Research Title: Decoding Place Functionality: Identifying Emerging Third Places and the Drivers of Remote Work Visitation using GPS signals and Google Places Reviews

Inessa Tregubova, Department of Geography, Hebrew University of Jerusalem

Abstract

The COVID-19 pandemic has reshaped urban dynamics by disrupting the traditional "home-work-third place" triad, giving rise to a new "home-third place-home" spatial pattern. In this structure, 'new third places' (non-traditional work places) increasingly function as office substitutes, moving professional activity from the central business district (CBD) closer to home locations. On the other hand, with the opportunity to work outside the office, home location is moving from the city center to suburbia where housing costs are lower. Existing research acknowledges that among two remote neighborhoods people choose one that provides a bigger diversity of urban amenities. However the specific influence of 'new third places' in this choice remains under-researched due to identification complexities. Therefore, this research develops a new methodology to identify emerging "new third places" within the Tel Aviv Metropolitan Area (TAM) and examine their role in supporting remote work at the neighborhood level. The area choice is explained by the negative migration dynamic in Tel Aviv between 2020-2022 and retaining work from home (WFH) pattern discovered in our earlier work. To answer the research question, for each place of interest (POI) suitable for remote work we construct the dataset of metrics derivded from the GPS-signals collected between January 2019 and September 2023 and extend with features available via Google Places API. The metrics construction involves different techniques starting from the estimation of visiting statistic to measuring the semantic distance of each POI to 'work' using Natural Language Processing Word2Vec model. The final identification of office substitutes we produce using the Logistic Regression classifier with Elkan and Noto adjustment methodology which allows to have only positive groundtruth. For this we are using POIs with the clear evidence of remote work from Google Reviews.

Finally, to evaluate the relative impact of third places versus general amenities on local WFH levels, we estimate independent OLS regressions for each quarter. We then compare the magnitude and statistical significance of their standardized coefficients. Ultimately, these findings provide a deeper understanding of how locations successfully transform into third places, and how these spaces relate to local WFH trends.

Introduction

Although the onset of the COVID-19 pandemic occurred six years ago, the hybrid and working from home (WFH) patterns that emerged during that time have now firmly established themselves as the new normal. Consequently, in the US, the share of hybrid and remote workers has remained steady at roughly 30% since 2023 (Barrero et al., 2025), while in Europe, the adoption of hybrid schedules varies between 20% and 52% in 2025 (Remotly, 2025). In Israel, the most recent official statistics from the end of 2021 suggest that 15% of the workforce is engaged in remote work (Zontag et al., 2022).

The impact of amenities on local remote work

The impact of remote work on residential sorting has been confirmed by previous research: residents who no longer need to regularly commute to the Central Business District (CBD) often relocate to the city outskirts to achieve a better trade-off between price and quality of life (Althoff et al., 2022; Brueckner et al., 2023; Monte et al., 2023). Studies name urban amenities as one of key drivers of this movement, together with house prices. Neighborhoods that offer a better functional substitution for the CBD attract more population and experience a higher growth in rental prices (Biagetti et al., 2024). However, few studies specify exactly which amenities positively impact the growth of remote workers in the neighborhood. Most studies argue that the high concentration of services itself will attract remote workers. For example, Gokan et al(2025) argue that presence of Neighborhood High Streets in remote areas, streets where all local businesses are clustered, define the extent of migration from CBD. They claim that area near such streets is the more preferable area for remote workers because they can balance housing costs and access local amenities. While most studies focus on US and EU case, Tregubova & Gdaliahu also proved the general positive relationships between local variety of amenities and WFH in TAM (Tregubova & Gdaliahu, 2025).

On the other hand, the classical urban economic literature argues that the ultimate welfare impact on local residents depends heavily on their specific preferences, as different demographic groups favor different types of amenities (Almagro & Domínguez-Iino, 2025; Couture & Handbury, 2023). Given that remote work arrangements primarily facilitate the mobility of highly educated individuals (Dingel & Neiman, 2020) rather than wealthy populations, traditional urban amenities that attract the wealthy, such as restaurants, schools, parks, or personal services, may hold less weight for the highly educated. So for policy making purposes more detailed investigations are demanded.

New working spaces

At the same time, there is another group of studies on remote work that claims the importance of so-called ‘new third places’, adapting the classic definition of Ray Oldenburg (1989). In the diversity of urban amenities they highlight the places where people who choose to perform their work outside the office can find the office replacement - in other words, they can comfortably sit there and work without being disturbed from the one hand, and observe people around from the other (Tomaz & Tabrizi, 2024). Among them coworkings, cafés, libraries, and community centers located inside residential areas that provide reliable Wi-Fi and power outlets are the most popular hybrid space options, meaning they serve for both leisure and work (W. Li et al., 2024; Tomaz & Tabrizi, 2024; Yang et al., 2025).

However the existing literature lacks robust quantitative proof of this phenomenon, largely because the empirical identification of remote work visits remains a highly challenging task. Consequently, only a limited number of studies have explored the connection between residents' home locations and third places, and their analyses are typically confined to physical distance while ignoring other important spatial, demographic and economic factors (W. Li et al., 2024). Also no research, to the best of our knowledge, has been done so far to understand to what extent specific third places besides coworkings (W. Li et al., 2024) may explain remote working level at the neighborhood, and what are the preconditions for the success of this process. The closest study that exists so far, that attempts to establish connection between specific types of third places and remote working days, captures only correlation associations and completely ignores the potential endogeneity problem (Yang et al., 2025). Lastly, there is no study yet that would explain what features of amenities let them successfully convert to new third places.

The lack of detailness about the attractors for remote workers limits policy makers in their ability to adapt the city to the new behavior of such residents. At the same time the positive influence of remote workers on the city has been widely acknowledged. Among recognised impacts are decrease of the burden on the commuting system (D. Li et al., 2026), the development of more livable neighborhoods in the suburbia, including better local services and new low-skilled workplaces (Gokan et al., 2025).

As such, this study aims to extend our understanding of new third places: trying to detect specific features that make some amenities popular among remote workers, understand new visiting behavior and test the relationships with local WFH level. Importantly, coworking spaces are out of scope of this research. Because they predominantly serve startup and remote workers lacking a local corporate base, they function as primary office equivalents (second places) rather than Oldenburg's 'third places'.

Description of the Data

GPS signals dataset

Empirical testing of the model requires a method for measuring the dynamics of remote work at the neighborhood level before and after the COVID-19 pandemic. To this end, we utilize a dataset of GPS signals collected via mobile applications in TAM between January 2019 and September 2023. The dataset contains geolocated 'stays' - clusters of consecutive signals occuring at the same location. For each stay we have User ID, the beginning and end of the stay and location centroid. After removing users with less than 4 'stays' per month, the data capture approximately 1% of the TAM population before 05/2020 and 5% after on a monthly basis (Fig1). The daily hours distribution is more stable but still reveals the outliers in the month. Such differences in data volume and signal variation significantly limit our ability to compare pre- and post-COVID periods in the next steps and require us to normalize singals volumes before drawing conclusions about temporal differences in hourly patterns.

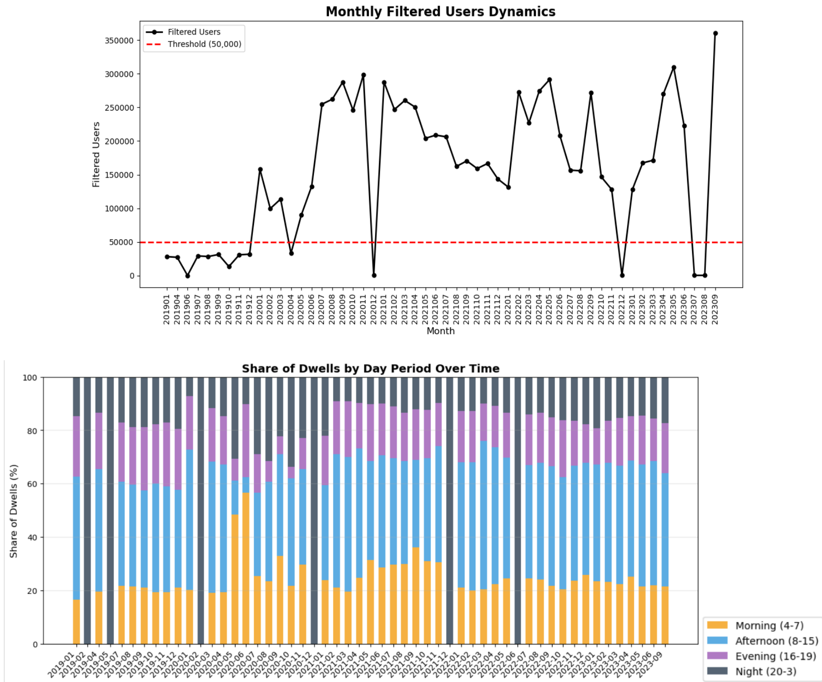


Fig. 1 Monthly temporal statistics of the GPS signals dataset. The top panel shows number of users with at least 4 stays in a month; the bottom panel shows the hourly distribution of stays across the period. Both diagrams show the significant variations in the stay counts suggesting the removal of some months and normalization from further analysis.

<u>POIS Datasets</u>

To map signals to physical locations, we utilize several sources for Points of Interest (POIs), specifically the Google Places API and OpenStreetMap (OSM). Although OSM is a frequent choice in academic research, it exhibits a high degree of spatial discrepancy that can potentially compromise mapping quality. Consequently, to ensure accuracy in identifying third places, we collected POIs from both sources and then select the best candidate for signal location during the later steps. First of all, we extracted 4902 potential third places within the TAM from Google Places (Fig 2) for the following keywords: *bakery, cafe, cafeteria, coffee_shop, tea_house, dessert_restaurant, cat_cafe, dog_cafe, breakfast_restaurant, brunch_restaurant, library, book_store, museum, cultural_center, community_center.* From them we keep 4176 places that have at least one review on Google Maps. Next, we collected POIs indicating work or education locations such as *business_center*, *industrial_area, corporate_office* (Google Places) and *university, school* (OSM). Lastly, we collected data on the rest urban and natural amenities from OSM to detect these stops in users daily trajectories. The full list can be find in the Appendix 1.

