האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM
الجامعة العبرية في القدس

Patterns of heterogeneity in the effect of public transportation services on residential rents

Gal Amedi[1]

Advisor: Michael Amior

A Dissertation Submitted to the Bogen Family Department of Economics and the Federman School of Public Policy and Government

December 2021

Jerusalem

[1] I would like to thank Michael Amior (Hebrew University) and Noam Zussman (Bank of Israel) for dedicated guidance throughout the project, Jonathan Dingel (Chicago University), Nick Tsivanidis (University of California at Berkeley), and various personnel at the Bank of Israel for helpful advice. I would also like to thank Ido Klein, Sarit Levy, and Vladimir Simon (Israeli Ministry of Transportation), Yakov Lev (Israel Railways Ltd.), Jonathan Brown (Jerusalem Transport Master Plan Team), Amir Shalev (Adalya), Yehoshua Shuki Cohen (Matat) and Bobi Lavi (Opisoft) for providing data and helpful professional background on urban planning and transportation in Israel.

Abstract

Accessibility from a given area is thought to be a key determinant of the utility from living in that area. Theoretically, this utility should be internalized by the rents market, increasing rents in areas with high accessibility. In this paper I utilize variation stemming from an exceptionally large increase in public transportation services in Israel in the years 2013-2019 to examine this effect. I use a theoretically grounded measure of accessibility and apply both causal machine learning and standard econometric methods on high-resolution nation-wide data to find that on average, public transportation had no effect on asked rents in the short-term, although there is important heterogeneity.

High residential density, Mixed-Use zoning, and a demographic composition more reflecting typical public transit users implies a larger elasticity of residential rents to public transit services. Proximity to Mass Transit Systems implies a positive effect on rents over and above the effect expected by a reduction in travel times alone. This last finding especially holds for new train stations inaugurated during the research period and for the Jerusalem Light Rail.

Introduction

Transportation costs are an important factor affecting the spatial organization of economic activity. These costs consist mainly of monetary costs and time spent on travel, and are mostly determined by distance, topography, infrastructure, and mobility technology. Importantly, costs can be dramatically altered by policy interventions, resulting in changes in the implied utility of living and working, and consequentially on residential rents, in different areas.

This paper utilizes a massive ongoing improvement in public transportation in Israel, driven by improvement of railways and bus services, to estimate the heterogeneity of the effect of accessibility by public transportation on residential rents. The improvement in services started around the turn of the millennia and accelerated in the 2010-2019 decade. Throughout the decade, train and bus activity increased by 66% and 53% accordingly.[2] Such a rapid improvement is highly unusual in developed countries, allowing a unique opportunity to examine the effect of transit using a large margin of change in a developed economy context.

I apply both causal machine learning and traditional econometric methods to find that on average, accessibility by public transportation had no short-term effect on residential rents. I estimate considerable heterogeneity in the treatment effect. Higher treatment effect is associated with high residential density, Mixed-Use zoning, demographic composition representing typical transit users, and proximity to newly inaugurated train stations or to the Jerusalem Light Rail.

This paper is part of a growing literature applying newly developed causal machine learning methods to the field of economic geography, and the first to apply causal machine learning methods to systematically explore patterns of heterogeneity in the effect of public transportation services on residential costs. This paper is also the first to use nation-wide high-resolution transportation data to examine the effect of transportation on economic phenomena in the Israeli context, and one of the few papers using a panel of high-resolution travel times on a nation-wide analysis.

[2] Activity In terms of total KM travelled. Other notable improvements are the inauguration of the first Light Rail (2011) and BRT (2013) systems in the country during the same period. See more at the empirical context section.

The theoretical framework relating transportation costs to spatial organization can be dated back to Rosen (1974, 1979) and Roback (1982). They present a no-spatial-arbitrage condition to develop a spatial equilibrium model. Recent developments include incorporation of commuting into the model.[3] In those models an increase in a geographical unit's accessibility to other units improves the attractiveness of residing in that unit by providing better employment opportunities and allowing residents and firms to utilize economies of scale.

