Estimating the impact of working from home on urban equilibrium: neighborhood scale effects using mobile data

Inessa Tregubova, Department of Geography, Hebrew University of Jerusalem

The COVID-19 pandemic has fundamentally reshaped urban dynamics, primarily through the accelerated adoption of remote and hybrid work. The new work schedule allows employees to work from home or from any other location outside the office at least couple of days a week and save time and money on commuting. These shifts challenge long-standing urban equilibrium frameworks, such as the Alonso-Muth-Mills model, which has traditionally guided understanding of residential choice in urban economics and regional science. Previously, proximity to the CBD served as a primary factor influencing residential desirability and housing prices. However, with the increase in remote work, this proximity becomes less relevant. At the same time other factors such as housing affordability and access to local amenities gain importance in the residential decision-making process.

This paper addresses the question of the effect of working from home (WFH) on urban equilibrium. This issue is investigated in terms of theory, method and empirics.

Theoretical model

At the outset, this paper presents a theoretical model that explains the consequences of remote work on the urban equilibrium, including the emergence of the so-called 'donut effect' by comparing urban equilibrium under normal conditions with urban equilibrium when people work from home. Equilibrium conditions exist when everyone tends to maximize utility:

$$U_i = f(s_i, q_i, t_c) \quad (1)$$

s_i– *composite good,* q_i – *size of housing, sq.foot,* t_c*–commuting time to CBD*

The study formulates the urban equilibrium using three equations:

1. A budget constraint extended with hedonic rent:

$$Y_i = S + q_i * \exp(\alpha + x_1 A_n + x_2 (A_{CBD} * (1-h)^{-k} t_c^{-k})_n + C_n (2)$$

where Y_i = income, C = commuting costs, S = is disposable income, k, α – model coefficients, A_{CBD} - amenities in CBD, A_n- local amenities, h = share of working days per week that a person works from home, h~0;

2. The equilibrium between spare time, spending and disposable income:

$$S = K, K = t_l * \left(\frac{p}{f}\right),\ t_l = T - t_c(1-h) + t_w \quad (3)$$

where K = total amenities spending, T= total active time per week, t_l - leisure time per week, t_w = work time per week, p = price of 1 man-made amenity fixed inside city, f = average time spent by individual on one service

3. The local market equilibrium for amenities:

$$z^n(1-h^n)K_W^n + g^n h K_H^n = A_n * p \ (4)$$

where $z^n =$ number of workplaces in the neighborhood n, $g^n =$ number of residents in the neighborhood n, K_W- workers spendings in the neighborhood n, $K_H -$ residents spending in the neighborhood n.

Violation of the equilibrium is assumed to be different across urban agglomeration. To account for this the paper presents three archetypal neighborhoods: the central business district, the residential neighborhood and the satellite city. Each of them has a unique combination of parameters which define how neighborhood recovers after pandemic.

The paper defines the impact of WFH over the short-term. This means that people may have already started working from home *(h>0)* but haven't made significant changes such as change of residence ($R_{WFH} = R$) or work ($Y_{WFH} = Y$). This shift brings different risks to neighborhoods.

As a result of the shift, the residents of the CBD start moving out from expensive downtown to more affordable suburbia as far as $\frac{dU}{dq} > |\frac{dU}{dt_c^{WFH}}|$. In the residential neighborhood, extra spare time (Δt_l) and accessibility of central amenities induced by the drop in commuting frequency (-Δh) increase residents' expectations from local accessibility. As long as $A_{CBD} * \Delta h * \Delta t_l \leq A_n (5)$ residents stay in the neighborhood but once this condition no longer exists they are likely to migrate to the areas with higher local accessibility. The urban equilibrium of the satellite city also depends on local accessibility. But in this case condition (5) is more likely to be satisfied as this type of neighborhood has been already disconnected from CBD and provides a sufficient level of local accessibility. Overall, the satellite city is expected to experience positive change after remote work and attract residents from the two other types of neighborhoods.

Data

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 use a unique dataset of mobile GPS signals collected between January 2019 and September 2023 by a commercial company relating to metropolitan Tel Aviv. The dataset contains geolocated signals for individuals with information about the beginning and end of the signal (which indicate stays in a given location). After removing noise, these data capture approximately 5% of the metro area population. The choice of area is dictated by its monocentric urban structure: most workplaces are concentrated in Tel Aviv, while residential neighborhoods are located in an area of 1516 sq.km. The high spatiotemporal granularity and accuracy of the data allow for precise estimation of remote work patterns.

For technical validation of the estimates, the study uses the Israeli Census 2022, a dataset of volume of office workplaces in 2019, and open dataset of monthly validations on public transportation provided by the Ministry of Transport for the period January 2019 – December 2023.