.

For each place int Google Places dataet, we retain a unique Google identifier, place name, Google primary place type, location, ranking, the number of reviews, and the data on availability of specific amenities - such as Wi-Fi or restroom. Aditionally, we collect five ‘most relevant’ reviews from Google Maps – the maximum number available through service API. The limitation of the data

Statistical datasets

For Land use classification we use two additional datasets: the layer of Statistical Areas (Sas) from Israel Census 2022 and Israel designated employment zones 2022. Within the boundaries of the TAM, there are a total of 1,223 residential and 101 commercial or institutional Sas as well as 17 employment zones.

Methodology

The identification of third places is the compex task that in the light of uncomplete information in available sources, requires to embrace multiple approaches and sources, and validate them versus each other.

For this matter, the indification strategy contains several steps:

- Defining a sample of candidate office substitutes
- Identifying the sample of *true third places* based on Google reviews
- Extracting visits metrics from the GPS-signals data
- Detecting work replacers in users weekday trajectories using location embeddings
- Building the model on collected metrics that identifies third places

Defining a sample of candidate office substitutes

The sample that we use to find third places candidates is limited by POIs collected from Google Places database as we already selected them based on the list of selected keywords presented in the previous chapter.

Next, we restrict the sample to residential and mixed-use areas, excluding formal workplaces and institutional sites. This approach aligns with existing literature indicating that third places are primarily located near homes outside of commercial hubs. Furthermore, removing such places prevents the potential misclassification of office workers’ daytime visits as remote work from the third place. For this matter we construct the *location_fits_remote_work* flag that indicates if surrounding building environment, official landuse or local POI composition refer to work location nearby.

To identify work zones, we spatially intersect these locations with datasets of dedicated employment zones and Census SAs. A location is designated as a work zone if it falls within an "industrial" land-use area according to Census or an employment zone entirely lacking residential buildings. Furthermore, if a geohash intersects with polygons representing universities, colleges, medical institutions, or transport hubs, it is classified as institutional zone. Lastly, we calculate the number of business centers and corporate offices within a 250-meter radius of the geohash; if this count exceeds nine (falling into the top decile), the area is considered a work zone. Areas containing one to nine of these establishments that are not otherwise classified as work or institutional zones are categorized as mixed-use. Finally, all locations classified as work zones or institutional are marked as unsuitable for remote work (*location_fits_remote_work* = False), the rest we consider candidates third places. Later, we consider only POIs with *location_fits_remote_work* =True.

Identifying the sample of ground truth third places based on Google reviews

The most reliable indicator of remote work in third places within our dataset is user reviews on Google Maps. When individuals mention working on laptops or use synonymous terms in a review, it provides a strong indication of remote work activity at that location. A limitation of this dataset is that the Google Places API does not return a complete list of reviews; instead, it yields a maximum of five reviews deemed "most relevant." Consequently, we are highly likely to miss reviews indicating remote work for many locations, resulting in potential under-identification (low recall). However, the accuracy rate for places that *do* possess such reviews is expected to be high (high precision), allowing us to utilize this subset as a ground-truth validation sample for evaluating other identification methods.

To operationalize this, we derive a text-based indicator *flag_keyword_remote_work*. First, *direct remote-work evidence* is identified using a dictionary of explicit work-related terms and phrases, such as "laptop" or "place to work" (the full list is presented in Appendix 2). A review contributes positive evidence only if it contains at least one such term and doesn't contain exclusion terms that indicate unrelated or contradictory contexts (e.g., "event venue" or "no Wi-Fi" statements; see

Appendix 2). A location is flagged as positive when at least one qualifying review is detected. All observations with *flag_keyword_remote_work* we will use as ground truth.

In addition to the ground truth flag, we extract a couple of more POIs features. First of all we construct another flag, *flag_atmoshere_remote_work*, that captures whether the reviews describe remote work 'vibe'. This indicator leverages keywords reflecting a calm and comfortable social ambiance—such as "quiet", "meeting" or "sitting with others" (see Appendix 2)—while explicitly excluding negative ambiance signals (e.g., "loud", "rush"). A location is designated as atmosphere-suitable if its reviews contain positive ambiance-related evidence without any accompanying negative signals. Next, we extend the sample with amenity availability parameters returned via Google Places API. The limitation of these two flags is that they filled by owners and are missing for about 50% of places.

Identifying 'remote work' visits and deviations in visiting schedule in GPS-signals data

The goal of this stage is to extract weekday visiting metrics for each place based on GPS signals. The process begins by classifying each *stay* as Home/Work/POI/Uknown. This involves a hierarchical spatial assignment: GPS signals are first matched against the user's Home or Work Geohash Spatial Index, and subsequently against a 20-meter POI radius if the initial match fails (detailed in Appendix 3). Home and work locations are recalculated monthly using our previously established frequency-based methodology (Tregubova & Gdaliahu, 2025). This specific radius was selected to balance the inherent inaccuracy of GPS positioning against the risk of misattributing signals to incorrect, neighboring POIs.

Matching the January 2019 to January 2020 signals to individual POIs revealed a very low average visit volume of fewer than 10 (Fig. 2). This reduction is a direct consequence of the fivefold difference in total stays between this period and next years identified earlier. Therefore, for identification task we use only post-COVID period data.

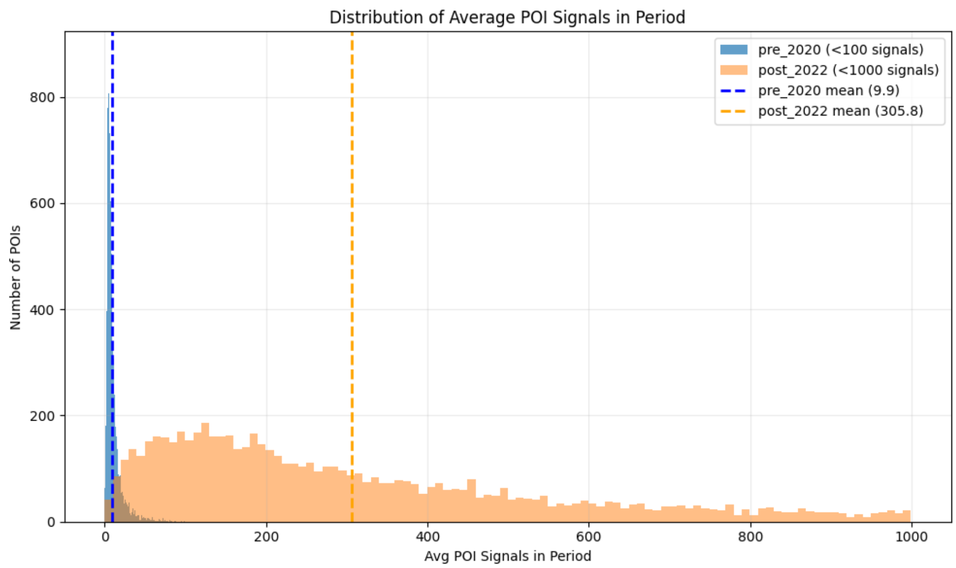


Fig 2. Distributions of average visits per individual POI in two periods: 01/2019-01/2020 and 01/2022-09/2023

Following this spatial attribution, we filter the dataset to keep only POIs with a minimum of ten visits during 01/2022-09/2023. The primary metric calculated is the *average share of remote work visits*. We define a "remote work visit" as a continuous stay lasting more than one hour that occurs between 10:00 AM and 1:00 PM. This classification aligns with remote worker behavioral patterns documented in the literature (Tomaz & Tabrizi, 2024). Specifically, remote workers tend to avoid peak commuting hours, arriving at third places between the morning rush and the lunch peak when venues are less crowded. Furthermore, their visits are typically characterized by longer durations, and having originated from a residential location.

Next, we construct the metrics that represent visitation stability at a Point of Interest (POI) relative to its daily peak: *hourly signal standard deviation* and flag *no_drops_in_daily_dynamic*. The first step is normalizing it against the total signal count to mitigate uneven hours representation inherent in GPS data collection. To mitigate the uneven hourly sampling inherent in GPS data, we first normalize the signal counts at a given POI against the total signal volume recorded during that corresponding hour. Then, the first metric calculates the standard deviation of these normalized hourly averages per POI. The second sets a binary flag to 1 if the minimum visit volume between 10:00 AM and 3:00 PM remains at or above 20% of the peak hour volume. This methodology aligns with literature indicating that alternative workspaces exhibit lower visitation variance than locations driven by discrete events (e.g., meal or coffee rushes), as remote workers typically avoid peak crowds and stabilize daytime foot traffic.

Beyond the described metrics, we extract a comprehensive list of POI-level features capturing localized visit intensity and temporal patterns. As such, we compute the average *visit frequency* per unique user, the absolute volume of weekday visits, the average dwell duration on weekdays, We also calculate the overall share of visits occurring on weekdays versus Fridays to capture overarching temporal distributions.

Finally, for each visit we check users *previous location*, home and work location. From this we derive the distance to user's work and home locations and their ratio, indicator if user came from. To assess the localization of these amenities and their integration into immediate residential neighborhoods, we also calculate the proportion of visits occurring within a one-kilometer radius of the user's home. Together, these temporal and spatial features yield a multidimensional dataset of 11 features describing new mobility patterns in urban amenities visits.

Detecting work replacers in users weekday trajectories using location embeddings

The final indicator is the most comprehensive - it reveales not only a specific location's potential to replace a formal office but also broader shifts in resident behavior. For this analysis, we utilize the previously generated dataset of anchored stays, where each instance is classified as home, work, POI, or unknown.