A recent influential model is presented in Ahlfeldt et al. (2015). In this model a continuum of agents simultaneously chooses a residence-workplace pair. The agents' preferences over the different possibilities are affected by wages offered at the workplace area, residential costs in the residence area, the cost of commuting between them, and by a vector of idiosyncratic preference shocks to each workplace-residence pair.[4] This model also predicts that an increase in the accessibility from a given area will cause an increase in the price of residence in the area. I rely on a variant of this model, developed by Tsivanidis (2019), to estimate Residential Commuter Market Access, a theoretically grounded measure of accessibility, by mode of transportation.

The empirical literature faces an inherent difficulty in identification: possible endogeneity of investments in transportation. Papers that examine relations between transportation and economic phenomena generally account for possible endogeneity by institutional arguments, or by using planned or historical routes as an instrumental variable for current transportation infrastructure.[5] Another approach is restricting the sample only to regions enjoying the infrastructure inconsequentially.[6]

In the specific literature on public transit-residential costs relationship, standard estimation constitutes of either a difference-in-differences design or cross-sectional hedonic regressions for the effect of a single transportation project to the value of nearby properties. Identification is usually claimed using institutional knowledge, or

[3] Albouy & Lue (2015), Ahlfeldt & Feddersen (2017), Monte et al (2018), Amior & Manning (2019), Severen (2021) are recent examples.
[4] A developing strand of the literature explicitly examines the importance of idiosyncrasy in locational preferences. See for example Diamond (2016), Büchel et al (2019) and Dingel & Tintelnot (2020).
[5] See review at Redding & Turner (2015), prominent examples are: Baum-Snow (2007, 2010), Duranton & Turner (2011,2012), Duranton et al (2014), Baum-Snow et al (2017) Severen (2021). Brooks & Lutz (2016) Argue that historical railways should be used for sample selection and not as instrumental variables due to path-dependence.
[6] E.g., Chandra & Thompson (2000), Mayer & Trevien (2017), Banerjee et al (2020). A less common approach is examination of obviously exogenous shocks to transportation opportunities; A key example is the division and re-unification of Berlin, see Redding & Sturm (2008), Ahlfeldt et al (2015).

without accounting for endogeneity. In this paper I make an institutional argument for exogeneity in the timing of allocation of public transportation in Israel. I estimate models both directly and using an Instrumental Variable approach.

Empirically, the literature generally finds a small positive residential costs-accessibility relationship.[7] Proximity to BRT, light rail or train stations implies a 12%, 4% or 6% increase in property values accordingly,[8] though there is considerable variation between studies, including many studies who find a zero, or even negative effect. The effect on residential rents is generally smaller than the effect on property values, though research on this relationship is rather scarce.

Papers that examine heterogeneity in the effect usually focus on a single dimension. Common findings include an effect rising with proximity to stations, stronger for smaller or more expensive apartments, and for suburbs compared to inner cities. The literature also finds a stronger effect for rail and BRT systems compared to regular bus services.[9] The literature lacks analyses that systematically tests for different sources of variation, and their relative importance in a combined framework.

In the Israeli context, some papers examine the effect of transportation on residential and employment location choices: Leck et al (2008) finds that rail transit diminished periphery-core wage disparities in southern and central Israel. Israel & Blankshtain-Cohen (2010) find a suburbanization and counter urbanization effects of rail services in the Tel Aviv metropolitan area. Frisch & Tzur (2010) finds that new road and rail infrastructure increased the share of long-distance commuters in their catchment areas. Bleikh (2018) explores long-term trends in commuting.

