The Differentiation and Spatial Sorting of Urban Amenities: Evidence from Madrid's Restaurants

April 2026

(Preliminary and incomplete)

Jorge Martín[†]
CEMFI

Álvaro Sánchez Leache[‡]
CEMFI

Abstract

We study the joint patterns of product differentiation and spatial location in the market for urban amenities. We bypass the limitations of traditional data sources by scraping unstructured microdata from a popular mapping platform, using the restaurant and bar market in Madrid as our case study. We construct a continuous measure of each venue's position in the product space by embedding the text of their menus into dense vector representations. We report three main findings. First, we find complex, high-dimensional, and highly asymmetric patterns of differentiation between venues. Second, we document rich patterns of co-agglomeration and spatial sorting across different types of venues. Third, we find an extensive degree of matching between venues and various sources of local demand, such as different residential demographic groups, tourists, and workers. Our results highlight how state-of-the-art measurement tools open the door to studying new questions of product differentiation and agglomeration in consumption.

1 Introduction

Over the past decade, a vast literature has studied the unequal spatial distribution of urban consumption amenities—restaurants, shops, bars—and its impact on residential choice and welfare inequality (Ahlfeldt et al., 2025; Almagro and Domínguez-Lino, 2025; Cook, 2025; Couture and Handbury, 2020; Couture and Handbury, 2023; Couture et al., 2024; Davis et al., 2019; Diamond, 2016). Despite this widespread interest, we know surprisingly little about how the many varieties within these markets differentiate from one another and locate across the city.

Uncovering the joint patterns of differentiation and location of consumption venues is necessary to accurately capture both consumer welfare and the spatial organization of supply. On the demand side, the spatial distribution of distinct varieties determines the actual consumption opportunities available to consumers, shaping urban accessibility and welfare. In turn, such allocation depends on how distinctly positioned venues choose to distribute themselves across the city. Empirical progress on these dimensions, however, has been constrained by the nature of traditional data sources. Because administrative records and censuses typically report only aggregate venue counts or broad industry classifications, the existing literature is forced to assume that all establishments within a given category are equally substitutable. This forced symmetry conflates the mere quantity of establishments with the true consumption diversity available in a neighborhood, while leaving the actual spatial sorting of differentiated products entirely unobserved.

To advance our understanding, we generate unstructured microdata on individual venues that we scrape from a popular mapping platform, which we use to create a continuous, high-dimensional representation of what each venue offers. Using the Madrid restaurant and bar sector as a case study, our dataset includes the full text of menus, spatial coordinates, and user-generated metadata for thousands of individual establishments. Because a restaurant's identity is inherently defined by the food and drink it serves, we locate each venue within the product space by embedding the text of their menus into dense vector representations—a machine learning technique that encodes text into a high-dimensional space where distances and directions capture semantic meaning.

Armed with our measure of venue positioning, we proceed in three steps. First, we characterize the structure of the product space within Madrid's restaurant sector. In this step, we validate that the embeddings agree with observable metadata—for example, cuisine categories—and demonstrate that they capture a complex, highly asymmetric network of semantic similarities from which we can extract interpretable dimensions of differentiation. Second, we leverage this granularity to uncover new empirical facts about the spatial organization of the market. By extending traditional tests of co-agglomeration to our continuous product space, we document rich co-localization patterns between both highly similar and highly dissimilar venues. Finally, we document an extensive degree of multidirectional sorting between specific types of amenities and various sources of local demand.

To capture the full scope of differentiation across venues, we exploit the continuous geometry of our vector representations. We demonstrate that distances within this space naturally cluster establishments into recognizable groups that strongly align with observable metadata, such as cuisine tags. Going beyond these discrete categories, however, reveals a highly asymmetric topology of differentiation: while traditional Spanish venues form a dense, overlapping continuum of highly substitutable menus, international and niche venues isolate into sharply defined clumps with varying degrees of internal cohesion and complex cross-category relationships. Finally, we show that we can still recover interpretable dimensions of differentiation from this complex space. By projecting venues against synthetic menus that act as semantic poles, we extract continuous economic axes—such as upscale versus budget or daytime versus nightlife—and validate that these purely text-derived rankings strongly predict external proxies like average prices and peak occupancy hours.

We then study the patterns of co-localization to determine *which types* of venues tend to locate together or apart. We extend traditional dartboard tests of co-agglomeration (Duranton and Overman, 2005; Ellison and Glaeser, 1997; Ellison et al., 2010): we evaluate spatial clustering across a kernel-smoothed version of the product space—rather than partitioning into discrete categories. We then draw random permutations where we redistribute restaurants randomly across locations. This allows us to evaluate whether venues at any two distinct points in the semantic space share geographic locations more often than would be predicted by chance. We document a high degree of sorting, with two dominant, mutually exclusive clusters. First, traditional Spanish restaurants cluster together among highly similar establishments. Second, international cuisines and nightlife venues locate together despite offering vastly different products. These patterns reveal that agglomeration in consumption is not just driven by clusters of close substitutes, but also by complementary, distinct venues sharing the same spaces.

Finally, we document an extensive degree of two-way sorting between different types of venues and the local composition of demand. We show that the intense spatial clustering of venues translates into distinct consumption profiles at the neighborhood level, with the majority of areas hosting venue compositions that diverge significantly from the citywide distribution. We also uncover how neighborhood demographics correlate highly with the local composition of supply. Strikingly, we find that while local income levels do not predict the *overall* magnitude of a neighborhood's specialization, they strongly correlate with its semantic dimension: high-income areas systematically skew toward upscale venues, just as high employment density predicts the presence of daytime-oriented venues and tourist activity correlates with nightlife consumption. Furthermore, the presence of foreign residents predicts semantic proximity with specific international cuisines.

2 Data and methodology

2.1 Administrative data

Our starting point is the Censo de Locales y sus Actividades, an administrative census maintained by the City of Madrid that records the universe of commercial establishments with an active economic license. The census provides the official name and street address of each establishment, along with its licensed activity category. We restrict attention to establishments classified under food-service categories—restaurants, bars, cafeterias and similar venues—which yields a population of $18,629$ establishments. This administrative register serves two purposes: it provides a comprehensive sampling frame that does not suffer from the selection inherent in user-generated platforms, and it supplies the exact addresses we use to query Google Maps.

2.2 Web-scraped venue data

We match each establishment in the Censo de Locales to its Google Maps listing using an automated scraping procedure. For each address in the administrative census, we query Google Maps and retrieve both structured metadata—price range, average user rating, number of reviews, Google-assigned category tags—and unstructured content in the form of *menu labels*: the list of food and drink items that each venue's page shows as photo tags. We successfully match $15,847$ of the $18,629$ census establishments to a Google Maps listing. Of these, $7,464$ have menu information available, which constitutes our main analysis sample.

Table 1 reports descriptive statistics for the full census population, the Google Maps matched sample, and the subsample with menus. The menu subsample is balanced relative to the matched Google Maps population in terms of average rating and price range, suggesting that menu availability is not strongly selected on venue quality or price tier. Restaurants with menus tend to have somewhat more reviews, consistent with the fact that more established or popular venues are more likely to feature a complete Google Maps profile.

	Censo de Locales	Google Maps	Menu Labels	Description	Desc & Menu Labels
N	18,629	15,847	7,464	3,795	2,331
Mean (sd) Rating	-	4.17 (0.51)	4.21 (0.39)	4.16 (0.38)	4.18 (0.36)
Mean (sd) N. Reviews	-	815 (1294)	1098 (1407)	1841 (1784)	2060 (1784)
Mean (sd) Price	-	14.81 (9.19)	15.21 (9.14)	18.42 (11.75)	18.48 (11.57)

Table 1: Sample composition and representativeness. Each column reports summary statistics for a progressively restricted subsample of Madrid food-service venues. *Censo de Locales* is the administrative universe of establishments with an active food-service license. *Google Maps* is the subset matched to a Google Maps listing. *Menu Labels* is the subset for which at least one menu item is available on Google Maps and constitutes our main analysis sample. *Description* and *Desc & Menu Labels* are subsets with owner-uploaded venue descriptions, reported for reference. Mean rating, review count and price are not available from the administrative census. Standard deviations in parentheses.

2.3 Measuring product positioning: menu embeddings

The central measurement challenge in this paper is to characterize where each restaurant sits in product space. Restaurants differentiate themselves along many dimensions: the cuisine tradition they belong to, the individual dishes and ingredients they feature, their food-to-drink ratio, the specificity or breadth of the menu, the balance between regional and cosmopolitan offerings, the degree of formality implied by the food choices. These dimensions are not orthogonal and do not collapse into a small set of types. A measure of product position therefore needs to be continuous and multi-dimensional—it must locate each venue in a space where multiple axes of differentiation are simultaneously represented.

We construct such a measure by encoding each venue's menu as a point in a high-dimensional semantic space—a dense vector representation in which directions encode meaning and geometric proximity reflects semantic similarity (Mikolov et al., 2013b; Reimers and Gurevych, 2019). Concretely, we concatenate each venue's menu labels into a text string, prepend the task instruction *"En este sitio se ofrecen los siguientes productos o servicios:" ("This venue offers the following products or services:")*, and pass the result through the Multilingual-E5 model (Wang et al., 2024). This model is trained through contrastive learning (Gao et al., 2021): it learns to map texts with similar meaning to nearby points in the output geometry while placing semantically dissimilar texts far apart. The result is a vector $\mathbf{e}_i \in \mathbb{R}^d$ for each venue i—a *menu embedding*—that encodes the full profile of the venue's offerings as a continuous coordinate in the semantic space.

Three properties of this representation make it well-suited to studying product differentiation. First, it yields a complete and symmetric network of pairwise product similarities. The cosine similarity

$$\text{sim}(i, j) = \frac{\mathbf{e}_i \cdot \mathbf{e}_j}{\|\mathbf{e}_i\| \, \|\mathbf{e}_j\|}$$

is defined for all $N(N-1)/2$ pairs in the sample, producing an $N \times N$ matrix that captures the full structure of product distances across the restaurant population without any pre-specified taxonomy. This parallels the approach of Hoberg and Phillips (2016), who construct continuous firm-level similarity measures from product descriptions in annual reports and show they outperform standard industry classifications; here the text source is menus, and the encoder is a modern transformer model rather than a bag-of-words representation.

The second and third properties concern the internal geometry of the embedding space. Directions in the space encode semantic content and can be anchored to interpretable concepts. Since Mikolov et al. (2013a) it has been known that semantic relationships in dense vector spaces are approximately linear—directions are consistent and meaningful across the geometry. Building on this, Kozlowski et al. (2019) and Grand et al. (2022) show that projecting document vectors onto a direction defined by contrasting reference texts recovers a scalar measure of each document's position along a specified conceptual axis. We apply this *semantic anchoring* procedure to our menus: constructing artificial menus that represent the poles of a given dimension—such as upscale versus budget dining, or international versus traditional Spanish cuisine—and projecting all venue vectors onto the resulting direction yields an interpretable, continuous score for each venue. Finally, the representation is genuinely high-dimensional: no small number of principal components captures the bulk of variation in the embedding matrix (Figure A.1 in the Appendix), confirming that differentiation across venues operates across many simultaneously active dimensions that cannot be collapsed without substantial information loss. We exploit all three properties in Section 3.