Detecting office substitutes requires constructing location embeddings: vector representations of each place derived from its sequential position within daily mobility trajectories (Esuli et al., 2018; Murray et al., 2023; Weng et al., 2025). Locations sharing similar trajectory contexts—meaning they are frequently preceded and followed by the same spatial sequences—yield mathematically proximate embeddings. As such, third places that play role of office substitutes should have emeddings similar to established work locations. we measure similarity using cosine distance (Crivellari & Resch, 2022).

We construct embeddings in three steps: first we create vocabularly of 'tokens' – desciptions of each places, then we build dataset of individual mobility trajectories, and last we train on them a Word2Vec Natural Language Processing (NLP) model (Mikolov et al., 2013) to create location embeddings.

Each location is encoded into a discrete 'token' two times based on two groups of encoding rules: one for aggregated naming and one for individual naming (Table 1). The first group of rules is applied to both pre- and post-COVID periods and is used for descriptive purposes: to identify changes in users behavior. The second dictionary of tokens we apply only to post-COVID data in order to calculate cosine distance between individual POI location in different time of the day and ‘work’.

For each vocabularly of ‘tokens’ we create a dataset of trajectories. Each trajectory represent chronologically ordered sequences of places (‘tokens’) visited by a user within a specific day. To ensure the robustness of this movement representation, we enforce a minimum stay duration of 10 minutes and a maximum temporal gap of three hours between visits. If the time gap exceeds three hours, the trajectory is terminated, and a new one is initiated. In the end, we retain only trajectories containing a minimum of three locations.

Table 1. Encoding rules for different types of locations

	Home or Work	**POI**	**Unknown location**
Aggregated locations	‘Home’/‘Work’	‘POI’+[POI category]	‘GH’ + [Day Period];
Individual locations	‘Home’/‘Work’ + [Statistical Area ID]	‘POI’ + [Unique ID] + [Day Period].	[Geohash Index] + [Day Period]

Following this procedure, we yield a final dataset of 5,744,898 ‘sentences’: 855,538 in pre-COVID period and 4,889,360 in Post-Covid. Aggregated version of tokens comprises 42 unique location tokens with an average frequency of 25,972 appearances in Pre-Covid period and 134,574 in Post-Covid period. While for the size of tokens with unique identifiers is 57,503 with an average frequency of 99.8 appearances in post-covid period.

Leveraging these discrete trajectories, we train a Word2Vec model using the Gensim Python library. Guided by literature best practices and scaled to the size of our dataset, we apply the following model parameters:

- Architecture: skip-gram
- Embedding dimensionality: 100
- Minimum token frequency: 5
- Negative sampling: 15 negative samples per positive pair
- Training epochs (passes over the corpus): 10

We execute the model 3 times: two times for aggregated tokens in two periods and one with individual tokens for post-Covid period. Then, for each model separately, we identify POIs that are semantically closest to "work" locations by calculating their cosine similarity. This metric mathematically translates associative strength into the angle between normalized vectors within the embedding space. Last, to account for the average differences in similarity to 'work' between POIs categories, we replace original cosine similarity of the place with its ranking inside the category.

The obtained results are utilized in two ways: first, we append the rank of cosine similarity metric calculated on individual tokens to our database of third places; second, we visualize the token embeddings from each model and the most frequent 3-grams (three consequatively visited locations) to descriptively illustrate the differences between two periods and conclude about emerging patterns in post-COVID era.

Identification of third places

Based on the collected dataset of third places with flags and metrics we run the identification algorithm. First, we eliminate POIs with more than 50% of missing values. Then, we formulated the third places identification as a Positive-Unlabeled (PU) learning problem. This approach is the binary classification task where instead of negative sample we have unlabelled which means we don't know if it is negative or postivie. It suits us because we have limited number of Google Reviews per place so we don't know if no remote work is happening there in reality or we just don't observe the reviews indicating it. Following Google API description we also assume the randomness of review selection, meaning that for each real third place chances to have in our dataset relevat review are the same. This assumtion allows us to use classifier with the adjustment methodology proposed by Elkan and Noto (Elkan & Noto, 2008). Following their paper, we train the classifier to estimate the probability that an observation is labeled positive, $P(s=1|x)$ and then, assume that this quantity is proportional to the target probability $P(y=1|x)$ with constant c. So the final probability is computed as: $P(y=1|x) = P(s=1|x)/c$.

Given the high dimensionality and likelihood of multicollinearity among the features, we utilized an Elastic-Net Logistic Regression model as classifier to mitigate coefficient instability. In order to interpret derive feature strength from the value of coefficients, we standardize the metrics using StandardScaler method.

Since the groundtruth negative examples are unavailable, we can't use neither the classic confusion matrix, nor standart cross-validation to evaluate the model's quality and stability. Instead, since c is constant and the ranking of observations by predicted score is preserved, we evaluate the model using repeated positive-holdout procedure. In each repetition, we randomly hide 30% of the trusted positive places by changing their observed label from positive to unlabeled, train the Elkan-Noto PU model on the remaining labeled positives and the full unlabeled pool, and then rank all candidate places by the model's estimated probability of being positive. Model quality is evaluated by whether the hidden positives were recovered near the top of this ranking, using *recall-at-k* measures and the median percentile rank of hidden positives. Such validation approach is standard and used for evaluating PU model without known negative samples by many papers (Saunders & Freitas, 2025). The choice of recall for ranking is explained by the higher priority of minimizing False Negatives rather than False Positives. We aim to identify as many third places as exist while allow small amount of 'wrong third places' caused by signals similar to remote work but coming from unemployed or retired people.

Additionally, we test the model sensitivity to different hyperparameters such as regularization coeffient C iterating it over [0.03, 0.1, 0.3, 1] and *l1-ratio* - the share of how much of the penalty is Lasso vs L2 Ridge iterating it over [0.25, 0.5, 0.75].

<u>Inspecting relationships between WFH, amenities and third places</u>

After identifying third places, we test the hypothesis of a stronger association between these locations and WFH than between general local amenities and WFH. To answer this question, we pool quarterly levels of WFH with the density of local amenities and detected third places at the SA level. Using this dataset, we run independent OLS regressions for each quarter, controlling for the distance to the center of Tel Aviv and the number of homes detected via GPS.

The construction of quarterly WFH metric by SA is explained in details in the previous research (Tregubova & Gdaliahu, 2025). Briefly, it calculates the individual's share of days with home signals during working hours, and then averages these values by SA. The density of local amenities and third places we estimate by creating a 1-km buffer around each SA, corresponding to a 15-minute walk in an urban area. This buffer ensures all neighborhoods capture amenities within walking distance. We then calculate the density of each amenity category as the count of POIs from both OSM and Google Places divided by the total area (neighborhood area + buffer area). For third places, we calculate two metrics: the average raw probability of a location functioning as a third place, and the absolute count of such venues derived from the model's binary classification. We estimate the impact of each of them independently in the regressions:

$$WFH_{ij} = \alpha + \beta_1 T_i + \beta_2 A_i + X_i (1)$$

Where T is one of two metrics associated with third places that we test, A – density of amenities within 1 km, X – list of control variables for specific SA.

Before runnig the regressions we normalize using min-max scaler and check mulitcollinearity issues: we expect all features to have mutual correlation below 0.8

Results

<u>Descriptive results</u>

In the original sample of candidate work substitutes from Google Places with at least one review we identified 662 located in workzones, inside institutions or big transport hubs. They were marked as impossible for remote work and excluded from the model sample. Therefore, the final sample contains 3524 candidates.

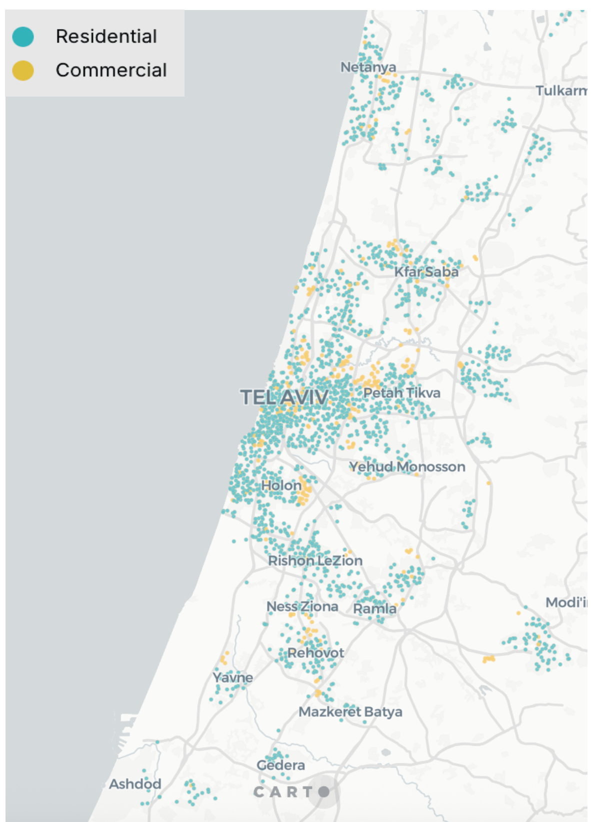


Fig.3. Spatial distribution of third place candidates in TAM, categorized by office and residential locations

In the sample we found 100 places with direct keywords from the reviews indicating remote work. The biggest number of ground truth third places belong to coffee shops (41 places) and cafes (34), then libaries - 12 places, cultural centers and bakeries – 6 places each, community centers – 4 and 1 bookshop. Keywords showing the right atmosphere for remote work were found in 532 places.

Visits distribution analysis

The descriptive analysis of the metrics collected based on GPS-signals show their potential to distinguish third places. Figure 3 illustrates the hourly dynamics of long visits by comparing two Tel Aviv cafes: a bakery lacking remote work amenities (left panel) and a spacious courtyard café highly conducive to laptop work (right panel). While both reside in mixed-use areas near—but not directly adjacent to—established workplaces, their visitation patterns diverge significantly. "In the left diagram, the absolute number of long visits remains stable—accounting for less than 30% of total visits until the 5:00 PM end of the workday—with a relative peak occurring during the 12:00 PM to 1:00 PM lunch hour. Conversely, the right diagram exhibits a marked increase in long visit counts between 10:00 AM and 3:00 PM, culminating at 2:00 PM when over half of all visits exceed one hour.