Other prominent papers include Ida & Talit (2018), who describes and examines the effect of an ongoing reform in bus operation in Israel. Soffer & Suhoy (2019) use survey data to construct relative accessibility indices by transportation mode, and examine determinants of modal transportation choice. They find an increase in recent

[7] See Wardrip (2011), Mohammd et al (2013), Ingvardson & Nielsen (2018) or Zhang & Yen (2020) for reviews. Rodriguez & Traga (2004), Dorantes et al (2011), De Bruyne & Van Hove, (2013), Brennan et al (2014), Cao et al (2019), Li et al (2019), Wittowski et al (2020) and Yang et al (2020) for specific examples. Portnov et al (2009) applies a hedonic model to examine the disutility from proximity to rails in the Israeli context.

[8] Median values from papers included in tables 2-4 in Ingvardson & Nielsen (2018).

[9] This might be a result of increased visibility of these projects, transportation mode preferences, or the more stable nature of rail networks.

years in rail use among all income groups. Several papers[10] examine employment and education effects of penetration or massive improvement of bus services to Arab localities following Israeli Government Decisions: No. 1539 (2009), and No. 922 (2015), aimed specifically to manifest economic development in those localities. General findings include a small positive effect of on employment.

The paper proceeds as follows: section 1 provides empirical context on housing and transportation in Israel, section 2 describes the data, section 3 describes the Commuter Market Access measure, section 4 the methodology, section 5 reports and discusses the results, and section 6 concludes.

1. Empirical setting

1.1 Housing and rents market

Economic activity and population and in Israel are mainly concentrated around three metropolitan areas of descending economic importance: Tel Aviv, Jerusalem, and Haifa. Rents and housing prices, as theory suggests, are higher around the metropolitan areas, especially Tel Aviv.

Table 1

Asked Rents for four room apartments by district

District	Jerusalem	North	Haifa	Center	Tel Aviv	South
2013	4.66	2.82	3.34	4.32	5.58	3.05
2019	5.24	3.03	3.66	4.86	6.34	3.64
Percent Change	12.3%	7.4%	9.3%	12.6%	13.6%	19.3%

Source: RENTS dataset. Raw average prices in thousand NIS

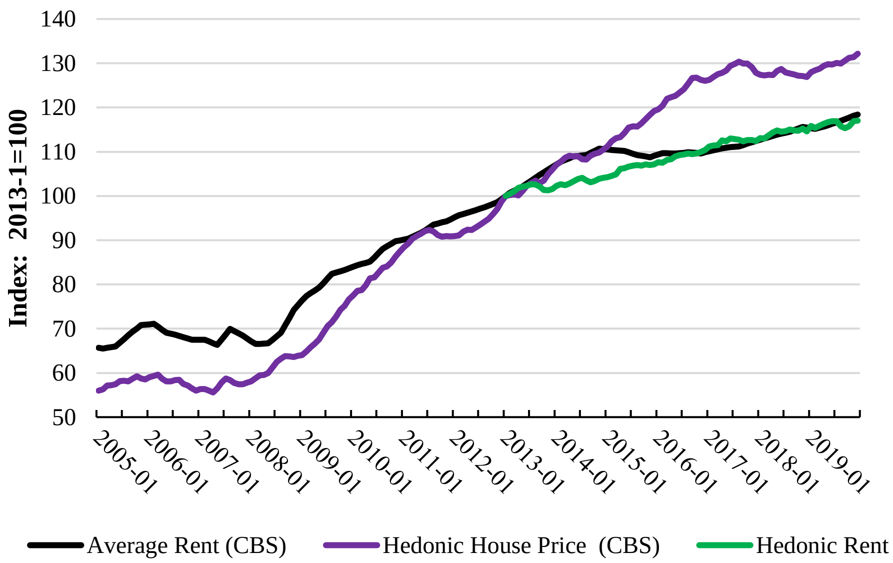
Residential costs significantly increased mainly before, but also throughout the research period (2013-2019).[11] The appreciation began in 2008, with house prices rising 111.3% in 2008-2019 (27% during the research period). Rents, the dependent

[10]Greenwald et al (2018), Abu qarn & Lichtman-Sadot (2019), Barak (2020).