Other datasets were used in the study to collect data on urban equilibrium components. They include POIs from OSM, GIS building layer from the Survey of Israel, housing transactions from a real estate portal, Labor force Survey of Israel, General Transit Feed Specification (GTFS) files of the Ministry of Transport.

Empirical strategy

Estimation of remote work

The estimation of remote work comprises 2 steps: the identification of individuals' home and work locations and the estimation of share of individual's work time spent at home. Calculations are done at the monthly level as few individuals occur in the dataset over several months.

The identification of home and work location incorporates a deterministic approach: home locations are characterized by an individual's presence during night hours and Saturdays while work locations are defined as the most frequent locations located outside the home neighborhood where signals are recorded only during workdays.

To estimate an individual's remote work hours, we use Bayesian conditional probabilities. For each hour during the time interval 8 am - 7 pm of the days without signals from a work location, we compare an individual's conditional probability of being at home $P_i(H|h)$, a work location$P_i(O|h)$ or a third place $P_i(A|h)$. For example, the formula for the work location is:

$$P_i(O|h) = P_i(O) * P_i(h|O)/P_i(h)$$

Where $P_i(O)$ denotes an individual's share of office location hours in a working day, $P_i(h)$ denotes an individual's share of exact hours in during office days, $P_i(h|O)$ is the share of exact hour at an office location in office days. In order to avoid giving high weights for very rare hours, probabilities are smoothed using general monthly probability:

$$P_{i_{smooted}}(O|h) = P_i(O|h) * P(O|h)$$

In hours where the probability of being at work is the highest but signals come from home locations, we label an hour as remote work. Days where an individual has at least one remote work hour are called remote work days. The monthly remote work level for a neighborhood is defined as:

$$Remote\ Work\ ratio = avg(\frac{\Sigma(1_{remote\ work\ day})}{\Sigma(1_{workday})}$$

where an individual's *workday* is the official workday with at least one hour where $P_i(O|h) > P_i(H|h)$ & $P_i(O|h) > P_i(A|h)$

Spatial panel regression

Next, to establish the existence of new urban equilibrium we integrate neighborhood-level insights on remote work to examine its impact on key component of urban equilibrium which is housing prices. To do this, we build spatial panel regression:

$$p_{it} = \tau p_{it-1} + \delta \widehat{p_{it}} + \eta \widehat{p_{it-1}} + \beta_1 t_{it} + \beta_2 \widehat{t_{it}} + \beta_3 t_{it-1} + \beta_4 \widehat{t_{it-1}} + v_{it}$$

$$v_{it} = p v_{it-1} + \rho \hat{v}_{it} + \mu_i + \lambda_t l_N + \varepsilon_{it}$$

$$\mu_i = \kappa \hat{\mu}_i + \xi;$$

Where $\widehat{p_{it}}$=the spatially lagged housing prices for neighborhood i and month t equal to Wp_{it}, where W is a matrix of spatially lagged values defined based on proximity; $\widehat{p_{it-1}}$ -the same as $\widehat{p_{it}}$ but also lagged in time; p_{it-1} = time-lagged housing prices; t_{it}- weekly commuting time, $\widehat{t_{it}}$=the spatially lagged weekly commuting time for neighborhood i and month t equal to Wt_{it}; t_{it-1} =time-lagged weekly commuting time; $\widehat{t_{it-1}}$ = the same as $\widehat{t_{it}}$ but also lagged in time;. τ, δ, η are the response parameters of respectively, the dependent variable lagged in time, space and both. $\beta_1, \beta_2, \beta_3, \beta_4$ are the response parameters of the exogenous explanatory variables; .v_{it} = serially and spatially correlated error term; $\rho \hat{v}_{it}$ = spatially lagged error term, γ is the serial autocorrelation coefficient and ρ is the spatial autocorrelation coefficient. λ_t = the time-period-specific effects, l_N = vector of ones, meant to control for all time-specific, unit-invariant variables whose omission could bias the estimates in a typical time-series study, μ_i = spatial specific effect which control for all spatial-specific, time-invariant variables, including local accessibility (A_i) and ξ_t = vectors of i.i.d. disturbance terms, whose elements have zero mean and finite variance r2 and r2n, respectively.