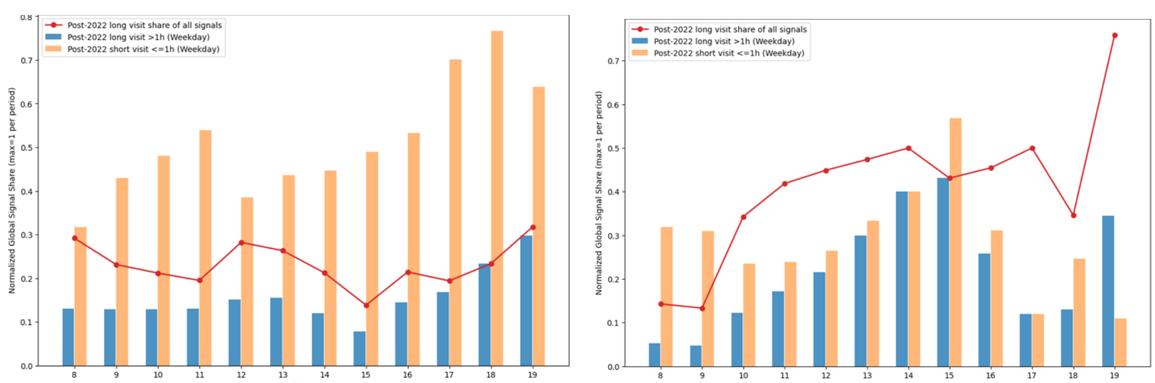


Fig. 4. Comparison of normalized visit counts between two locations. A visit is attributed to a specific hour if it overlaps with that hour (i.e., started before or during, and ended during or after). Presented values are normalized against total signals at specific hour and the maximum hourly visit count recorded within the 01/2022–09/2023 study period ar specific location.

Daily trajectories analysis

The analysis of daily trajectories constructed based on GPS signals revealed several shifts in users behavior after COVID-19 that may be indication of remote work presence.

First of all, top 3-grams constructed from aggregated token reveal two new popular trajectories in post-COVID time: Home->Home->Home (2.4% of trajectories) and Home->Park during weekday late morning -> Home (0.7% of trajectories) (Fig.5, right panel).

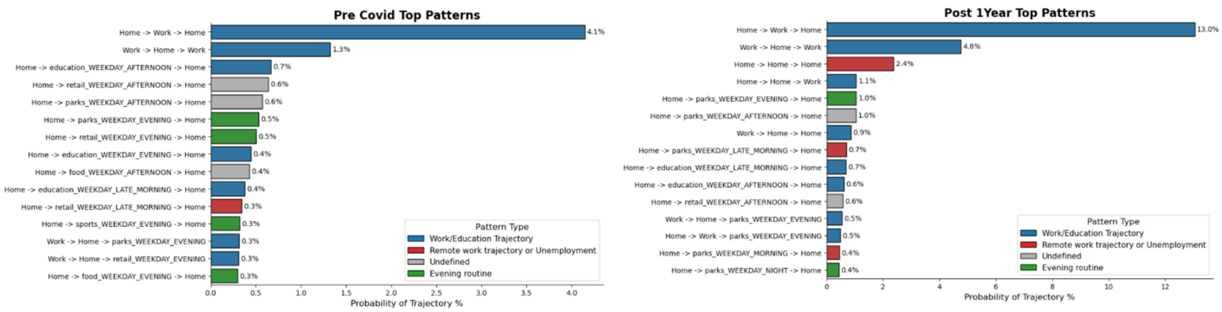


Fig. 5 Top 3-grams extracted from the residents daily trajectories, comparing the 01/2019–01/2020 and 01/2022–09/2023 periods, colored by the type of the trip

Additionally, we notice that residents of CBD area (within 5 km from Tel Aviv Azrieli center) have started to show a unique new behavior pattern: Home -> food place during late morning hours - > Home (0.27%) (Fig.6). This may signals for remote work from the third places.

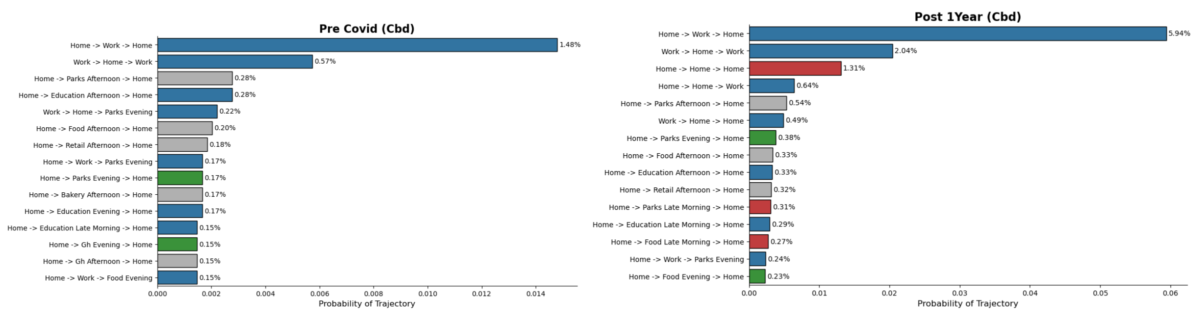


Fig.6 Top 3-grams extracted from the daily trajectories of residents whose detected home location falls within 5 km of the Azrieli Center, comparing the 01/2019–01/2020 and 01/2022–09/2023 periods, colored by the type of the trip

The Word2Vec model results show how close POI locations are to work locations in terms of their positions within daily trajectories before and two years after COVID-19. The identified list of places semantically closest to 'work' in 2019 includes transition points where people stop on the way to work, as well as public and religious places. As such, Donut Shops and Transition points have a cosine similarity of more than 50%, while Community Centers, Religious places, and Public Services have more than 40%. One unexpectedly high place is Entertainment spots, which also indicate a similarity higher than 40% (43.1%).

The list of places semantically close to 'work' in 2022 differs from the previous list. It includes educational places (59%), medical institutions (51%), and parks and sports locations (55%), as well as places suitable for remote work: Coffee Shops (55%) and Bakeries (54%). A t-SNE representation of the embeddings on aggregated tokens, built using https://projector.tensorflow.org/, is presented in Appendix 4.

Final sample

After putting all metrics together, we end up with 27 features describing each of the 3,524 locations, where each location has values for at least 50% of the features. Our dataset contains columns with missing values, which we process using different strategies (Appendix 5). Because we cannot determine with high certainty why certain locations lack sufficient signals for inclusion in the trajectories, we apply two parallel imputation strategies—mean substitution and zero imputation—allowing the model to determine which approach holds more predictive value. Additionally, we introduce a missing value indicator flag to capture any potential significance associated with the absence of data itself.

Model results

Evaluating the model over 100 runs with 30% of the positives hidden and C=0.1 and l1 ratio=0.5 revealed an average median rank in the top 11%, with a worst-case performance bounded within the top 18% (Fig.8)

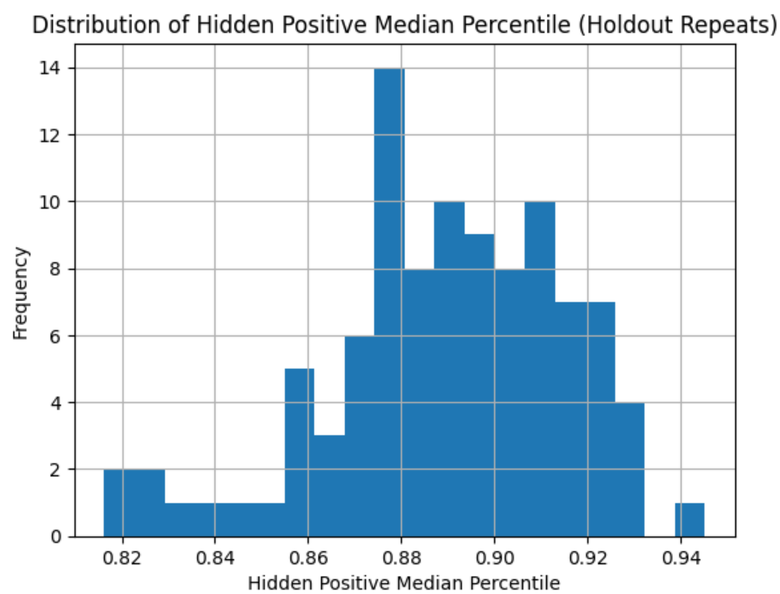


Fig.7 Distribution og Hidden Positive Median Percentile on 100 iterations with 30% of the hidden positive.

By averaging the coefficients obtained at each iteration, we reveal the model's strongest features (absolute coefficient value > 0.1). The results are presented in Figure 9. Six of these come from Google Places, with keywords indicating atmosphere being the strongest, while two features represent cosine similarity with work during morning hours and the share of 'remote work' visits during weekdays. Important to note, that all of them were different from zeros at least in 95% of iterations.