[11] Several papers examined whether this hike represents the development of a price bubble and concluded that it is not the case. Yakhin & Gamrasni (2021) estimate that the price level in 2019 is only 5.5% higher than the estimated long-run equilibrium price. Also see Dovman et al (2012), Caspi (2016), Arestis & Gonzalez-Martinez (2017) for analysis of the major hike in the early period.

variable in this paper, also increased, but at a much lower rate; 68.3% in 2008-2019 (only 12.9% during the research period). The appreciation is visualized in figure 1.[12]

Figure 1
Housing price indices, 2005-2019



Source: Israeli CBS, hedonic rents estimated with data in the paper.

1.2 Transportation in Israel

Until the late 1990's, bus services were operated almost exclusively by two cooperatives: Egged and Dan who provided 96% of all bus services in Israel in 1997. The large market share along with weak regulation led to complete dependence on the cooperatives which in turn led to a gradual decline in the level of services. Following government decision 1301 (1997), the right to operate public transportation lines was gradually tendered to new firms, inducing improvement in services and efficiency.[13]

After the Six days war (1967) investment in rails was neglected in favor of an auto-oriented policy. Toward the turn of the millennia, roughly at the same time the bus reform was initiated, a policy shift led to large successful investments in rail infrastructure specifically intended to serve passengers (and not freight, see figure 2).

[12] The hedonic rental price index produced by the Israeli CBS have been shown to be problematic due to exclusion of new tenants from the estimation (See Raz-Dror, 2019). Therefore, the reported CBS figure is the simple average price index. I also report a hedonic regression estimated with my data, including all physical and spatial variables described in the data section, and area fixed effects.

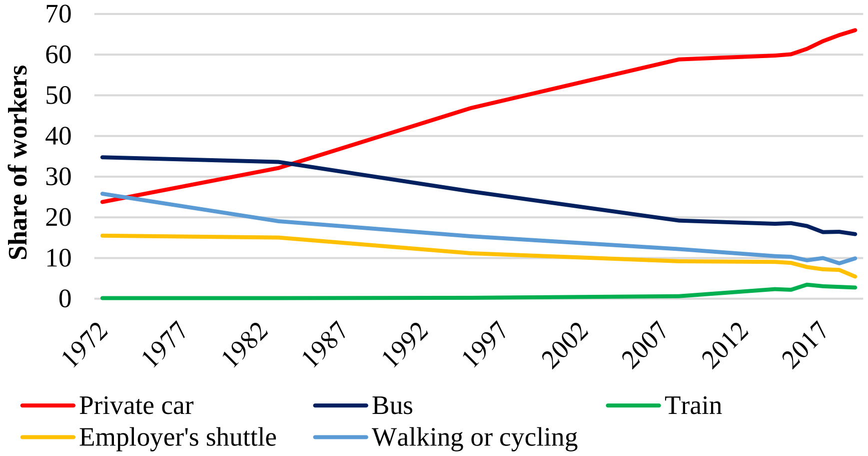
[13] As a side note, the process of tendering all services takes much longer than originally expected and is still undergoing. Tenders that took place are considered a success with a large increase in the level of service and decrease in costs. Thorough review can be found at Ida & Talit (2018).

Figure 2
Annual totals of rail passengers and freight in Israel, 1950-2019

Source: Israeli Central Beaurau of Statistics annual reports

The rise in the number of passengers by train towards the end of the millennia is impressive, but train passengers still constitute only a small share of the number of commuters. An increase in the standard of living, alongside the aforementioned auto-oriented policy, resulted in a consistent and significant upward trend in the motorization rate and the use of private cars for commuting (figure 3).[14]