2.4 Other data

We complement the venue-level data with neighborhood-level information on the composition of local demand. We obtain data on residential demographics—income distribution, age structure, household composition and nationality of residents—from the municipal statistics portal of Madrid. To measure tourist presence, we collect data on active Airbnb listings at the neighborhood level. To construct neighborhood-level measures of total employment, we use data from the Madrid Business

Establishments and Employment Dataset (Colectivo Empresarial), an administrative dataset covering the universe of establishments in the Community of Madrid with information on their location, sector, and employment.

3 Capturing differentiation between venues

In this section we show that our measure captures rich patterns of asymmetric differentiation across venues, and that it allows to recover meaningful dimensions of differentiation between them. We proceed in three steps. First, we show how our measure allows us to construct a network of product similarity across restaurant venues. Second, we show how this network is highly asymmetric across venue categories. Third, we show how we can obtain interpretable dimensions of differentiation by exploiting the properties of the embeddings through *semantic projection*.

3.1 Differentiation across individual venues

The most basic requirement of our measure is that it captures the degree to which any given pair of venues' offerings are similar or different from each other. As detailed in section 2.3, the semantic proximity between two restaurants' offerings is captured by the *cosine similarity* of their embeddings $\text{sim}(i,j) = \frac{\mathbf{e}_i \cdot \mathbf{e}_j}{\|\mathbf{e}_i\| \, \|\mathbf{e}_j\|}$.

Table 2 provides an intuitive example from four sample restaurants: a Japanese restaurant (*Restaurante Kippu*), a Chinese restaurant (*Wok Man Express*), and two Spanish restaurants from different regions (*Bar Entre Cáceres y Badajoz, El Madroño*). On panel (a) it displays the name of each restaurant, their category tags, ratings, price ranges, and the first ten items of the menus that were used to construct the product positioning embeddings. On panel (b), it displays the 4×4 matrix of semantic similarities, *normalized to percentiles over the whole distribution of bilateral similarities*. Among these four examples, the two Spanish restaurants are classified as being extraordinarily similar among all the city's restaurants, since they share very semantically similar offerings. *Restaurante Kippu* and *Wok Man Express* are, however, more similar between each other than they are to any of the two Spanish restaurants in this example.

Generalizing these bilateral comparisons across all restaurant pairs, we obtain figure 1. To produce this figure, we project each high-dimension embedding onto a two-dimensional plane using the t-SNE algorithm, which preserves local similarity in the form of proximity in the 2-D space. Each

Restaurant 1: Restaurante Kippu	
Category tag:	Restaurante japonés
Rating:	4.8 ⋆ (6,455 reviews)
Price range:	30-40 €
Menu (Spanish):	Vino, Makis, Salsas, Tempura Roll, Chupitos Sake, Tarta de Queso, Yakimeshi Pato, Kimchi Yakiudon, Mochi de Cheesecake, Coulant de Chocolate, Amplia Selección de Sushi Variado A la Carta Niguiri Y MAKI
Restaurant 2: Wok Man Express	
Category tag:	Restaurante chino
Rating:	4.4 ⋆ (321 reviews)
Price range:	1-10 €
Menu (Spanish):	Gyozas, Pato Kung Pau, Sopa de Wantun, Arroz de la Casa, Rollito Primavera, Verduras Con Gambas, Tallarines Con Cerdo, Arroz Frito Con Verduras, Tallarines Fritos Con Ternera, Arroz Blanco Con Ternera Y Verduras
Restaurant 3: Bar Entre Cáceres y Badajoz	
Category tag:	Bar de tapas
Rating:	3.9 ⋆ (5,801 reviews)
Price range:	10-20 €
Menu (Spanish):	Barra, Tapas, Cañas, Rabo de Toro, Tinto de Verano, Pulpo A la Gallega, Torreznos Con Padron, Lomo de Buey A la Plancha, Mollete Con Tomate Y Jamón, Tocino Con Pimientos de Padrón, Huevos Rotos Con Jamón Ibérico
Restaurant 4: El Madroño	
Category tag:	Restaurante de comida madrileña
Rating:	4.0 ⋆ (1,440 reviews)
Price range:	20-30 €
Menu (Spanish):	Sangria, Torrijas, Rabo de Toro, Patatas Bravas, Cocido Madrileño, Pulpo A la Gallega, Croquetas de Jamón, Callos A la Madrileña, Calamares A la Andaluza, Chupitos de Licor de Madroño, Huevos Rotos Con Jamón Ibérico

(a) Representative examples of restaurants from the Madrid dataset.

	Restaurante Kippu	**Wok Man Express**	**Bar Entre Cáceres y Badajoz**	**El Madroño**
Restaurante Kippu	—	87.8	4.5	9.3
Wok Man Express	87.8	—	40.0	12.8
Bar Entre Cáceres y Badajoz	4.5	40.0	—	99.4
El Madroño	9.3	12.8	99.4	—

(b) Examples of pairwise restaurant similarities based on menu embeddings.

Table 2: Restaurant examples and pairwise menu similarities. Panel (a) describes four representative establishments chosen to illustrate contrasting positions in the differentiation space. Panel (b) reports the full 4×4 matrix of pairwise cosine similarities between their menu embeddings, standardized across the full sample. The matrix validates the measure: venues that share a culinary tradition are assigned high similarity, while venues from distinct culinary traditions are assigned low similarity, and partial overlaps are reflected in intermediate values.

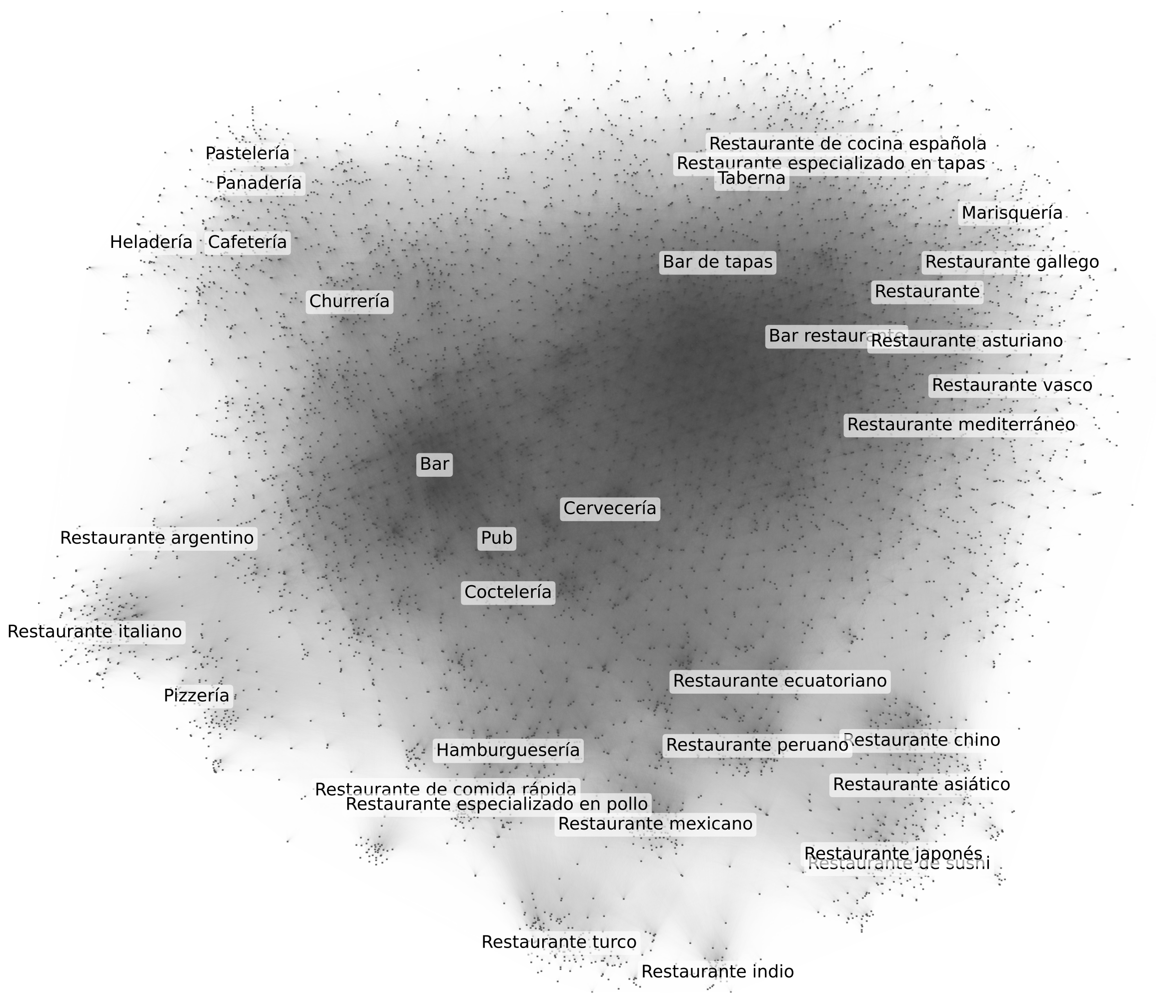


Figure 1: Restaurant differentiation network. Each point in the graph is a restaurant, whose coordinates are projected from the high-dimensional embedding space to a 2D plane using the t-SNE algorithm, which aims to preserve semantic proximity. Shades connecting points are similarity relationships, with darker shades indicating a higher degree of similarity. Tags are placed at the local density peaks of major cuisine type clusters.

point in this plot represents a restaurant, and connecting shades represent similarity, with darker shades denoting a higher degree of similarity between venues or groups of venues. We also label the density peak of restaurants for major groups, as denoted by their Google Maps tags. Note that, throughout the rest of the paper, we will represent the restaurant product space through this projection, so that individual venues and regions of the product space are consistent across figures.

Analyzing figure 1 reveals that our approach identifies both the functional differentiation of venues and the specific clustering of regional cuisines. The plot roughly organizes into four main areas: breakfast and pastry venues occupy the top-left, bars and pubs form a dense core in the center, traditional Spanish restaurants populate the top-right, and international cuisines stretch across the bottom periphery. Notably, Spanish establishments blend into a smooth, continuous mass of differentiation, reflecting highly overlapping menus across venues, whereas international venues form distinct, isolated clumps. Within this international section, semantically related cuisines sit closer together: Asian restaurants (chino, japonés) form one neighborhood, Latin American venues (ecuatoriano, peruano, mexicano) group in another closer to the center, and Italian restaurants, Pizzerias and Argentinian restaurants cluster on the left of the 2D projection. It is important to remark that this topology emerges entirely unprompted, since the restaurant categories themselves are not fed to the embedding model.

3.2 Asymmetric differentiation across venue types

We perform further analysis of our measure to confirm that it captures coherent groupings of venues, and to understand how venues relate to each other within and across categories. On the first account, we confirm that groupings of venues created from the embeddings are consistent with Google Maps tags, which are the best alternative measure of restaurant type. On the second account, we find that the degree of differentiation is highly heterogeneous and asymmetric within and across categories, formalizing the intuitions from figure 1 and showing how our measure provides additional information relative to discrete categorizations.