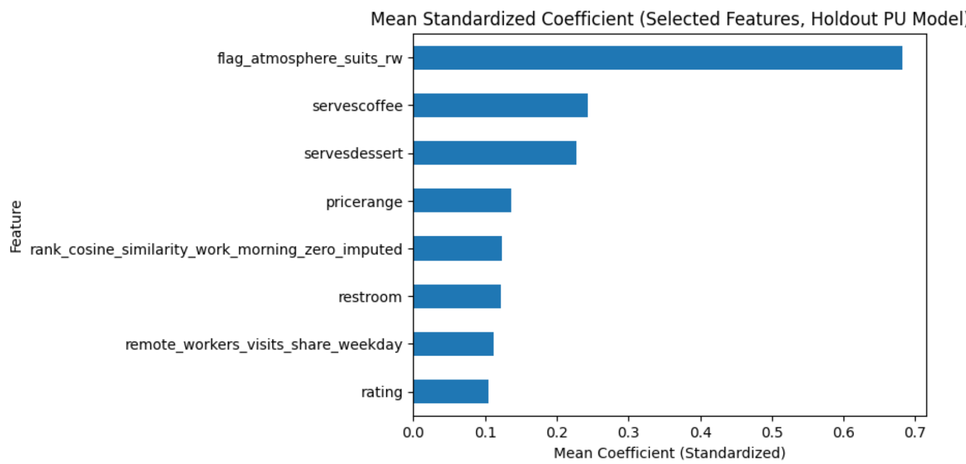


Fig.8 Average feature importance on 100 repetitions with 30% of hidden groundtruth

Then, the iteration over the grid of hyperparameters helped us find the best combination, while generally demonstrating that the results are not highly sensitive to their choice. As such, ranking differences exist in the top 20% to 40% range and are fully eliminated by the top 50%, where all specifications demonstrate a 95% recall level. The optimal model configuration for the top 20% of the ranking relies on stronger regularization with an equal contribution from Ridge and Lasso penalties (C=0.03, L1 ratio=0.5). At the top 20%, it reaches a recall of 77%. This configuration we use to produce the final identification.

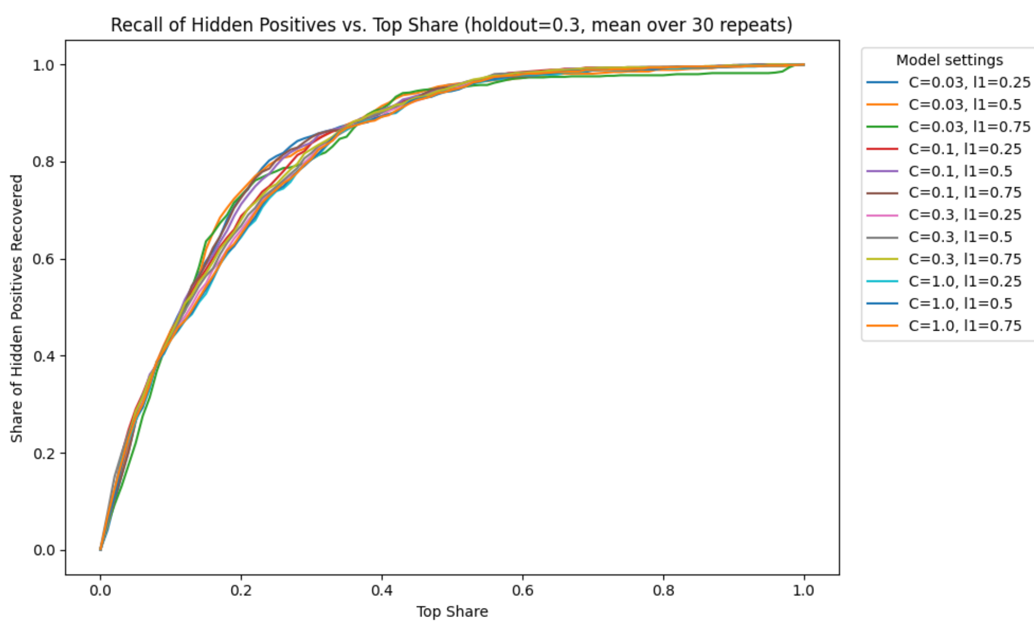


Fig. 9 Recall of hidden positives by top share ranking across evaluated hyperparameter configurations (mean over 30 repeats, 0.3 holdout). Each line represents the results for unique combination of hyperparameters. Model performance exhibits low sensitivity to parameter variations, with all

specifications converging to 95% recall at the top 50% threshold.

As detailed in Appendix 6, we evaluate the model excluding Google Places attributes. The resulting drop in classification quality confirms that successful third-place detection requires combining intrinsic venue characteristics with human visitation patterns.

Detected third places.

The model detected 550 third places, which represents 15.6% of the total candidate sample. The highest shares are observed in the central Tel-Aviv area—22% in the CBD and 19.7% in residential areas—while in satellite cities, we identified 12% of candidates as real third places, regardless of how active the neighborhood is. Therefore, these results suggest that in general the chances of finding work substitutes are higher in central areas rather than suburbia in TAM.

Categorically, we observe that the absolute values are highest among coffee shops (174 places) and cafes (128 places), while the share of places suitable for remote work is highest for libraries.

Table 3

primarytype	**total_places**	**model identified places**	**% of total**	**groundtruth count**	**groundtruth identified by model**	**recall**
coffee_shop	581	174	29.95%	41	29	70.73%
cafe	621	128	20.61%	34	17	50.00%
bakery	1115	66	5.92%	6	2	33.33%
cultural_center	392	62	15.82%	6	6	100.00%
library	145	50	34.48%	12	12	100.00%
community_center	243	27	11.11%	0	0	NaN
book_store	246	20	8.13%	1	1	100.00%
museum	126	15	11.90%	0	0	NaN
breakfast_restaurant	11	3	27.27%	0	0	NaN
dessert_restaurant	15	2	13.33%	0	0	NaN
tea_house	23	2	8.70%	0	0	NaN
dog_cafe	3	1	33.33%	0	0	NaN
brunch_restaurant	2	0	0.00%	0	0	NaN

We also note that the model detects remote work behavior in categories absent from the ground truth sample, such as museums, community centers, and specific niche establishments like breakfast and brunch restaurants, dessert shops, and dog cafes. The model exhibits the highest level of noise within the bakery category, where only 2 of the 6 locations from the ground truth were correctly identified. Spatially, while the absolute number of identified third places is highest in Tel-Aviv, the highest proportional shares are found in satellite cities such as Kfar Saba, Raanana, Kiryat Ono, and Rehovot.

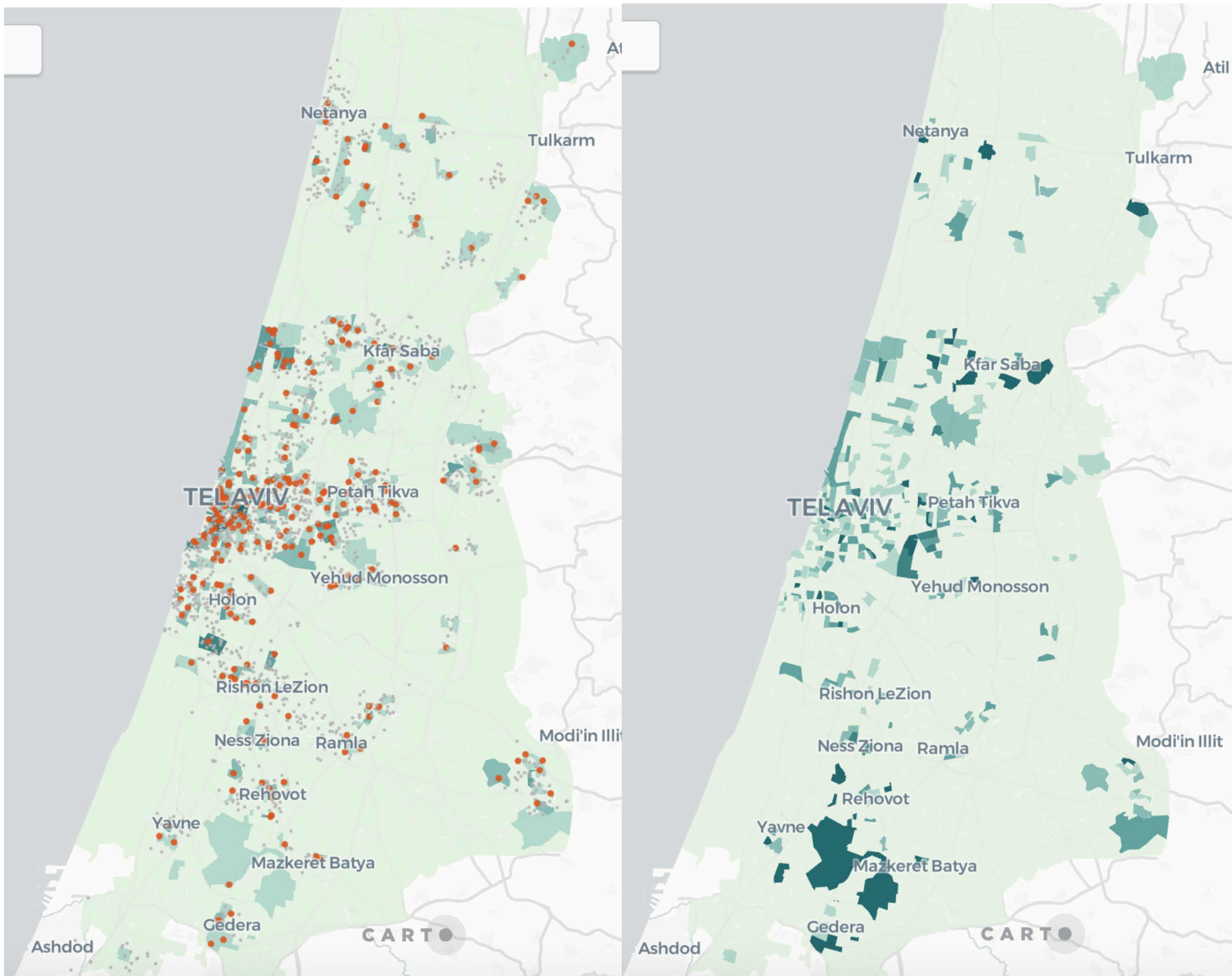


Fig. 10 **Distribution of identified third places (absolute counts)** **Distribution of identified third places (percentages)**

According to the distribution of identified third places by POI category and neighborhood type (Figure 14), coffee shops are the most popular option for remote work across all neighborhoods in absolute terms. Interestingly, however, the highest share of coffee shops utilized for remote work is found in central residential areas rather than the CBD (36.2% vs. 33.5%), where one might typically expect the greatest concentration of such venues. Furthermore, residents of quiet suburbs choose libraries and cultural centers much more frequently than residents of other areas (35 locations versus 10 elsewhere), while the percentage of utilized libraries is highest in the centers of satellite cities.