Figure 3
Travel to work mode in Israel, 1972-2019

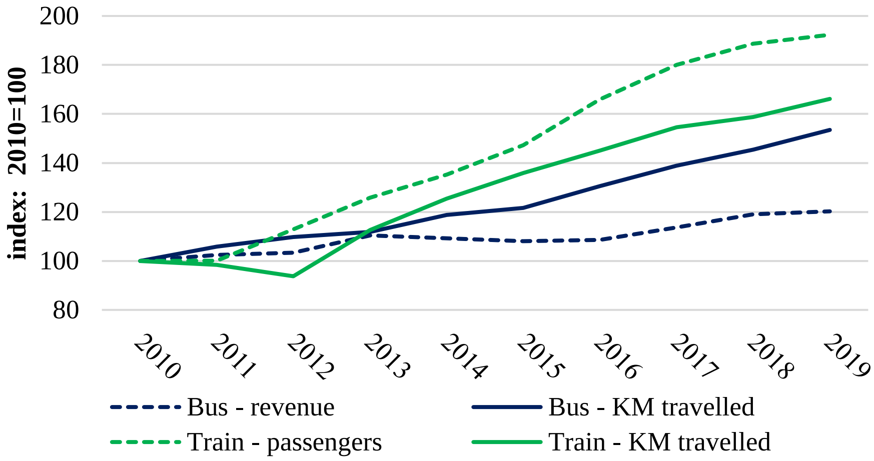


Note: The 1972 census had no seperation between public buses and employer's shuttles. I divided the unified category based on the stable ratio between them in later years.The 1983 survey had no seperate category for train passengers, I've assumed linear progress between the 1972 and 1995 censuses.
Source: Israeli Central Beaurau of Statistics censuses and social surveys

[14] These trends in the past two decades are discussed at Friedman (2019).

The last decade demonstrated considerable growth in the supply of public transportation, and accordingly in the number of passengers (figure 4). The improvement in the bus network was more pronounced in Haifa and its surroundings, in Judea & Samaria, and in the Ashdod greater area (figure 5). Improvement in rail services focused on similar regions. Out of 68 now-active stations in Israel, 15 were inaugurated during the research period: stations in the new "Rakevet HaEmek" rail connecting Haifa to the Jezreel valley and Bet-Shean, the rail to Karmiel, the new southern rail, a new station in Jerusalem,[15] and several suburban stations in central Israel. It seems that improvements in train services were more effective, with ridership by train hiking considerably higher than bus ridership in the last decade.[16]

Figure 4
Transportation statistics by mode, 2010-2019



Source: Israeli Central Beaurau of Statistics annual reports

1.3 Process of public transportation allocation

To identify the effect of public transit on rents I rely on the exogeneity of the timing of public transit allocation. This section argues for exogeneity in the timing of allocation of both bus and train services.[17]