First, we follow Hoberg and Phillips (2016) to validate that our measure clusters venues into meaningful groups[1]. We cluster venues using the K-Means algorithm, and validate them against the human-labeled tags from Google Maps. We first create a 35-group classification which we validate

[1] Hoberg and Phillips (2016) create a matrix of manufacturing firm similarities based on a bag-of-words model of their 10-K reports product descriptions, and they validate that industry classifications based on their measure of similarity are consistent with traditional industry classifications

<table>
<tr><td colspan="2">Submarket 1: Restaurante japonés (123 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>Shao grill, Ta-Kumi | Restaurante Japonés - Madrid, Torijiro Madrid, SUMO Juan Bravo, Soy</td></tr>
<tr><td>Google Maps tags:</td><td>Restaurante japonés (53), Restaurante de sushi (20), Restaurante asiático (15), Restaurante (11), Restaurante japonés auténtico (6)</td></tr>
<tr><td>Common dishes:</td><td>Sushi Variado (24), Takoyaki (16), Sopa Miso (12), Yakisoba (11), Tartar de Atún (10), Postre (10), Tartar de Salmón (9), Sopa de Miso (7), Tarta de Queso (6), Ramen (6), Pollo Al Limón (6), Gyozas de Pollo (6)</td></tr>
<tr><td colspan="2">Submarket 2: Restaurante chino (95 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>Palacio Imperial, 68 Yong Tai, Restaurante Royal Cantonés, Wok 4 You Bravo Murillo, Sabor Cantonés</td></tr>
<tr><td>Google Maps tags:</td><td>Restaurante chino (53), Restaurante asiático (18), Restaurante (10), Restaurante de cocina mandarina (3), Bar (3)</td></tr>
<tr><td>Common dishes:</td><td>Arroz Tres Delicias (36), Pollo Al Limón (23), Pan Chino (17), Rollito de Primavera (15), Pollo Con Almendras (14), Gyozas (13), Cerdo Agridulce (12), Cerveza China (10), Pato A la Naranja (10), Ensalada China (10), Rollitos Vietnamitas (8), Rollo de Primavera (8)</td></tr>
<tr><td colspan="2">Submarket 3: Restaurante italiano (144 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>Grosso Napoletano, La Bruschetta, BIZZO! | Pizzería en Madrid | Pizza | Masa & Ensalada | Restaurante Italiano | Bernabéu, La Tagliatella Senza Glutine (Sin Gluten) | Avda. de América, Madrid, Fellina Chamberí | Restaurante Italiano Madrid</td></tr>
<tr><td>Google Maps tags:</td><td>Restaurante italiano (79), Pizzería (32), Restaurante (24), Restaurante de alta cocina (1), Restaurante de comida para llevar (1)</td></tr>
<tr><td>Common dishes:</td><td>Tiramisú (47), Tarta de Queso (27), Pizza Diavola (17), Pasta Carbonara (15), Tiramisú Casero (14), Lasaña (13), Vitello Tonnato (12), Pizza Calzone (12), Tiramisu (11), Calzone (10), Pizza Carbonara (10), Spaghetti Carbonara (9)</td></tr>
<tr><td colspan="2">Submarket 4: Restaurante mexicano (91 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>Malpica, Restaurante Tepic, The Toast Taproom, TKO Tacos Iglesia, Tierra Burrito Bar</td></tr>
<tr><td>Google Maps tags:</td><td>Restaurante mexicano (64), Restaurante (9), Restaurante de comida rápida (4), Taquería (4), Restaurante de burritos (4)</td></tr>
<tr><td>Common dishes:</td><td>Michelada (24), Tacos Al Pastor (18), Nachos Con Guacamole (13), Tacos de Cochinita Pibil (12), Quesadillas (10), Tarta de Queso (10), Sopa Azteca (8), Tortillas (8), Cochinita Pibil (8), Nachos (8), Postre (7), Margaritas (7)</td></tr>
<tr><td colspan="2">Submarket 5: Cafetería (65 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>Miga Bakery, Open25 Coffee Market, 80's coffee, Harina Alfonso XII, Santagloria Coffee & Bakery Conde de Peñalver</td></tr>
<tr><td>Google Maps tags:</td><td>Cafetería (43), Panadería (13), Pastelería (2), Restaurante de brunch (2), Bar (2)</td></tr>
<tr><td>Common dishes:</td><td>Café Con Leche (13), Croissant Mixto (13), Tarta de Zanahoria (11), Tarta de Queso (10), Cinnamon Roll Canela (9), Mini Croissant (8), Cafe Con Leche (7), Chocolate Caliente (6), Ensaimada (6), Matcha Latte (5), Flat Croissant Choco Blanco Y Lotus (5), Palmeritas (4)</td></tr>
<tr><td colspan="2">Submarket 6: Restaurante asturiano (71 restaurants)</td></tr>
<tr><td>Sample restaurants:</td><td>La Fistiella, Casa de Asturias, Restaurante Teitu, Casa Mingo, Asador El Molinón</td></tr>
<tr><td>Google Maps tags:</td><td>Restaurante asturiano (30), Restaurante (17), Sidrería (8), Bar restaurante (5), Restaurante de cocina española (4)</td></tr>
<tr><td>Common dishes:</td><td>Chorizo A la Sidra (31), Fabada Asturiana (21), Patatas Al Cabrales (14), Pulpo A la Gallega (13), Tarta de Queso (12), Arroz Con Leche (11), Arroz Con Bogavante (11), Sidra Natural (10), Cachopo de Ternera (9), Tabla de Quesos Asturianos (9), Huevos Rotos Con Jamón (8), Pulpo A la Brasa (8)</td></tr>
</table>

Table 3: Examples of restaurant clusters based on menu embedding similarity. Clusters are created with the K-Means algorithm, creating 35 clusters from the menu embeddings. Each cluster groups restaurants with similar menu offerings. For each cluster we report a sample of member restaurants, the most frequent Google Maps category tags assigned to members (with counts in parentheses), and the most common menu items across members (with frequency counts). Each cluster label is the most common Google Maps tag.

manually. Table 3 shows six such groups, with examples of venues, the most common offering in their menus, and the counts of Google Maps tags for restaurants in those groups. We also run the classification with 169 clusters, the same as Google Maps tags, and we validate that the distribution of restaurant category sizes matches up to those provided by Google. This result is in figure A.2 in the Appendix.

While our method produces results which are close to human-labeled tags, it also reveals why discrete categories are insufficient to capture the full product space. Specifically, our continuous measure uncovers a high degree of asymmetry in how venues differentiate, both within and across these categories. To demonstrate this, we group venues by their Google Maps tags and construct a matrix of average similarities, visualized as a heatmap in Figure 2.[2]. Analyzing this matrix gives us rich information about how each type of venue differentiates. First, looking at the diagonal of the matrix, we have the *own-similarities* for each category, that is, the average similarity between all the venues that share a particular tag. Looking at the off-diagonal elements of the matrix, we have *cross-category* degrees of similarity, that is, for categories i and j, the average of the cosine similarity of all elements of category i with all elements of category j.

The results from figure 2 highlight two ways in which differentiation is highly asymmetric. First, when we look across different categories in the diagonal of the matrix, there are very different degrees of own-similarity for different categories. Second, when we look at the off-diagonal elements, the degree of cross-similarity is highly heterogeneous. It is evident that there are several distinct blocks[3] that feature high degrees of cross-similarity between their members, but there is also heterogeneity outside of these broad blocks. It is also interesting to note that all of these patterns are visible in the t-SNE visualization in figure 1.

3.3 Semantic projection: interpretable dimensions of differentiation

Having shown that our measure captures venue similarity and aligns with standard categories, our next step is to validate that we can extract interpretable dimensions of differentiation. Fundamentally, the model compresses a multitude of semantic features from the menu text into dense continuous vectors. The crucial property of these vector spaces is that their linear geometry preserves semantic relationships (Kozlowski et al., 2019; Mikolov et al., 2013a). This is a highly desirable feature: it implies that moving along a specific linear direction corresponds to a systematic shift in an underlying

[2]For visualization purposes, we restrict the sample to tags with at least 30 venues.
[3]Spanish restaurants, Latin-American restaurants, Asian restaurants, fast food, Italian-Argentinian, bars, breakfast

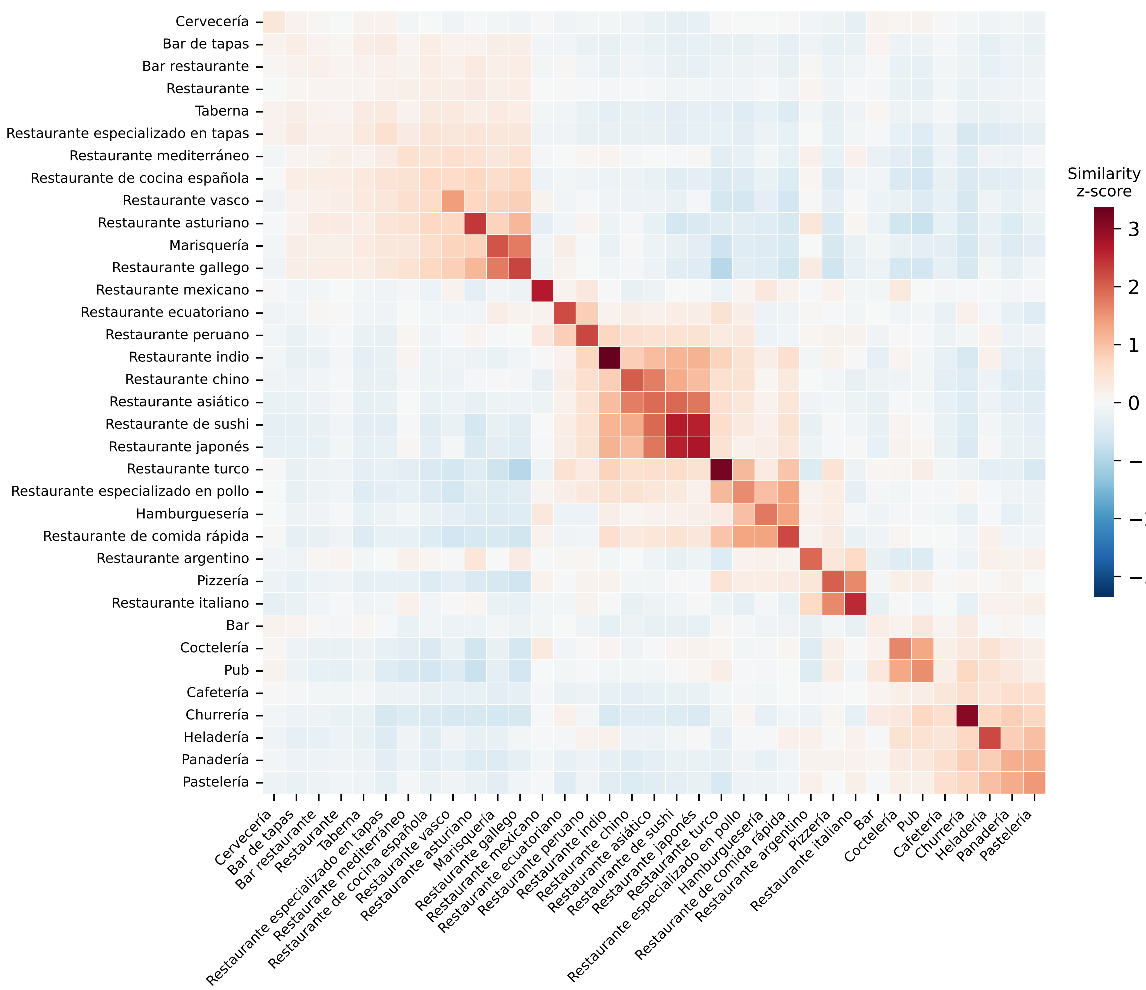


Figure 2: Cross-cluster similarity matrix. We compute the cosine similarity between each venue and aggregate at the mean for every pair of Google Maps tag. The matrix reveals that differentiation is highly asymmetric across clusters: some venue types are close substitutes for multiple other categories, while others occupy more isolated market positions.

concept. The challenge, however, is the extreme dimensionality of the space.[4] Because semantic information is distributed across hundreds of latent variables, no single coordinate maps directly to an observable trait. Consequently, the main hurdle is identifying a direction that isolates the specific dimension we want to evaluate.