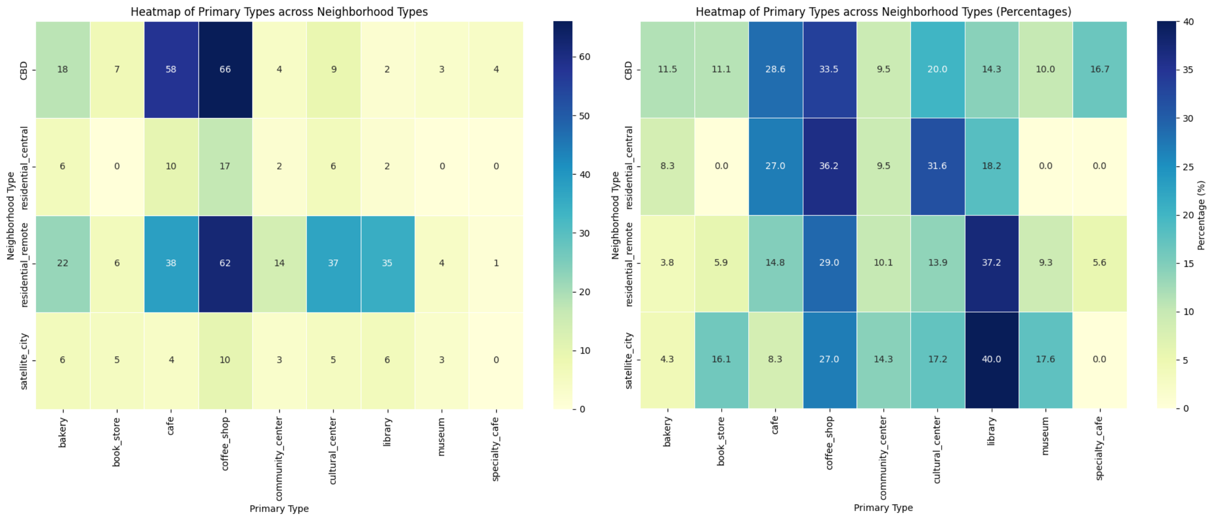
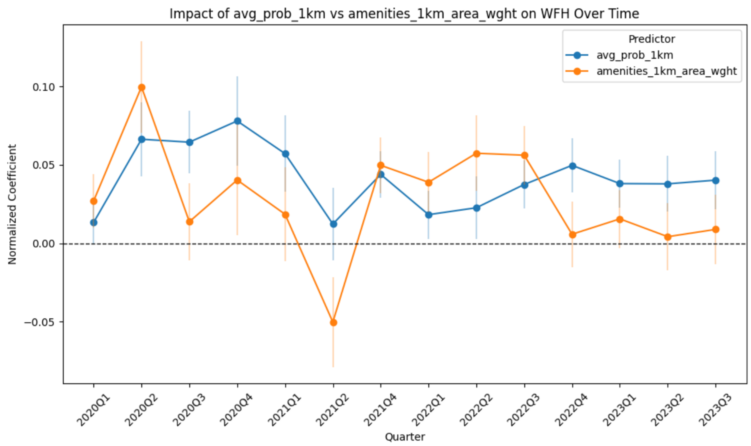


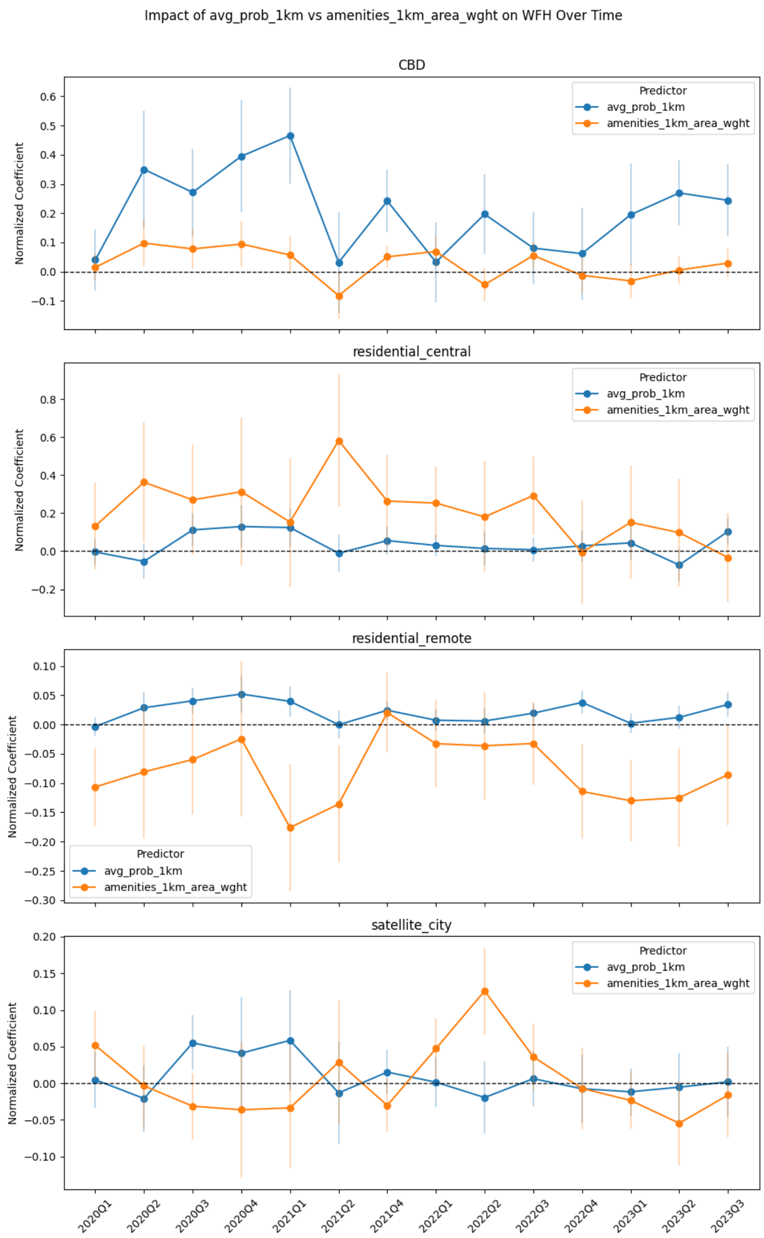
Fig. 14

<u>OLS regressions results</u>

First, the mulitcollinearity check didn’t reveal the problem: the highest correlation 0.41 found between density of amenities and average probability for being a third place is below the threshold – 0.8.

At the same time, the results of running Eq.2 shows that the coefficient before average probability in the neighborhood is significant and positive all the quarters except 2020Q1 and 2021Q2. Moreover it mostly exceeds the coeffecent of all total amenities, the exception is the period 2021Q4-2022Q3. After the 3rd quarter the probability for third places shows stable impact near 0.05 while total number of amenities is insignificant . However if we build the same model for remote and central areas based on users’ home locations independently we only notice the significan effect for CBD. In residential central in opposite general amenities have higher positive and significant effect which may be explained by distance to park and beaches that are not office substitutes but can also be used during workday.


Impact of avg_prob_1km vs amenities_1km_area_wght on WFH Over Time
Predictor
avg_prob_1km
amenities_1km_area_wght
Normalized Coefficient
Quarter
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q4
2022Q1
2022Q2
2022Q3
2022Q4
2023Q1
2023Q2
2023Q3


Impact of avg_prob_1km vs amenities_1km_area_wght on WFH Over Time
CBD
residential_central
residential_remote
satellite_city
Predictor
avg_prob_1km
amenities_1km_area_wght
Normalized Coefficient
Quarter
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q4
2022Q1
2022Q2
2022Q3
2022Q4
2023Q1
2023Q2
2023Q3

Discussion of the results

The Functional Utility of Amenities

While previous literature identifies amenities as drivers of residential sorting among remote workers (Biagetti et al., 2024; Gokan et al., 2025), it often conflates specific functional needs with general commercial density. Our findings challenge the assumption that proximity to high-concentration service corridors, such as Neighborhood High Streets, is the primary spatial attractor for the WFH demographic. Instead, we demonstrate that the presence of dedicated "third places"—venues actively functioning as office substitutes—exhibits a stronger, more resilient association with localized WFH patterns than generic amenity density. This indicates that functional utility, rather than mere commercial proximity, fundamentally drives the modern geography of remote work.

Behavioral Trajectories and Methodological Innovations

Capturing this functional reality requires moving beyond static point-of-interest (POI) density models. By explicitly incorporating localized mobility patterns through advanced NLP and trajectory analysis, this study retheorizes the traditional role of commuting in urban development. We pioneer a methodological bridge by applying Word2Vec embeddings to physical mobility, demonstrating how the semantic functions of physical spaces continuously adapt to macro-behavioral trends. This approach reveals that third places seamlessly integrate into daily commuter-like sequences, functionally replacing the traditional Central Business District (CBD) destination. While the long-term urban fallout of the COVID-19 pandemic remains fluid, our research successfully unravels the complex spatio-temporal dynamics of this evolving home-third place relationship.

Localized Sorting in the TAM Context

Within the Tel Aviv Metropolitan (TAM) area, these dynamics provide a nuanced spatial resolution to the established 15% national WFH baseline (Zontag et al., 2022). Rather than a uniform dispersion to the urban periphery, residential sorting among hybrid and remote workers exhibits distinct micro-clustering around functional third places. This localized sorting mechanism refines current models of post-pandemic migration, illustrating that remote workers do not entirely abandon the need for workspace infrastructure; rather, they decentralize it into their immediate residential environments.