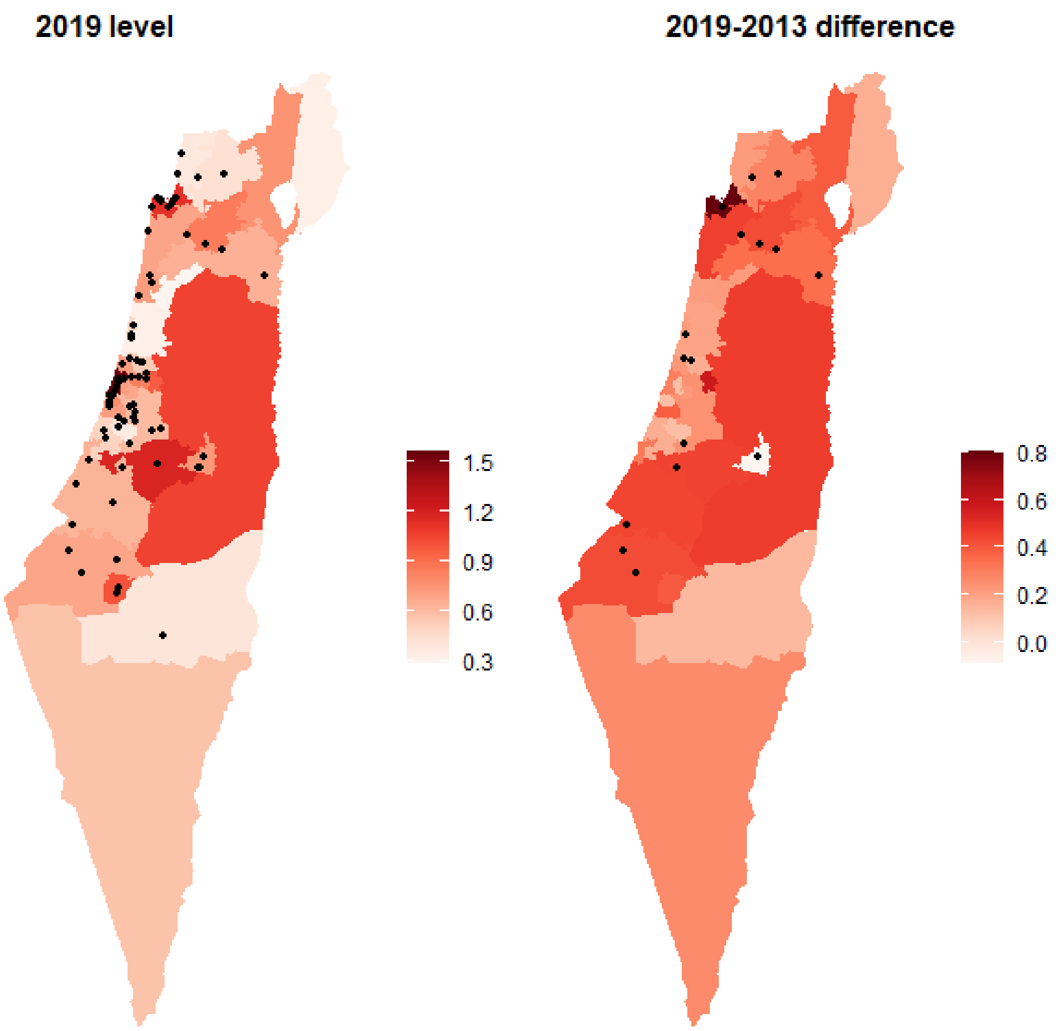
[15] Jerusalem was connected to rail services since 1892, but the service wasn't effective since the old rail and stations' location didn't allow quick travel to major economic centers. Many of the new rails follow the general path of historical rails built by former sovereigns of the region as an extension of the Hejaz railway and for British military purposes (Falkov, 1982).

[16] Note that there is no available data for total number of passengers by bus, I proxy for this using the total revenue of bus operators from regular bus lines.

[17] The Jerusalem Light Rail will not be described here. It is operated by a private firm under the supervision of the Jerusalem Transport Master Plan Team. There was no change to its rails since its inauguration in 2011, though there is as an improvement in travel times and frequency.

Figure 5

Bus activity per capita and active train stations, 2019 level and change during the period



Note: Activity defined as the number of times a bus stops at any station in the region.

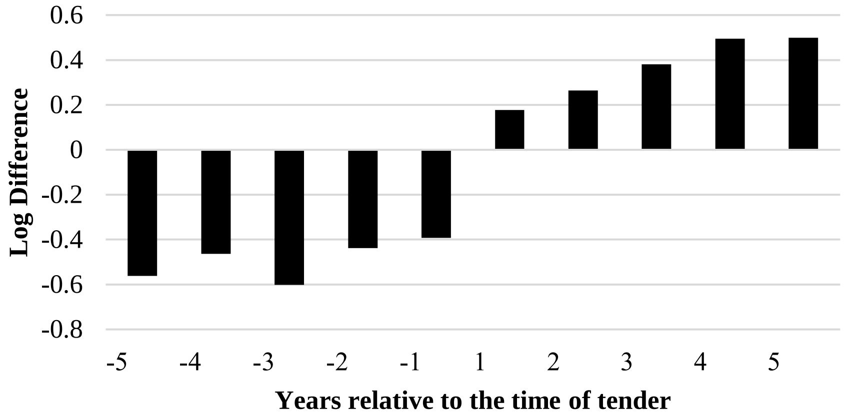
<u>Bus[18]</u>

Planning and allocation of the bus network in Israel, including determination of routes, frequency, and schedule is under the responsibility of the National Public Transport Authority (NPTA).[19] The bus network is divided into operation clusters of