To identify these directions, we turn to the method of *semantic projection*. The intuition is straightforward: if we can formulate two text descriptions that capture the pure, opposing extremes—or "poles"—of a characteristic (for example, an archetypal "upscale" menu versus a "budget" menu), the vector difference between these two points forms an isolated, one-dimensional axis representing that exact trait. By orthogonally projecting each restaurant's embedding onto this line, we obtain a continuous scalar indicating its position along that dimension. Historically, this method has been applied to word embeddings—where pairs of opposing words serve as poles—to quantify psychological biases (Caliskan et al., 2017), uncover structural class and gender stereotypes (Ash et al., 2024; Kozlowski et al., 2019), and reconstruct human conceptualizations of objects (Grand et al., 2022). By adapting this approach to our menu space, we can query the model to extract specific, unobserved variables of economic interest.

Formally, for each dimension of interest, we define a positive anchor text $\mathbf{a}^+$ and a negative anchor text $\mathbf{a}^-$, encoded using the same model applied to the venues' menus. The semantic anchor score for restaurant i on that dimension is the projection of its menu embedding onto the unit direction defined by the anchor contrast:

$$s_i = \mathbf{e}_i \cdot \frac{\mathbf{a}^+ - \mathbf{a}^-}{\|\mathbf{a}^+ - \mathbf{a}^-\|}.$$

A higher score indicates that the menu of restaurant i is semantically closer to the positive pole; a lower score indicates proximity to the negative pole.

To create semantic poles, we construct synthetic menus that represent the extremes of three fundamental dimensions of differentiation: upscale vs. budget venues, traditional Spanish vs. international venues, and nightlife-oriented vs. day-oriented consumption. To build these synthetic menus, we use the observable metadata to identify real restaurants situated at the empirical extremes of our data—for example, extracting items exclusively from the cheapest and most expensive venues to define the budget and upscale anchors, respectively. We then use only the text of these items to form the anchors. By embedding these polar archetypes, the semantic projection effectively ranks all other

[4] Figure A.1 in the Appendix shows how a Principal Component decomposition needs many dimensions to explain the variation in the menu embeddings.

restaurants in the sample based entirely on how closely their actual product offerings resemble these known extremes. Table 4 outlines the representative items used to generate these textual anchors. We will leverage these exact dimensions later in Section 5.2 to analyze how neighborhoods specialize across different regions of the product space.

Anchor	Pole	Menu
Upscale (vs. Budget)	Upscale	Arroz Negro Y Paella Valenciana, Tarta de Queso, Falso Risotto, Filloas Crujientes Rellena Crema Pastelera Al Punto de Azúcar Y Canela, Truffle Pasta, Musaka Vegetal, Puerros Miga Cana, Grilled Meat on the Table, Tortilla de Patatas Y Cebolla Confitada, Filete de Ternera de Asturias
	Budget	Bocadillo de Lomo Con Queso, Patatas Fritas, Café Con Leche Y Bollería, Pincho de Tortilla de Patatas, Roscón de Reyes, Bagel de Salmón Y Aguacate, Nachos Con Queso, Tosta de Jamon, Cachopo Asturiano, Flat White
Spanish (vs. International)	Spanish	Caña, Entrecot de Ternera, Parrilluda de Verduras, Cachopo, Caparazón de Centollo, Pan, Boquerones Fritos, Tarta de Queso, Fabada Asturiana, Jamón Ibérico de Bellota
	International	Maki, Amazing Choco Ball, Ravioli Di Salmone, Enmoladas, Huevo Y Trufa, Crocante de Atún, Tarta de Oreo, Hamburguesa Wallace, Coxinhas de Pollo Y Croquetas de Carne, Hawaiana Crispy
Nightlife (vs. Day)	Nightlife	Korean Spicy Fried Chicken, Tortilla de Patatas, Patatas Bravas Y Ali-oli, Cócteles Deliciosos, Huevos Rotos Con Gambones Al Ajillo, Nachos Con Queso, Gin Tonic, No Capricciosa, Tartar de Atún, Flamenquines
	Day	Desayuno Café Con Leche Y Tostada de Tomate, Montado Beicon Con Queso, Sopa Castellana, Caldo de Pollo, Cafe 1 2 Tostada, Cortado, Tarta de Bizcocho Con Fresa, Flat White, Chuletón, Cocido Madrileño

Table 4: Examples of artificial menus we created to compute the semantic anchors. First row menu includes items of a high price restaurant. Second row menu includes items of a low price restaurant. Third row menu includes items of an international restaurant. Fourth row menu includes items of a traditional Spanish restaurant. Fifth row menu includes items of a night consumption restaurant. Sixth row menu includes items of a day consumption restaurant. We can embed these menus, subtract one from the other and obtain the directional embedding that represents the dimension of interest.

To validate that these projections recover meaningful economic dimensions, we compare the text-derived scores against observable venue metadata. Critically, we compute each venue's anchor score via a 50-fold cross-validation procedure: the anchor direction for each fold is estimated solely from the menus of venues outside that fold, so no restaurant's own items ever contribute to the direction it is scored against. This design rules out mechanical self-inclusion as a driver of any correlation we document. Figure 3 shows how these purely product-based, cross-validated projections align with external metadata. We find that venues with a higher mean price per meal score systematically higher on the *upscale vs. budget* dimension; venues tagged as international cuisines separate cleanly from traditional Spanish venues along the *Spanish vs. international* dimension; and venues with peak occupancy in the evening score higher on the *night vs. day* dimension.

Note that our semantic scores are constructed entirely from the text of the menus—representing the actual products offered—whereas the validation metrics (user-reported prices, explicit categorical tags, and hourly occupancy trends) are external variables that did not enter the embedding model at all. The cross-validated design ensures that any alignment we observe is not an artefact of a venue's own menu items being used to define the anchor it is scored against; the associations survive even when each venue is scored by a direction constructed from entirely different establishments. We therefore interpret these results as genuine evidence that the embeddings encode economically meaningful characteristics of the product space.

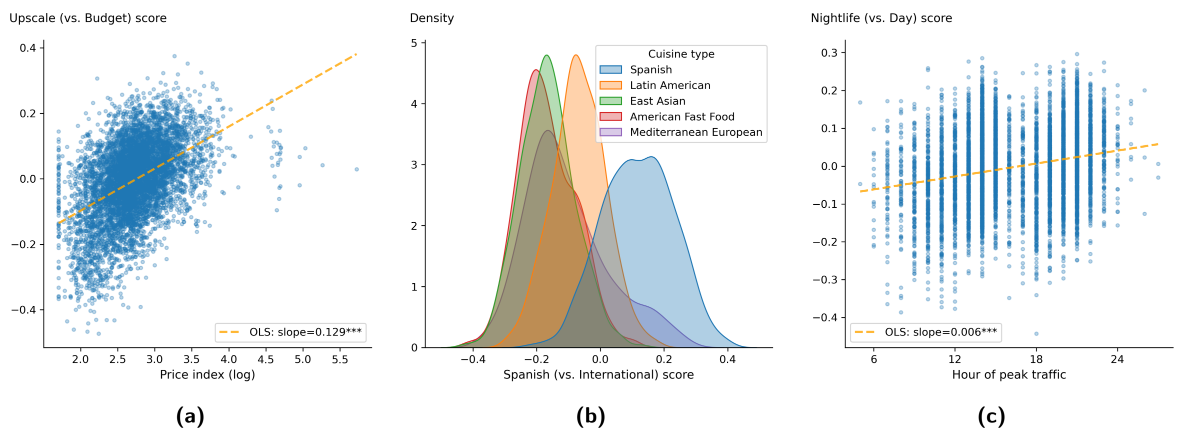
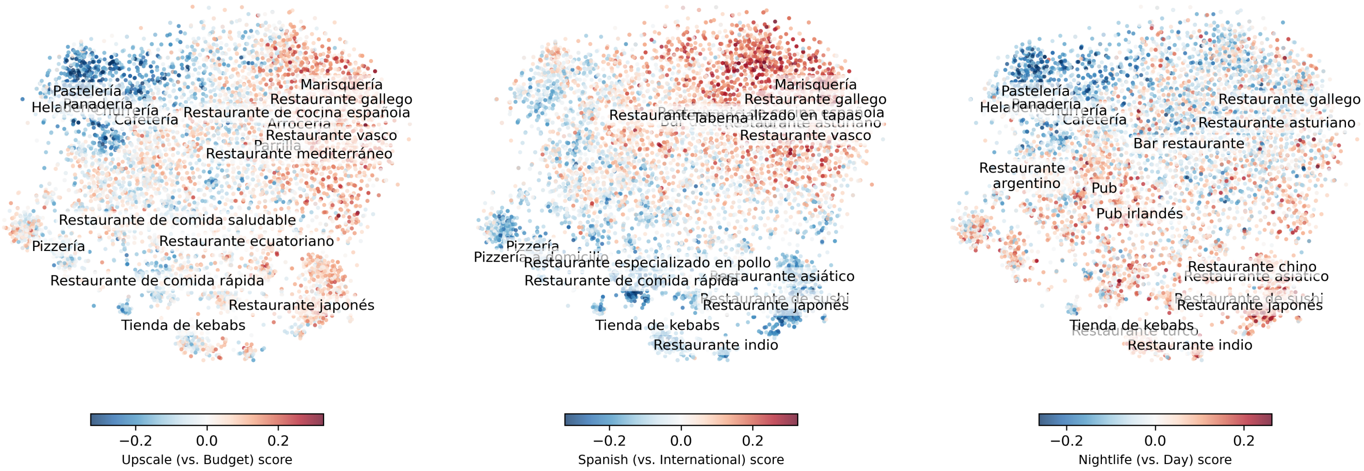


Figure 3: Panel (a) shows the correlation between the *upscale vs. budget* anchor scores and the mean price per meal for each restaurant. Panel (b) compares the distribution of *international vs. Spanish* anchor scores for restaurants tagged as international cuisine versus those tagged as traditional Spanish cuisine. Panel (c) plots the *night vs. day* anchor scores against the median peak-occupancy hour across days of the week for each restaurant. The price index is built using the reviewer-reported price range per person after a meal. We take the mid-point of the ranges and weight by the share of reviews in each price range. The Spanish-Other Country tag is obtained mapping the tags found in Google Maps data, which usually include the region the venue specializes in. The hour of most occupancy is obtained using the occupancy data from Google Maps for each day of the week.