Actionable Policy Implications for Urban Resilience

Beyond theoretical advancements, this study yields direct, actionable policy implications for municipalities seeking to enhance post-pandemic urban resilience. Because emerging third places are highly utilized yet unevenly distributed, identifying their spatial distribution and the micro-level drivers of destination choices provides critical empirical evidence for local planning. We directly meet decision-makers' needs by operationalizing these complex econometric findings through an accessible, interactive dashboard and mapping tool.

Specifically, linking visitor profiles with location externalities equips planners with the precise data needed to tailor environments to the diverse needs of individual neighborhoods. Understanding these localized choice drivers dictates where additional public service investments are most critical. Ultimately, this guidance allows municipalities to prioritize targeted infrastructure—such as the safe pedestrian walkways and cycle lanes essential for actualizing a "15-minute city"—ensuring these highly valued localized workspaces remain accessible and continue to sustain modern urban vitality.

Bibliography

Almagro, M., & Domínguez-Iino, T. (2025). Location Sorting and Endogenous Amenities: Evidence From Amsterdam. *Econometrica*, *93*(3), 1031–1071. https://doi.org/10.3982/ecta21394

Althoff, L., Eckert, F., Ganapati, S., & Walsh, C. (2022). The Geography of Remote Work. *Regional Science and Urban Economics*, *93*, 103770. https://doi.org/10.1016/J.REGSCIURBECO.2022.103770

Brueckner, J. K., Kahn, M. E., & Lin, G. C. (2023). A New Spatial Hedonic Equilibrium in the Emerging Work-from-Home Economy? *American Economic Journal: Applied Economics*, *15*(2), 285–319. https://doi.org/10.1257/APP.20210190

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

Crivellari, A., & Resch, B. (2022). Investigating functional consistency of mobility-related urban zones via motion-driven embedding vectors and local POI-type distributions. *Computational Urban Science*, *2*(1), 19. https://doi.org/10.1007/s43762-022-00049-8

Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, *189*, 104235. https://doi.org/10.1016/J.JPUBECO.2020.104235

Elkan, C., & Noto, K. (2008). Learning classifiers from only positive and unlabeled data. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 213–220. https://doi.org/10.1145/1401890.1401920

Esuli, A., Petry, L. M., Renso, C., & Bogorny, V. (2018). *Traj2User: exploiting embeddings for computing similarity of users mobile behavior*.

Gokan, T., Kichko, S., Matheson, J. A., & Thisse, J.-F. (2025). *Teleworking, local labor markets, and the structure of cities. **.

Li, D., Zheng, Y., Wang, S., Guo, X., & Zhao, J. (2026). Quantifying the nonlinear causal impact of commute time on U.S. remote work. *Transportation Research Part D: Transport and Environment*, *151*, 105153. https://doi.org/10.1016/j.trd.2025.105153

Li, W., Zhang, E., & Long, Y. (2024). Unveiling fine-scale urban third places for remote work using mobile phone big data. *Sustainable Cities and Society*, *103*, 105258. https://doi.org/10.1016/j.scs.2024.105258

Monte, F., Porcher, C., & Rossi-Hansberg, E. (2023). *Remote Work and City Structure*. https://doi.org/10.3386/w31494

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2023). Unsupervised embedding of trajectories captures the latent structure of scientific migration. *Proceedings of the National Academy of Sciences*, *120*(52). https://doi.org/10.1073/pnas.2305414120

Ray Oldenburg. (1989). *The Great Good Place: Cafés, Coffee Shops, Bookstores, Bars, Hair Salons, and Other Hangouts at the Heart of a Community.* (2nd ed.). Marlowe&Company.

Saunders, J. D., & Freitas, A. A. (2025). Automated machine learning for positive-unlabelled learning. *Applied Intelligence*, *55*(12), 875. https://doi.org/10.1007/s10489-025-06706-9

Tomas Mikolov, Kai Chen, Greg Corrado, & Jeffrey Dean. (2013). Efficient Estimation of Word Representations in Vector Space. *ArXiv: 1301.3781*. https://doi.org/10.48550/arXiv.1301.3781

Tomaz, E., & Tabrizi, H. A. (2024). *The Evolution of Non-traditional Workplaces: From Third Places to Hybrid Places* (pp. 7–20). https://doi.org/10.1007/978-3-031-50868-4_2

Tregubova, I., & Gdaliahu, M. (2025, August). *Estimating the impact of working from home on urban equilibrium: neighborhood scale effects using mobile data*. https://www.researchgate.net/publication/395945061_Estimating_the_impact_of_working_from_home_on_urban_equilibrium_neighborhood_scale_effects_using_mobile_data

Weng, G., Kim, M., Ahn, Y.-Y., & Moro, E. (2025). *Beyond Distance: Mobility Neural Embeddings Reveal Visible and Invisible Barriers in Urban Space*.

Yang, M., Zheng, Y., Wang, S., Guo, X., & Zhao, J. (2025). Impacts of the built environment on remote work choices and preferences. *Cities*, *163*, 106034. https://doi.org/10.1016/j.cities.2025.106034

Zontag, N., Madhala, S., & Bental, B. (2022). *Working From Home in Israel*. www.taubcenter.org.il

Appendix 1

POIs categories	POIs keywords
Health	hospital, clinic, pharmacy, dentist, doctors
Education	school, university, college, kindergarten
Entertainment	cinema, theatre, nightclub, arts_centre, community_centre, library, museum
Food	bar, cafe, restaurant, fast_food, pub, food_court
Outdoor attractions	attraction, memorial, ruins, artwork, fountain, archaeological
Parks and sport	park, pitch, sports_centre, stadium, playground, swimming_pool, dog_park,viewpoint
Retail	supermarket, mall, convenience, clothes, pharmacy, bakery, electronics, furniture, books, shoes, jewelry, sports, toys, optician, stationery, department_store, florist, kiosk, car_dealership
Services	bank, hairdresser, nails, spa, massage, tattoo, photography, travel_agency, bicycle_repair, car_repair, car_wash, post_office, bicycle_rental

Appendix 2
Exact direct remote-work keyword list used in the review matching:

- "laptop",
- "computer",
- "business meeting",
- " place for work",
- " place to work",
- " get the work done",
- "remote work"

And the exclusion list (reviews containing these are filtered out):

- events
- funeral
- repair
- wifi doesn't work
- on the go
- doctor
- wifi only
- no wifi

Keywords describing atmosphere suitable for remote working used to create a metric flag_atmosphere_keyword:

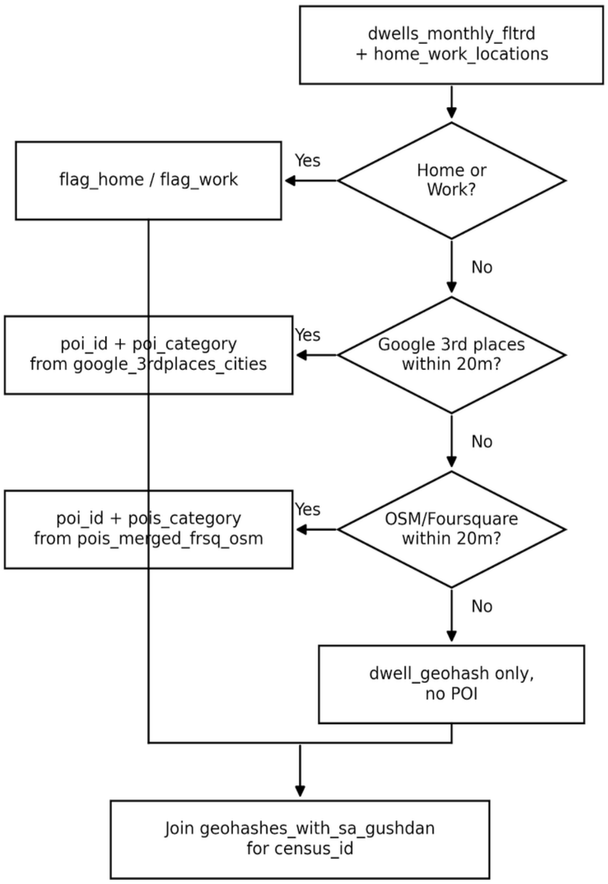
- " peace"
- " quiet"
- " relaxing"
- "relax"
- " comfortable"
- " calm"
- "meeting friends"
- "sit with"
- "sitting with"
- "wifi"
- "wi-fi"
- "charger"
- "socket"

And the exclusion list:

- "not quiet"
- "rush"
- "loud"
- "not comfortable"
- "not relaxing"
- "not calm"