Weekly commuting time (t_{it}) is calculated using the Origin-Destination(OD) matrix based on individuals home and work locations. For each pair of statistical areas, we calculate commuting time at specific month (based on GTFS files) and weight it by share of individuals whose home and work are in one of these areas and by (*1- Remote Work ratio*):

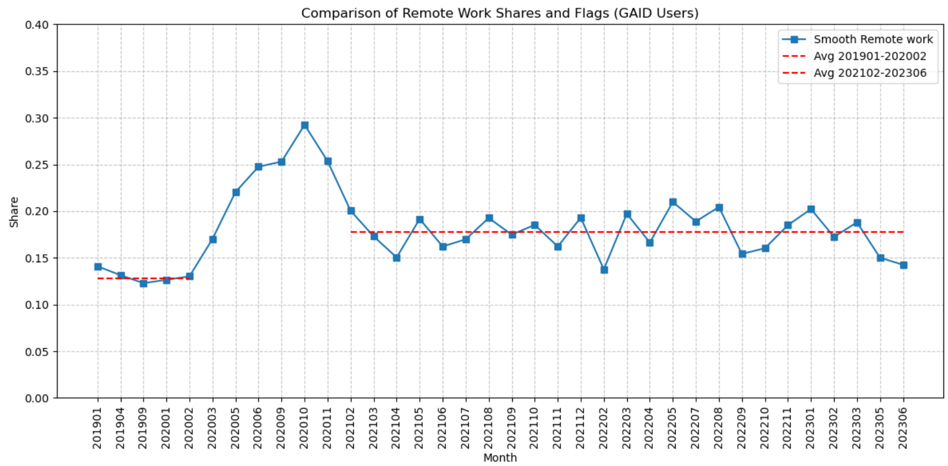
$$t_{it} = t_{c\_it} * (1 - h_{it})$$

Results

Preliminary results show high accuracy in identifying home and work locations, as well as remote work levels, using mobile data. Home location estimates aggregated by statistical area correlate 57% with the census population while city-level aggregation shows correlation of 90%. The lowest correlation appears in religious neighborhoods due to low smartphone penetration and non-use on Saturdays. Additionally, mapping home locations against nonresidential buildings confirms the quality of this identification with a monthly mismatch below 2.3%.

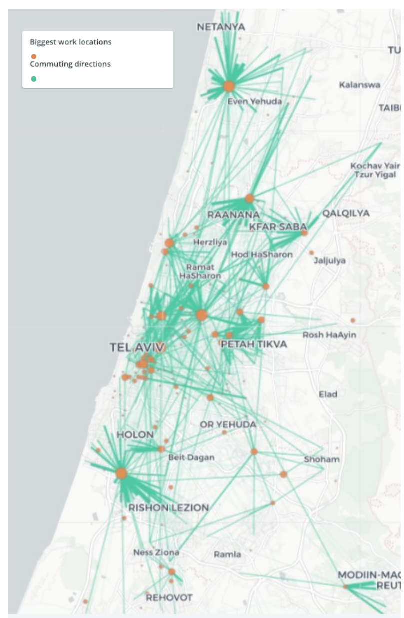
Work location identification also shows strong representativeness. Comparing mapped work locations with official work zones yields Type I error = 96% and Type II error = 69%. The lower Type II error is explained by the fact that some work zones are missing in official data.

Remote work estimates for the Tel Aviv Metropolitan area align with Central Bureau of Statistics data on average weekly work-from-home hours. A sharp rise to 30% occurred between 03/2020-11/2020, stabilizing at ~17% from 2021-2023, with fluctuations during vacation months (*Pic 1)*.



Pic 1 Remote work level dynamic

The Origin-Destination matrix, based on home and work locations, spans 1,398×1,398 statistical areas, with user flows ranging from 1 to 9. The average commuting distance is 8.1 km (*Pic.2*).



Pic 2 Map of commuting trajectories

For the second part of the study, housing transactions from 2018-2023 were collected and aggregated by statistical area. While neighborhood-level analysis is still underway an initial plot of 2019 price per sq m versus commuting distance reveals a significant inverse relationship.Further analysis on the dynamics of housing prices and commuting time, including spatial panel regression modeling, is in progress.

Contribution

This study has direct policy implications. Municipalities are currently working to undertake appropriate actions to enrich the resilience of cities and our empirical findings are likely to help prioritize these actions. To be more specific, comparing across different types of neighborhood for local public services will define the necessity of additional investments in maintaining the environment for remote workers and help prioritize them based on expected effects.

Additionally, this analysis meets decision maker's needs for assessing local policy measures at the microscale. Implementation of accurate commuting time, share of remote workers and numerous local socio-economic features into the model allows them to create the environment that reflects the particular needs of specific places with a high level of precision.

Finally, the explicit consideration of remote and hybrid workers is a big step towards re-theorizing the role of commuting and unique daily activities in city development. While the urban fall-out from Covid-19 is not yet fully understood, this research goes some way in unraveling the spatio-temporal dynamics of the home-work relationship that fashion the process.