Figure 4 maps these cross-validated scores back onto the t-SNE embedding, coloring each venue by its position along each of the three dimensions. The figure confirms that the three dimensions are mutually consistent: breakfast venues (*Pastelería*, *Panadería*, *Cafetería*) score low on both the upscale and the nightlife dimensions, placing them in the budget and daytime corner of the embedding space; traditional Spanish regional restaurants (*Marisquería*, *Restaurante gallego*) score high on both the upscale and Spanish dimensions; and nightlife venues (*Pub*, *Pub Irlandés*) score high on the nightlife dimension. The dimensions can also combine in less obvious ways: Japanese restaurants (*Restaurante japonés*), for instance, score low on the Spanish dimension—as expected for an international cuisine—but rank among the highest on both upscale and nightlife, occupying a distinct corner of the product space that no single categorical label would identify.

Taken together, the results in this subsection establish that semantic projection can recover in-



(a) **(b)** **(c)**

Figure 4: Visualization of the semantic anchors in the embedding space. Panel (a) colors restaurants by their *Upscale vs. Budget* score, panel (b) by their *Spanish vs. International* score, and panel (c) by their *Nightlife vs. Day* score. Each point represents a restaurant in the low-dimensional projection of the embedding space using t-SNE. Google Maps Tags are only shown for the top and bottom deciles of the anchor scores.

terpretable, economically meaningful dimensions of product differentiation directly from menu text. We exploit this property in Section 5.2, where we use the upscale, Spanish, and nightlife scores to characterize how neighborhoods specialize across different regions of the product space. The three dimensions studied here were chosen deliberately because each has a clean observable proxy—prices, cuisine tags, and peak-hour traffic—which allowed us to validate the approach rigorously. Having validated the method on observable traits, future applications could leverage survey data or expert opinion to construct anchors for more abstract, subjective dimensions that lack obvious empirical proxies.

4 Agglomeration - co-localization between venues

Having established that venues occupy a rich and structured product space, we now ask whether this structure has a spatial counterpart: do certain types of venues systematically tend to share the same spaces, and do other types tend to locate apart? This question has two important implications. First, if there is significant spatial sorting between venues, then the composition of a location's restaurant amenity supply cannot be summarized simply by the counts of restaurants. Two locations with identical restaurant counts may offer fundamentally different access to the city's consumption landscape if they offer distinct bundles of venues. Second, by learning *which types* of venues tend to locate together, we learn about which types of venues are likely creating positive externalities for each other, helping to understand the microfoundations of agglomeration in consumption.

While empirical evidence exists that consumption venues benefit from proximity to other establishments (Koster et al., 2019; Leonardi and Moretti, 2023; Shoag and Veuger, 2018), it is not clear ex-ante *which* types of venues should locate together. Recent micro-founded evidence on consumer behavior sheds light on two forces that generate opposing predictions. Miyauchi et al. (2025) document that consumers frequently chain trips, making sequential stops at multiple venues within walking distance during the same outing. If consumers are combining visits to venues that serve distinct but sequentially complementary purposes—an aperitivo, a dinner, a post-dinner drink—then *complementary* venues benefit from being close to each other, creating an incentive to co-locate. Vitali (2025) instead documents that consumers often travel to a location before deciding exactly where to purchase, preferring areas that offer many similar options to compare. This generates incentives for *substitute* venues to cluster: concentration of interchangeable options makes a location more attractive to uncertain consumers. Both forces plausibly operate simultaneously, and their relative strength across different regions of the product space is an empirical question.

To answer this question, we construct a dartboard test on the joint spatial-semantic distribution of venues that extends the co-agglomeration framework of Ellison et al. (2010) to continuous physical and product space. Each restaurant i is represented as a point (s_i, v_i) in the product of physical space and menu embedding space. The central object is a pairwise co-agglomeration score C_{ij}, which measures, for each pair of venues i, j, whether venues similar to i tend to locate in close proximity to venues similar to j. We evaluate C_{ij} against a dartboard null in which menu embeddings are randomly reshuffled across the fixed set of observed restaurant locations. Without this permutation test, a given C_{ij} might be high simply due to idiosyncratic features of the data. Permuting creates a pair-specific null distribution that filters out these idiosyncrasies, isolating colocation patterns that are genuinely unlikely to occur by chance.

To operationalize this test we need two proximity measures: one in physical space and one in product space, which are the continuous analogues of the spatial unit and industry category assignments in Ellison, Glaeser, and Kerr. For physical proximity we use a bisquare kernel $K_S(s_i, s_j) = \left(1 - \left(\frac{d(s_i,s_j)}{h}\right)^2\right)^2 \mathbf{1}\{d(s_i, s_j) < h\}$, with a bandwidth of 500 meters. The compact support is deliberate: spillovers between consumption venues are known to be highly local, operating within walking distance, and typically decay completely beyond 500 meters (Koster et al., 2019; Miyauchi et al., 2025; Shoag and Veuger, 2018). For product-space proximity we use an exponentiated-cosine kernel $K_V(v_i, v_j) = \exp\{-\kappa(1 - \cos(v_i, v_j))\}$. We calibrate the bandwidth κ from the empirical distribution of pairwise cosine distances using the median heuristic of Gretton et al. (2012), setting it so that the kernel reaches half its maximum value at 0.2 times the median pairwise cosine distance between venues.

Using these kernels we estimate the joint and marginal distributions as

$$\hat{P}_{\text{true}}(s, v) = \frac{1}{N} \sum_i K_S(s, s_i) K_V(v, v_i), \qquad \hat{P}_{\text{null}}(s, v) = \hat{P}_S(s) \hat{P}_V(v),$$

where $\hat{P}_S(s) = \frac{1}{N} \sum_i K_S(s, s_i)$ and $\hat{P}_V(v) = \frac{1}{N} \sum_i K_V(v, v_i)$ are the estimated marginal distributions of locations and venue types, respectively. The discrepancy $\Delta(s, v) = \hat{P}_{\text{true}}(s, v) - \hat{P}_{\text{null}}(s, v)$ measures whether a given semantic profile is over- or under-represented at a given location relative to independence. We define the co-agglomeration score for any pair (i, j) as the inner product of their discrepancy profiles across physical space,

$$C_{ij} = \int \Delta(s, v_i) \Delta(s, v_j) \, ds$$

which is positive when both types are over-concentrated in the same parts of the city, negative when one is over-concentrated where the other is absent, and near zero when their spatial patterns are unrelated.[5]

We assess each C_{ij} against the dartboard null by permutation: in each draw, menu embeddings are randomly reassigned across the fixed set of observed restaurant locations, and C_{ij} is recomputed. Because the scale of C_{ij} is determined by the marginal spatial and semantic properties of each type rather than by its co-agglomeration pattern, raw values carry no natural benchmark and are not comparable across pairs. The permutation generates a pair-specific empirical null distribution that absorbs these differences. Comparing the observed scores against these null distributions yields a standardized co-agglomeration matrix that clearly identifies which pairs of venue types are highly co-localized, which are systematically dispersed, and which co-localize at a rate indistinguishable

[5]To see how C_{ij} nests the Ellison et al. (2010) co-agglomeration index as a special case, replace $K_S(s_i, s_j)$ with $\mathbf{1}\{s_i \text{ and } s_j \text{ are in the same barrio}\}$ and $K_V(v_i, v_j)$ with $\mathbf{1}\{v_i \text{ and } v_j \text{ share a Google Maps tag}\}$. Under these hard-assignment kernels, Δ reduces to barrio-level deviations of tag shares from city-wide tag shares, and C_{ij} collapses to the Ellison et al. (2010) size-weighted covariance of those deviations across barrios—exactly their numerator. Substituting continuous kernels for the hard-assignment indicators generalizes both dimensions simultaneously.

from random chance.

Before breaking down individual pairs, our first statistic is a global test for semantic sorting across the entire city. Intuitively, if restaurant types are randomly assigned across the observed restaurant map, the joint distribution of locations and types will exactly match the independence benchmark, meaning the discrepancy surface $\Delta(s, v)$ should be close to zero everywhere. To measure the total amount of spatial sorting in the data, we simply aggregate this squared discrepancy across all locations and all semantic profiles. This aggregate measure—formally known as the Hilbert-Schmidt Independence Criterion (HSIC)—equals zero if and only if location and restaurant type are completely independent. As with our pairwise scores, we evaluate this single global statistic against the fixed-location permutation null to test whether the overall spatial structure of the city deviates significantly from random chance. We reject the null of spatial independence at the permutation p-value of $p = 0.0003$, confirming that the city-wide distribution of venue types is far from random.

We then look at the degree of co-localization across pairs. Of the 27,851,916 restaurant pairs, 41.4% are significantly co-agglomerated and 41.5% are significantly dispersed at the 5% level (see figure A.3 in the Appendix for the global HSIC permutation null and the full distribution of pairwise scores). To visualize the full $N \times N$ matrix of pairwise co-agglomeration z-scores, we aggregate by Google Maps tag. For each pair of tags (a, b), we compute the average standardized co-agglomeration score across all restaurant pairs (i, j) with $i \in a$ and $j \in b$. The result is the tag-level heatmap in Figure 5. It is important to note that this aggregation is purely for visualization: the underlying z-scores are computed at the level of individual restaurant pairs, so restaurants that belong to different tags but have semantically similar menus still contribute to the co-agglomeration signal. This distinguishes our approach from a co-agglomeration measure computed directly at the tag level, which would miss cross-tag affinities and treat all venues within a tag as identical .Figure A.4 in the Appendix plots the network of significant co-location and dispersion links without any aggregation, projected to the same t-SNE coordinates as the similarity network from figure 1.

Figure 5 reveals two dominant co-agglomeration clusters, each with a strikingly different internal structure. The first comprises traditional Spanish venues—specialised tapas restaurants, regional cuisines (Basque, Asturian, Galician), and seafood restaurants—which are also highly similar to each other in product space (Figure 2). Here, within-cluster co-agglomeration aligns naturally with within-cluster similarity, consistent with comparison-shopping by consumers seeking interchangeable options (Vitali, 2025). The second cluster is more surprising: it groups international cuisines—Asian,

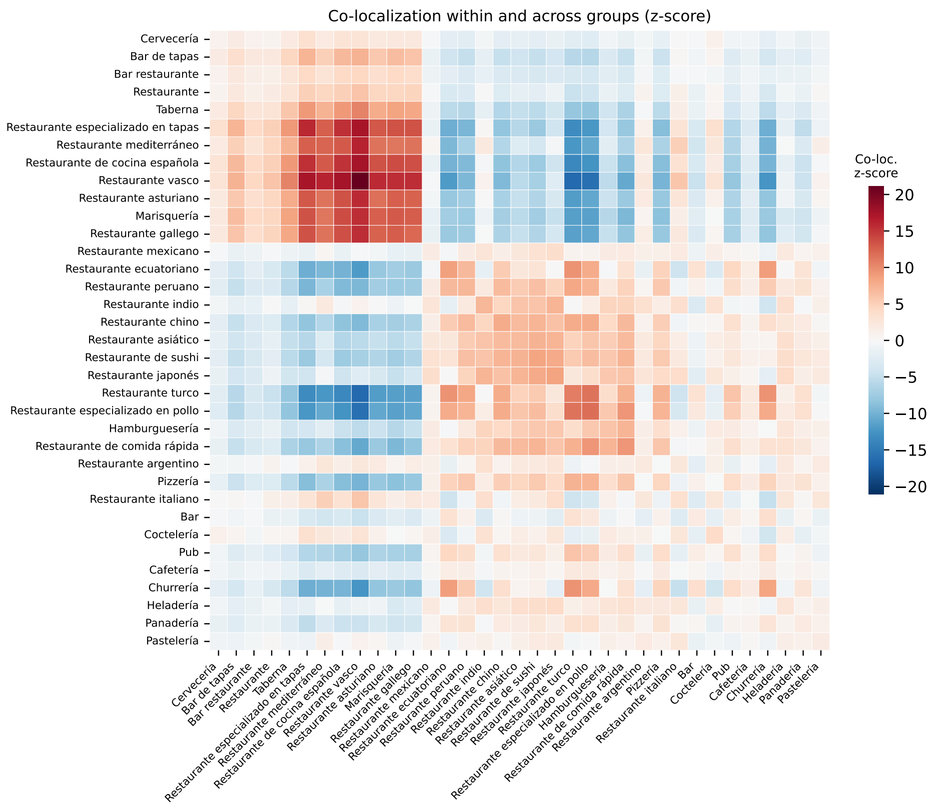


Figure 5: Co-location z-score matrix across restaurant tags.