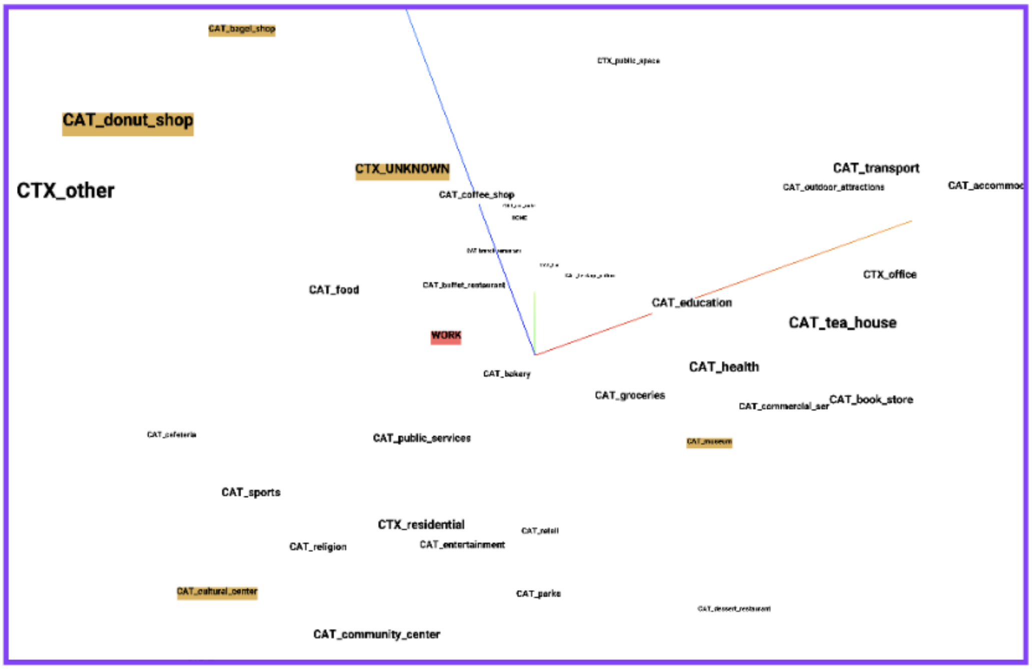
Appendix 3

Anchoring GPS stays with Home, Work, POIs locations

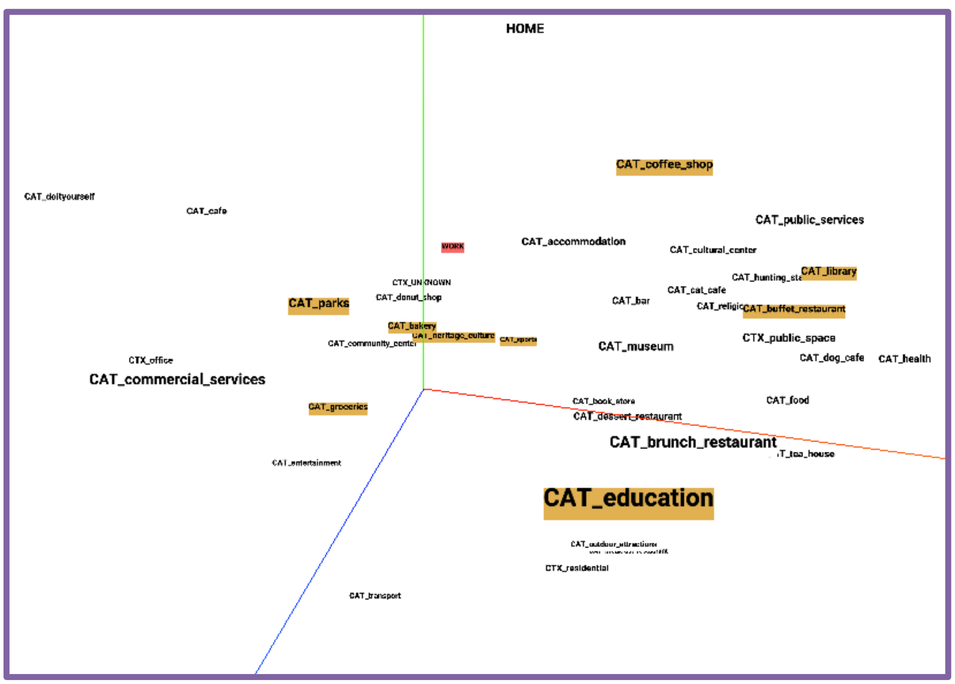


Appendix 4

T-SNE visualization of the Word2Vec embeddings build on pre-treatment sample. Locations that are semantically related to the 'work' are highlighted within the vector space



Before Covid-19



In 2022-2023

Appendix 5

Feature name	Datasource	null_share	Replacing strategy
rank_cosine_similarity_work_morning	Word2Vec results	76.3%	0.5 or 0, and flag missing value
pricerange	Google Places API	71.8%	average
livemusic	Google Places API	71.4%	0
servesdessert	Google Places API	66.2%	0
rank_cosine_similarity_work_late_morning	Word2Vec results	53.7%	0.5 or 0, and flag missing value
restroom	Google Places API	52.2%	0
servescoffee	Google Places API	50.5%	0
avg_rw_visits_share_after2022	GPS-signals analysis	46.5%	0
rank_cosine_similarity_work_afternoon	Word2Vec results	46.3%	0.5 or 0, and flag missing value
visit_frequency_rw	GPS-signals analysis	35.8%	0
remote_workers_visits_share_weekday	GPS-signals analysis	35.8%	0
avg_dist_home_rw	GPS-signals analysis	35.8%	average
share_dwells_home_lt_1km_rw	GPS-signals analysis	35.8%	average
avg_duration_hours_weekday_rw	GPS-signals analysis	35.8%	average
n_visits_weekday_rw	GPS-signals analysis	35.8%	0
n_visits_weekday	GPS-signals analysis	35.8%	0
avg_duration_hours_weekday	GPS-signals analysis	35.8%	average
share_friday	GPS-signals analysis	35.8%	average
share_weekday	GPS-signals analysis	35.8%	average
share_dwells_home_lt_1km	GPS-signals analysis	35.8%	average
visit_frequency	GPS-signals analysis	35.8%	0
avg_dist_home_work_ratio	GPS-signals analysis	35.8%	average

avg_dist_home	GPS-signals analysis	35.8%	average
share_visits_after_work	GPS-signals analysis	35.8%	average
share_visits_after_home	GPS-signals analysis	35.8%	average
min_norm_share_10_15	GPS-signals analysis	14.5%	average
max_n_signals	GPS-signals analysis	12.6%	average
hourly_signal_std_percentile_by_keyword	GPS-signals analysis	12.6%	average
hourly_signal_std	GPS-signals analysis	12.6%	average
offices_proximity	Google Places API	0.0%	
flag_keyword_review_rw	Google Places API	0.0%	
userratingcount	Google Places API	0.0%	
rating	Google Places API	0.0%	
flag_atmosphere_suits_rw	Google Places API	0.0%	
flag_rw_visits_share_after_2022	GPS-signals analysis	0.0%	
flag_office_substitute_daily_dynamic	GPS-signals analysis	0.0%	

Appendix 6

To test the pure ability of GPS signal data to detect third places, we remove all features obtained from the Google Places API and rerun the tests.The average median percentile of hidden positives drops from the top 12% to the top 35%, with the lowest median equal to 55%.

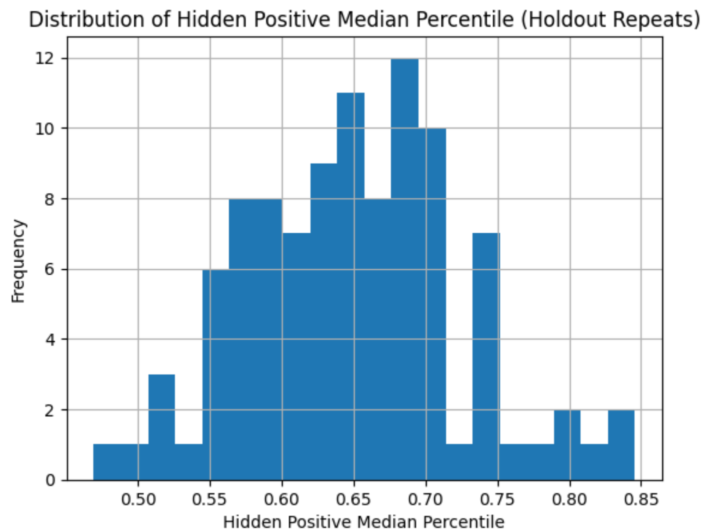


Fig.1

The top two features were also present in the first version of the model, indicating model stability, while four new features emerged: the visit duration of identified remote workers, a flag indicating if the lowest number of visits between 10 am and 3 pm represents at least 20% of the daily maximum, distance to home, and the share of weekday visits. Unexpectedly, the distance to home exhibits a direction opposite to what was anticipated: while the literature suggests that third places tend to be closer to homes in residential neighborhoods, the model consistently shows a higher probability for places requiring a longer commute. At the same time, although not among the top features, the share of remote worker visits where home is less than 1 km away also shows a positive association; however, it is weak, as its coefficient is not significantly different from 0 in 39% of iterations. The distance to home for identified remote workers is not significant at all—in 97% of tests, the coefficient is zero.

The cumulative recall diagram has an almost linear shape, further suggesting a weaker ability of the model specification without Google Places metrics to identify remote work substitutes. While 60% of hidden positives are covered by the top 40% of the data, covering 80% requires the same share of data. In this case, the choice of hyperparameters did not affect the results at all.

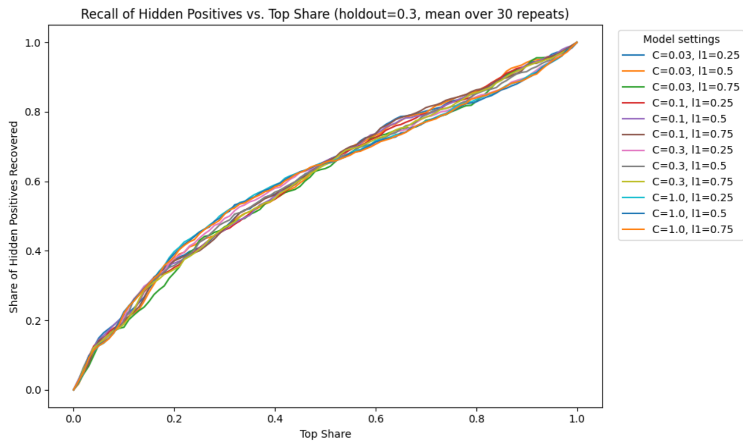

Recall of Hidden Positives vs. Top Share (holdout=0.3, mean over 30 repeats)
Share of Hidden Positives Recovered
Top Share
Model settings
C=0.03, l1=0.25
C=0.03, l1=0.5
C=0.03, l1=0.75
C=0.1, l1=0.25
C=0.1, l1=0.5
C=0.1, l1=0.75
C=0.3, l1=0.25
C=0.3, l1=0.5
C=0.3, l1=0.75
C=1.0, l1=0.25
C=1.0, l1=0.5
C=1.0, l1=0.75

Fig.2