environment and Planning B: Urban Analytics and City Science, 46*(1), 9–26. https://doi.org/10.1177/2399808317690157
Grinberger, A. Y., Schoot, M., Raifer, M., Troilo, R., & Zipf, A. (2019). Analyzing the spatio-temporal patterns and impacts of large-scale data production events in OpenStreetMap. In M. Minghini, A. Y. Grinberger, L. Juhasz, G. Yeboah, & P. Mooney (Eds.), *Proceedings of the Academic Track at the State of the Map 2019. Heidelberg, Germany* (pp. 9–10).
Hillier, B. (1996). *Space is the machine: A configurational theory of architecture*. Cambridge: Cambridge University Press.
Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: Or, configuration and attraction in urban pedestrian movement. *Environment and Planning. B, Planning & Design, 20*(1), 29–66. https://doi.org/10.1068/b200029
Honey-Roses, J., Anguelovski, I., Chireh, V. K., Daher, C., Konijnendijk, C., Litt, J. S., Mawani, V., et al. (2020). The impact of COVID-19 on public space: An early review of the emerging questions—Design, perceptions and inequities. *Cities Health*. https://doi.org/10.1080/23748834.2020.1780074
Hurst, N. B., & West, S. E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics, 46*(1), 57–72. https://doi.org/10.1016/j.regsciurbeco.2014.02.002
Jacobs, J. (1961). *The death and life of great American cities*. New York: Random House.
Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods, 17*(2), 137–152.
Lee, C. C., Liang, C. M., & Chen, C. Y. (2017). The impact of urban renewal on neighborhood housing prices in Taipei: An application of the difference-in-difference method. *Journal of Housing and the Built Environment, 32*(3), 407–428. https://doi.org/10.1007/s10901-016-9518-1
Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.
Mueller, N., Rojas-Rueda, D., Khreis, H., Cirach, M., Andrés, D., Ballester, J., Bartoll, X., et al. (2020). Changing the urban design of cities for health: The superblock model. *Environment International, 134*. https://doi.org/10.1016/j.envint.2019.105132
Newman, P., & Kenworthy, J. (1989b). Gasoline consumption and cities: A comparison of US cities with a global survey. *Journal of the American Planning Association, 55*(1), 24–37. https://doi.org/10.1080/01944368908975398
Newman, P., & Kenworthy, J. (1989a). *Cities and automobile dependence: An international sourcebook*. https://doi.org/10.1016/0957-1787(91)90079-k
Newman, P., & Kenworthy, J. (2015). *The end of automobile dependence: How cities are moving beyond car-based planning*. Island Press—Center for Resource Economics. https://doi.org/10.5822/978-1-61091-613-4
Nieuwenhuijsen, M. J. (2020). Urban and transport planning pathways to carbon neutral, liveable and healthy cities: A review of the current evidence. *Environment International*. https://doi.org/10.1016/j.envint.2020.105661
Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., et al. (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies, 49*(7), 1471–1488. https://doi.org/10.1177/0042098011422570
Porta, S., Strano, E., Iacoviello, V., Messora, R., Latora, V., Cardillo, A., Wang, F., & Scellato, S. (2009). Street centrality and densities of retail and services in Bologna Italy. *Environment and Planning. B, Planning & Design, 36*(3), 450–465. https://doi.org/10.1068/b34098
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.
Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association, 75*(371), 591–593.
Sadik-Khan, J., & Solomonow, S. (2016). *Streetfight: Handbook for an urban revolution*. New York: Viking Penguin.
Sevtsuk, A. (2014). Location and agglomeration: The distribution of retail and food businesses in dense urban environments. *Journal of Planning Education and Research, 34*(4), 374–393. https://doi.org/10.1177/0739456X14550401
Solnit, R. (2001). *Wanderlust: A history of walking*. New York: Penguin Press.
Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*(1), 1–21. https://doi.org/10.1214/09-STS313
Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services & Outcomes Research Methodology, 14*(4), 166–182. https://doi.org/10.1007/s10742-014-0123-z
Wang, H., Chen, Z., Wu, X., & Nie, X. (2019). Can a carbon trading system promote the transformation of a low-carbon economy under the framework of the porter hypothesis?-Empirical analysis based on the PSM-DID method. *Energy Policy, 129*, 930–938.
Wetwitoo, J., & Kato, H. (2019). Regional and local economic effects from proximity of high-speed rail stations in Japan: Difference-in-differences and propensity score matching analysis. *Transportation Research Record*. https://doi.org/10.1177/0361198119844757
Whyte, W. H. (1980). *The social life of small urban spaces*. Washington, DC: The Conservation Foundation. https://doi.org/10.4324/9780203873960
Wilkinson, M. (2016). *Restless creatures: The story of life in ten movements*. New York: Basic Books.
Yoshimura, Y., Santi, P., Arias, J. M., Zheng, S., & Ratti, C. (2020). Spatial clustering: Influence of urban street networks on retail sales volumes. *Environment and Planning B: Urban Analytics and City Science, 48*(7), 1926–1942.
Yoshimura, Y., Sobolevsky, S., Bautista Hobin, J. N., Ratti, C., & Blat, J. (2018). Urban association rules: Uncovering linked trips for shopping behavior. *Environment and Planning B: Urban Analytics and City Science, 45*(2), 367–385. https://doi.org/10.1177/0265813516676487