Latin American, and Italian-Argentinian venues—alongside bars, coctelerías, and pubs, yet these venue types are semantically *distant* from each other in product space. Their co-agglomeration therefore cannot be driven by substitutability alone, and may instead reflect a common draw to locations with high foot traffic, nightlife activity, or a cosmopolitan consumer base. The two clusters actively disperse from each other, with blue off-diagonal cells confirming that areas dense in traditional Spanish venues tend to be sparse in international and nightlife venues, and vice versa.

These patterns illustrate that co-agglomeration in the consumption space is rich and heterogeneous: similar venues cluster, but so do dissimilar ones, and the forces generating each type of clustering need not be the same.

5 Two-way sorting between venues and local demand

The co-agglomeration patterns in the previous section show that venues of different types are far from randomly distributed across the city, and that the two dominant clusters occupy largely non-overlapping areas. A natural question is, does this spatial sorting reflect in the local composition of residents and visitors?

Prior work has documented that consumer demographics and tourist intensity predict which broad types of venues are present in a neighborhood (Almagro and Domínguez-Lino, 2025; Couture and Handbury, 2020; Couture and Handbury, 2023; Couture et al., 2024; Hidalgo, 2024; Hidalgo et al., 2024; Lanzara and Minerva, 2018), and that consumers differ in their preferences across cuisine categories (Cook, 2025; Davis et al., 2019). We ask whether these correlations extend to the finer dimensions of product differentiation captured by our semantic embeddings. We find that they do: neighborhood demographic and visitor characteristics predict not only which broad category of venue is present, but the full semantic profile of the local amenity mix, along dimensions that are invisible to discrete category-based measures.

We proceed in three steps. First, we measure the extent to which each neighborhood hosts a composition of venues that is distinctive relative to the citywide distribution, using a distributional discrepancy measure built on the same semantic embeddings and kernel from the previous section. Second, we examine whether neighborhoods with more distinctive demographic or economic characteristics exhibit higher levels of this specialization. Third, we study the *direction* of specialization by projecting neighborhood-level venue profiles onto our semantic anchors, and ask whether different types of neighborhoods specialise in systematically different directions of the product space—connecting the fine-grained sorting patterns back to the observable characteristics of local demand.

5.1 Neighborhood specialisation

We begin by testing which neighborhoods exhibit a composition of venues that differs from the city-wide distribution. If amenities were randomly distributed across space, the local distribution of embeddings would simply mirror the overall city benchmark; systematic deviations indicate local specialisation. To formalise this comparison, we use the Maximum Mean Discrepancy (MMD) statistic from Gretton et al. (2012). Just as our HSIC statistic previously measured global dependence, this metric detects local specialisation by evaluating the distance between the distribution of venues within a specific neighborhood and those in the rest of the city.

A key advantage of this approach is that it admits a closed-form empirical estimator based purely on pairwise similarities. Let $V_n = \{v_1, \ldots, v_n\}$ be the semantic embeddings of venues located in the target neighborhood, and $V_m = \{v_1, \ldots, v_m\}$ be the embeddings of venues in the rest of the city. The squared discrepancy is given by:

$$\mathrm{MMD}^2(V_n, V_m) = \frac{1}{n^2} \sum_{v_i, v_j \in V_n} K_V(v_i, v_j) + \frac{1}{m^2} \sum_{v_i, v_j \in V_m} K_V(v_i, v_j) - \frac{2}{nm} \sum_{v_i \in V_n, v_j \in V_m} K_V(v_i, v_j)$$

Crucially, we evaluate these similarities using the exact same exponentiated-cosine kernel K_V and median-heuristic bandwidth κ defined in the previous section, ensuring that our definition of semantic proximity remains strictly consistent throughout the analysis. Each term in the equation has a natural interpretation: the first captures the average semantic similarity of venues within the neighborhood, the second measures the average similarity across the rest of the city, and the third represents the cross-similarity between local and outside venues. If a neighborhood contains a random sample of the city's amenities, within-group and across-group similarities balance out, leaving the statistic close to zero. Conversely, if a neighborhood specialises in a distinct profile, its internal similarity will heavily outweigh its cross-similarity to the rest of the city, driving the estimator strictly positive.

Each term in the MMD statistic has a natural interpretation. The first term measures the average semantic similarity between venues within the neighborhood, while the second measures the average similarity of venues across the rest of the city. The third term captures the average cross-similarity between neighborhood venues and outside venues. If a neighborhood contains a random sample of the city's amenities, within-group and across-group similarities will balance out, and the statistic will be close to zero. Conversely, if a neighborhood specialises in a distinct profile of venues, its internal similarity will heavily outweigh its cross-similarity to the rest of the city, driving the MMD strictly positive.

To assess statistical significance, we implement a permutation test analogous to our dartboard null. We repeatedly draw random samples of size n from the citywide distribution and recompute the discrepancy to generate a neighborhood-specific null distribution. Large standardised values of the statistic indicate that the neighborhood contains a set of venues that differs systematically from the citywide benchmark. As shown in Figure 6, roughly two thirds of neighborhoods are significantly specialised according to our measure, confirming that the spatial distribution of amenities is highly non-random and that distinct venue mixes are a common urban feature.

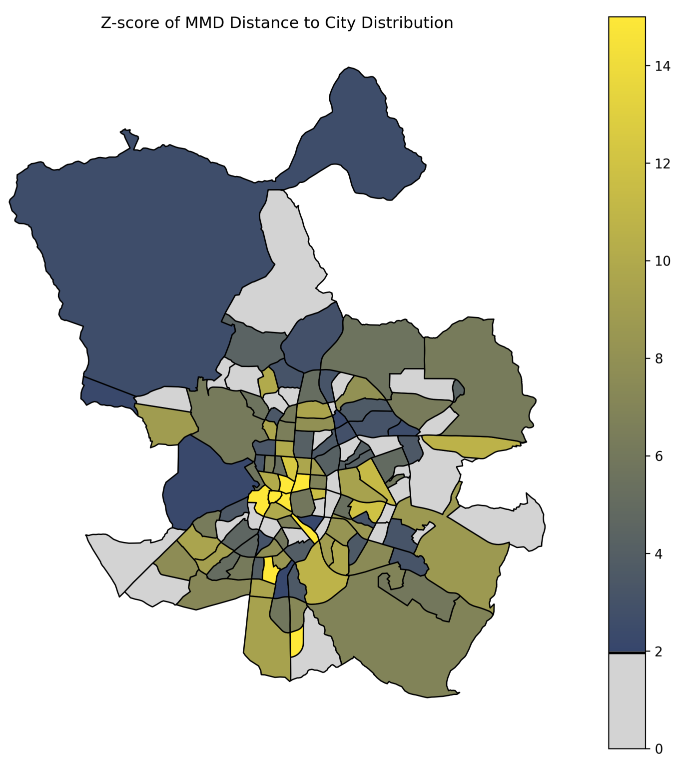


Figure 6: Neighborhood specialisation across the city. Neighborhoods with venue distributions that do not significantly differ from the citywide benchmark are shown in grey. Specialised neighborhoods are colored according to their standardised discrepancy (z-score) from the null distribution.

Having established that most neighborhoods host distinct amenity profiles, we next examine whether this specialisation correlates with the composition of local residents, visitors, and workers. In Table 5, we regress our measure of neighborhood specialisation against standardized measures of income, tourist activity, age, household composition, local employment, and distance to the city center. When considered in isolation (Columns 1–7), almost all of these characteristics strongly predict specialisation. Areas with higher tourist activity, younger populations, a larger presence of foreign residents, fewer children, and higher employment density are all significantly more specialised. Interestingly, income alone (Column 1) is not a significant predictor, indicating that distinctive amenity profiles emerge in both low- and high-income areas.

However, when we evaluate these characteristics jointly in Column 8, a fundamentally different picture emerges. The share of foreign residents and local employment density remain strong predictors of specialisation, but the previously significant effects of age and household composition disappear entirely, suggesting they were proxying for other spatial demographics. Strikingly, once we control for these local characteristics alongside the neighborhood's distance to the center, the association with tourist activity flips to negative and significant, while income remains statistically indistinguishable from zero.

Table 5: The dependent variable in all models is the neighborhood-level mean of its venues' Maximum Mean Discrepancy (MMD) with respect to the city-level distribution. All independent variables are standardized.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Household income	-0.585							0.527
	(0.484)							(0.495)
Log tourist activity		2.221***						-1.235**
		(0.446)						(0.580)
Young share			2.827***					-0.053
			(0.418)					(0.586)
Spanish share				-2.979***				-3.249***
				(0.410)				(0.618)
Share with kids					-1.673***			0.045
					(0.464)			(0.465)
Log employment						1.645***		1.743***
						(0.465)		(0.465)
Log distance to center							-2.938***	-2.205***
							(0.412)	(0.473)
R^2	0.011	0.161	0.261	0.290	0.092	0.088	0.282	0.515
Adj. R^2	0.004	0.155	0.256	0.285	0.085	0.081	0.277	0.487
Observations	131	131	131	131	131	131	131	131

Standard errors in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Taken together, these results demonstrate that the spatial distribution of urban amenities is highly structured, with the majority of neighborhoods hosting significantly specialised venue profiles relative to the citywide baseline. Furthermore, the degree of this specialisation correlates with local neighborhood characteristics. Specifically, areas with more foreign residents and higher employment density are more specialised, while local income levels have no significant effect. Having shown that neighborhood demographics predict the level of specialisation, we next examine some specific dimensions of specialization.

5.2 Heterogeneous specialisation

Knowing that neighborhoods host systematically different venue compositions raises the question of whether this specialization is systematically directed: do wealthier neighborhoods lean towards upscale venues, do tourist-heavy areas skew towards nightlife, do neighborhoods with more foreign residents host more international cuisines? To answer this, we project each neighborhood's venue embeddings onto the semantic anchors that we defined in Section 3.3 and test whether the resulting anchor scores are correlated with neighborhood characteristics, using the same regression framework as before.

Table 6 shows the results. First, along the upscale–budget dimension, higher-income neighborhoods and those with a greater share of foreign residents are systematically associated with more upscale venues. This resolves an apparent paradox from the previous section: while local income does not predict the *magnitude* of a neighborhood's specialisation (its overall MMD distance from the citywide average), it strongly dictates its *direction*. A high-income area and a low-income area may be equally specialised overall, but they diverge from the city average in entirely different semantic directions.

Beyond income, we observe two other distinct patterns. Along the time-of-day dimension, neighborhoods with higher employment density specialise in morning-oriented venues, while those with more Airbnb listings skew towards night consumption. When we consider a single Spanish–international axis, higher-income and more peripheral neighborhoods lean international, while central neighborhoods tilt towards traditional Spanish cuisine, but the presence of foreign residents is not significantly associated with the overall international score. This changes when we disaggregate by cuisine type: foreign residents are significantly associated with a higher share of Latin American and East and South Asian venues, but not with Italian or American fast food. The latter two are instead associated with higher employment density.

Table 6: Dependent variables are neighborhood-level z-scores of specialisation along different semantic dimensions. Lower values indicate specialization toward the negative pole of each dimension (top label), while higher values indicate specialization toward the positive pole (bottom label label). For cuisine-specific dimensions, the negative pole (Spanish) is omitted for brevity. All independent variables are standardized.

	Broad Semantic Dimensions			Specific International Cuisines (vs. Spanish)			
	Upscale (vs. Budget)	Nightlife (vs. Day)	Spanish (vs. Int'l)	Latin American	East Asian	American Fast Food	Med. European
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income and demand							
Household Income	1.262***	-0.243	-0.408**	0.206	0.234	0.060	0.537***
	(0.176)	(0.151)	(0.166)	(0.174)	(0.161)	(0.187)	(0.168)
Log tourist activity	0.400*	0.319*	0.143	0.054	-0.144	-0.313	-0.134
	(0.206)	(0.177)	(0.194)	(0.204)	(0.188)	(0.220)	(0.197)
Demographics							
Young Share	0.288	-0.180	-0.216	-0.001	0.181	0.155	0.178
	(0.208)	(0.179)	(0.196)	(0.206)	(0.190)	(0.222)	(0.199)
Share with Kids	-0.036	-0.163	-0.114	0.074	0.147	0.159	0.092
	(0.166)	(0.142)	(0.156)	(0.164)	(0.151)	(0.176)	(0.158)
Foreign Share	0.411*	-0.238	-0.317	0.468**	0.475**	0.256	0.331
	(0.220)	(0.188)	(0.207)	(0.218)	(0.201)	(0.234)	(0.210)
Economic activity and location							
Log employment	-0.206	-0.308**	-0.254	0.076	0.208	0.330*	0.319**
	(0.166)	(0.142)	(0.156)	(0.164)	(0.151)	(0.176)	(0.158)
Log distance to center	0.378**	-0.079	-0.505***	0.695***	0.555***	0.389**	0.469***
	(0.168)	(0.144)	(0.159)	(0.167)	(0.154)	(0.179)	(0.161)
R^2	0.356	0.124	0.181	0.186	0.197	0.123	0.209
Adj. R^2	0.319	0.074	0.135	0.139	0.152	0.073	0.164
Observations	131	131	131	131	131	131	131

Standard errors in parentheses.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

To summarize this section, we find three main results. First, the intense spatial sorting of venues that we described in Section 4 translates into distinct amenity profiles at the neighborhood level. Second, neighborhoods with distinct demographic and economic characteristics host similarly distinct venue compositions. Finally, these local characteristics predict not just the overall magnitude of specialisation, but its specific direction along interpretable dimensions of product differentiation.

6 Conclusion

In this paper, we set out to capture differentiation in the market for urban amenities, and how the diverse options in this market are distributed across space. To do so, we propose a measure of product positioning and differentiation by using dense vector representations of product descriptions—in our case, the menus of restaurants.

Through an extensive process of validation, we find that this measure captures high-dimensional and highly asymmetric patterns of differentiation across venues, but also that it captures interpretable dimensions such as the upscale-ness of a venue. Looking at the spatial distribution of restaurants, we find that co-localization patterns are highly structured: some venue types concentrate together while others actively avoid each other, and this spatial sorting translates into high degrees of specialization at the neighborhood level. The patterns are consistent with multiple forces operating simultaneously—comparison-shopping, trip-chaining, sorting toward shared demand environments—but disentangling their relative contribution remains an open question. Local income, tourist intensity, and the share of foreign residents each predict the direction of a neighborhood's semantic specialization in distinct ways.

Our results show that measures such as our own will prove useful to study local consumption markets. Standard frameworks use counts of venues and utility aggregators like CES where each venue is equally substitutable with every other, but substitutability is in fact highly heterogeneous and asymmetric across venue pairs—which means that establishment counts are a misleading proxy for the consumption variety and welfare a neighborhood actually offers. Our method also opens a set of questions that were previously out of reach: how co-localization patterns and the degree of diversity scale across cities of different sizes, whether agglomeration forces allow new varieties to emerge as markets grow, and how the composition of local amenity supply responds to shocks such as gentrification or large-scale immigration.

Finally, this approach holds broad promise for the study of product differentiation itself. Theories of multidimensional differentiation have long existed but lacked a tractable empirical counterpart. Text embeddings provide exactly this missing infrastructure, capturing a parsimonious, high-dimensional proxy of product positioning. Yet, fully integrating this approach demands new economic and econometric frameworks to formalize their scope and limitations.

References

Ahlfeldt, G., Bald, F., Roth, D., & Siedel, T. (2025). *Measuring the urban quality of life premium*. https://rationality-and-competition.de/wp-content/uploads/discussion_paper/544.pdf

Almagro, M., & Domínguez-Lino, T. (2025). Location sorting and endogenous amenities: Evidence from amsterdam. *ECTA*, *93*(3), 1031–1071. https://doi.org/10.3982/ECTA21394

Ash, E., Chen, D. L., & Ornaghi, A. (2024). Gender attitudes in the judiciary: Evidence from US circuit courts. *American Economic Journal: Applied Economics*, *16*(1), 314–350. https://doi.org/10.1257/app.20210435

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Cook, C. (2025). Heterogeneous preferences for neighborhood amenities: Evidence from GPS data. *Review of Economics and Statistics*, 1–43. https://doi.org/10.1162/rest_a_01551

Couture, V., Gaubert, C., Handbury, J., & Hurst, E. (2024). Income growth and the distributional effects of urban spatial sorting. *Review of Economic Studies*, *91*(2), 858–898. https://doi.org/10.1093/restud/rdad048

Couture, V., & Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, *119*, 103267. https://doi.org/10.1016/j.jue.2020.103267

Couture, V., & Handbury, J. (2023). Neighborhood change, gentrification, and the urbanization of college graduates. *Journal of Economic Perspectives*, *37*(2), 29–52. https://doi.org/10.1257/jep.37.2.29

Davis, D. R., Dingel, J. I., Monras, J., & Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, *127*(4), 1684–1738. https://doi.org/10.1086/701680

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000. *American Economic Review*, *106*(3), 479–524. https://doi.org/10.1257/aer.20131706

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *Rev Econ Stud*, *72*(4), 1077–1106. https://doi.org/10.1111/0034-6527.00362

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in u.s. manufacturing industries: A dartboard approach. *Journal of Political Economy*, *105*(5), 889–927. https://doi.org/10.1086/262098

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, *100*(3), 1195–1213. https://doi.org/10.1257/aer.100.3.1195

Gao, T., Yao, X., & Chen, D. (2021). SimCSE: Simple contrastive learning of sentence embeddings. In M.-F. Moens, X. Huang, L. Specia, & S. W.-t. Yih (Eds.), *Proceedings of the 2021 conference on empirical methods in natural language processing* (pp. 6894–6910). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.552

Grand, G., Blank, I. A., Pereira, F., & Fedorenko, E. (2022). Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nat Hum Behav*, *6*(7), 975–987. https://doi.org/10.1038/s41562-022-01316-8

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkkopf, B., & Smola, A. (2012). A kernel two-sample test. *Journal of Machine Learning Research*. https://www.jmlr.org/papers/volume13/gretton12a/gretton12a.pdf

Hidalgo, A. (2024). Your room is ready: Tourism and urban revival. *Regional Science and Urban Economics*, *109*, 104059. https://doi.org/10.1016/j.regsciurbeco.2024.104059

Hidalgo, A., Riccaboni, M., & Velázquez, F. J. (2024). The effect of short-term rentals on local consumption amenities: Evidence from madrid. *Journal of Regional Science*, *64*(3), 621–648. https://doi.org/10.1111/jors.12685

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423–1465. Retrieved March 13, 2026, from https://www.jstor.org/stable/26549905

Koster, H. R., Pasidis, I., & Van Ommeren, J. (2019). Shopping externalities and retail concentration: Evidence from dutch shopping streets. *Journal of Urban Economics*, *114*, 103194. https://doi.org/10.1016/j.jue.2019.103194

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *Am Sociol Rev*, *84*(5), 905–949. https://doi.org/10.1177/0003122419877135

Lanzara, G., & Minerva, G. A. (2018). Tourism, amenities, and welfare in an urban setting. *SSRN Journal*. https://doi.org/10.2139/ssrn.3198062

Leonardi, M., & Moretti, E. (2023). The agglomeration of urban amenities: Evidence from milan restaurants. *American Economic Review: Insights*, *5*(2), 141–157. https://doi.org/10.1257/aeri.20220011

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013a). Efficient estimation of word representations in vector space. https://doi.org/10.48550/arXiv.1301.3781

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. In C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 26). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

Miyauchi, Y., Nakajima, K., & Redding, S. J. (2025). The economics of spatial mobility: Theory and evidence using smartphone data. *The Quarterly Journal of Economics*, *140*(4), 2507–2570. https://doi.org/10.1093/qje/qjaf038

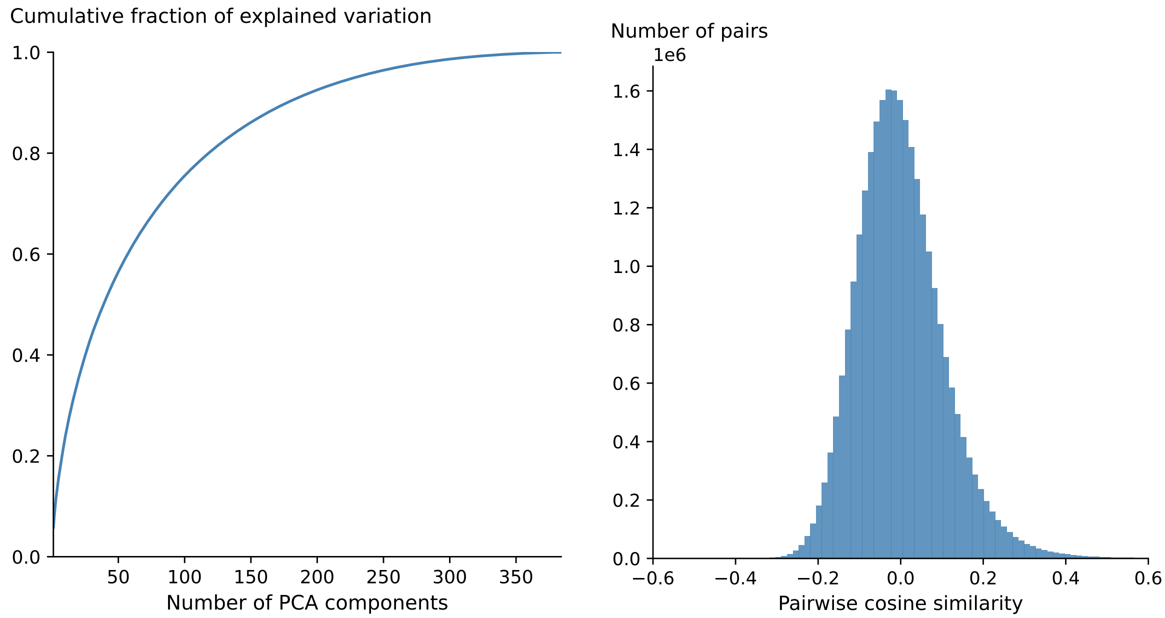
Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3980–3990. https://doi.org/10.18653/v1/D19-1410

Shoag, D., & Veuger, S. (2018). Shops and the city: Evidence on local externalities and local government policy from big-box bankruptcies. *The Review of Economics and Statistics*, *100*(3), 440–453. https://doi.org/10.1162/rest_a_00703

Vitali, A. (2025). Consumer search and firm location: Theory and evidence from the garment sector in uganda.

Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., & Wei, F. (2024). Multilingual e5 text embeddings: A technical report [Version Number: 1]. https://doi.org/10.48550/ARXIV.2402.05672

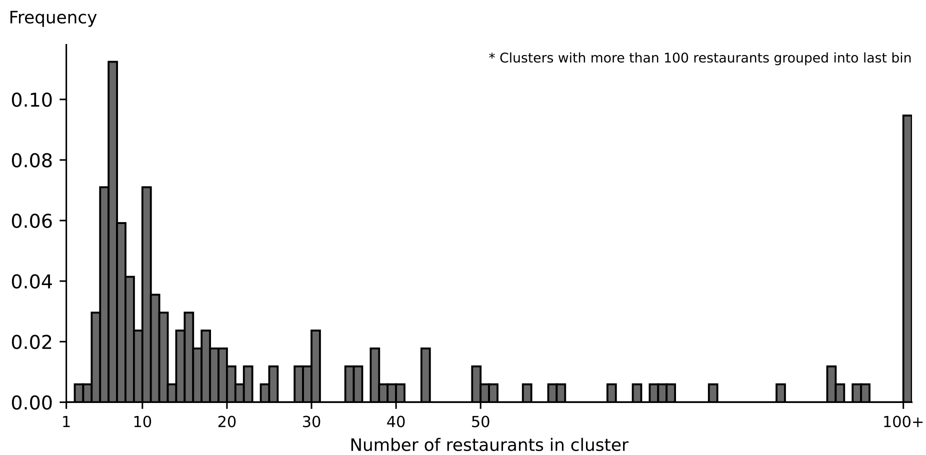
A More information on the embeddings measure



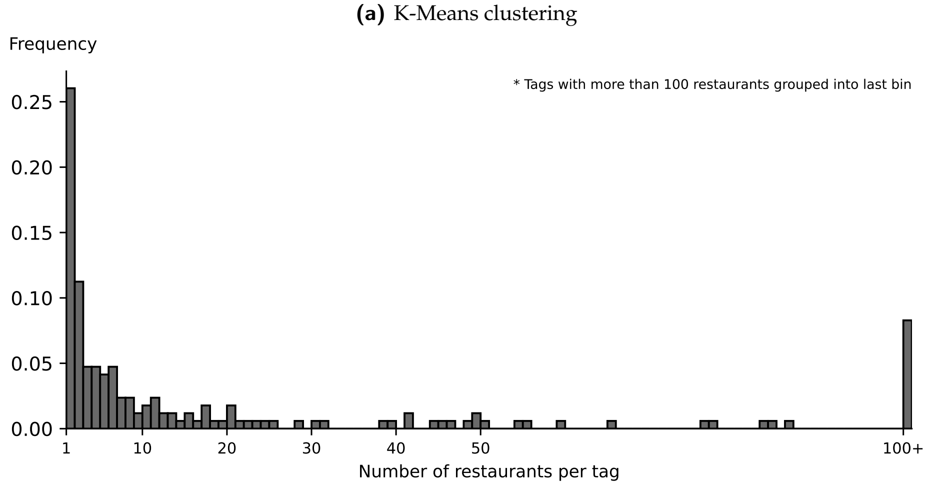
(a) Restaurant differentiation network. **(b)** Distribution of pairwise menu similarities.

Figure A.1: Differentiation network and similarity distribution for Madrid restaurants.

B Differentiation and clustering



(a) K-Means clustering



(b) Google Maps tags

Figure A.2: Distribution of cluster sizes and tag sizes across restaurant clusters.

C Co-localization between venues

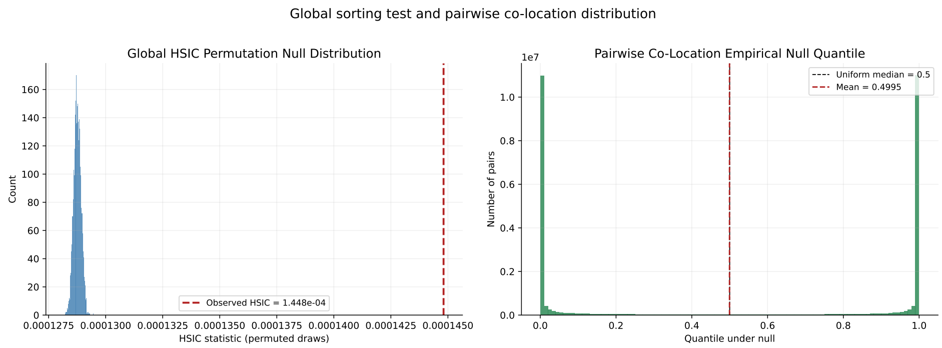
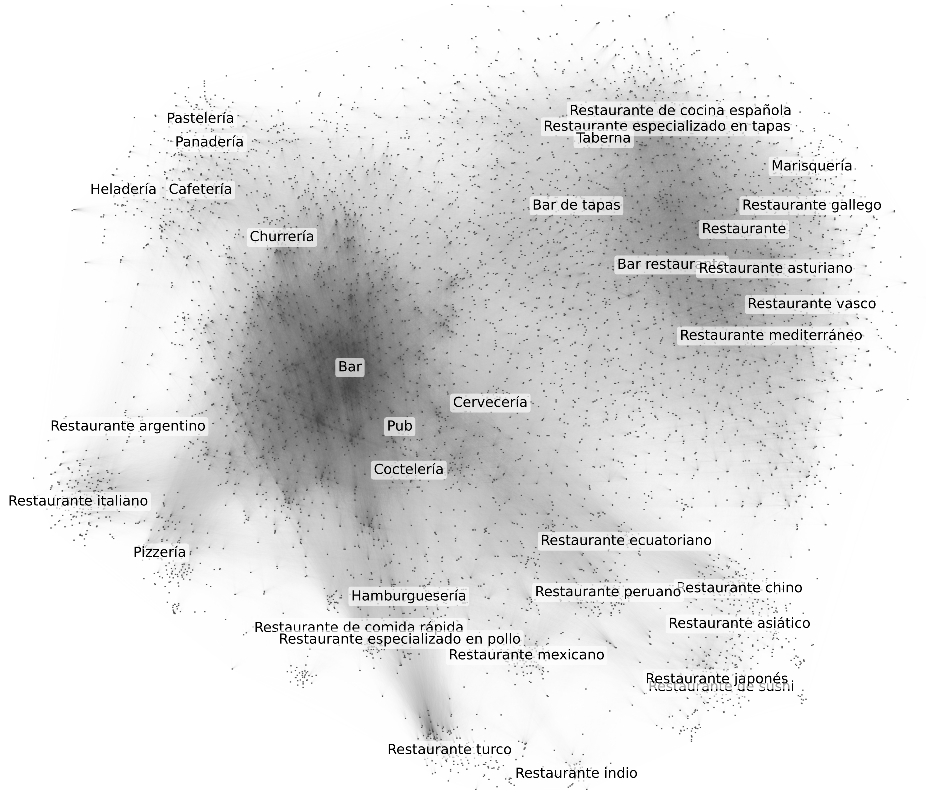
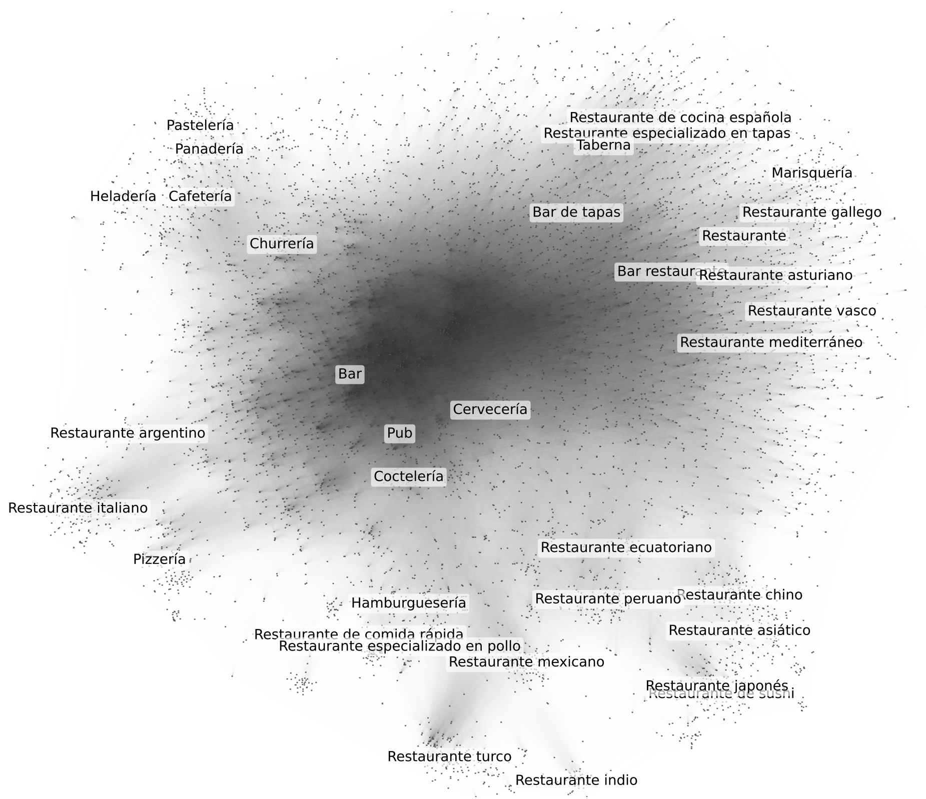


Figure A.3: Global HSIC and pairwise distribution.



(a) Top co-location links.



(b) Top dispersion links.

Figure A.4: Restaurant co-agglomeration and spatial sorting patterns.