[18] This section relies on Ida & Talit (2018), and on conversations with officials at the Ministry of Transportation and Adalya, a consulting firm providing services to the NPTA.
[19] A rather new authority under the responsibility of the Ministry of Transportation. established in 2012 as a result of government decision No. 3988 (2011). This section is also relevant to the Metronit, Israel's only BRT system operating in Haifa.

different size.[20] The operation is done by private firms, competing off-the-road in public tenders for exclusive rights to operate all bus lines in a cluster for 12 years.[21] The long duration of the operation agreements implies that the starting date of a new operation agreement is predetermined over a decade before actually taking place.

Figure 6
Log difference of average bus activity compared to time of tender



Note: Activity is defined as the number of times a bus stops at the station during a regular weekday.The presented difference is the average of log differences in each station's activity relative to the time of tender

A new operation agreement typically implies an immediate improvement, followed by an upward trend in services in the cluster. Figure 6 displays the average of log differences in a station's activity by time from the tender relevant to the station taking place.[22] In the end of 2019, the bus network was divided to 71 clusters, 18 of which, covering 44% of all weekday activity in the network, were tendered during the research period (2013-2019).

Train

The development of new railways in Israel is co-planned by Israel Railways Ltd., and the NPTA. Operation and schedule decisions are under the responsibility of Israel Railways, with NPTA supervision. Like similar transportation projects worldwide, the time between the beginning of the planning process of a new station to planned

[20] A cluster of lines usually includes a share of services in a metropolis, all service in a large locality, a group of close localities or a specified non-urban region, or all interurban services between two areas.
[21] Formally the winner will operate the cluster for 6 years. At the end of the first 6 years, if there were no severe violations, the NPTA can choose to extend the operation period twice for 3 years at a time. The NPTA never chose not to extend an operation period. Towards the end of the research period the NPTA changed the operation period in new tenders to 10 years, with no extensions.
[22] Time of tender for a given station defined as the first transportation period in which over 50% of the activity is tendered. Activity is defined as the number of bus stops-at-station.

inauguration is long.[23] On top of the long planning time, there is large uncertainty on the projects' schedule. The Bank of Israel (2010) puts a lower bound on the duration of average schedule overrun for rail projects in Israel at 72%.[24]

2. Data

The data used in this paper was received from different organizations, as summarized in table 2. In this section I will briefly describe the assembled dataset by main subjects: transportation, rental ads, the Origin-Destination matrix, and additional data.

Table 2

Summary of data sets

Dataset	Source	Range	Relevant Variables
TRAIN_RIDES	Israel Railways Ltd.	2013-2019	Actual and planned time for each stop-at-station in each train ride
LIGHT_RAIL	Jerusalem Transport Master Plan Team	2013-2019	Actual time of the start and end of each light rail ride
BUS_RIDES	Israeli Ministry of Transportation	2016-2019	Actual time of the start and end of each bus ride
BUS_SCHEDULE	Israeli Ministry of Transportation	2013-2019	Planned time of the start and end of each bus ride
BUS_ROUTES	Israeli Ministry of Transportation	2013-2019	Complete description of each line's route: stations location, and road distance and planned travel time between stations. Received twice a year
ROADS_NETWORK	Survey of Israel (Mapi), part of the BENTAL dataset	2013-2019	GIS of all roads in Israel including number of lanes in each direction, received quarterly.
RENTS	Private firm	2013-2019	Price, size, number of rooms, floor, number of floors in the building, number of toilet rooms. Dummies for renovation status and the existence of: air conditioner, lift in the building, parking, balcony, security room, new kitchen, barred windows.
ADDRESSES	Survey of Israel (Mapi)		Exact coordinates of addresses
OD_MAT	Israeli Ministry of Transportation	2018-2019	Period average by time of day of people making the journey (1210 polygons)
CBS_DATA	Israeli CBS	2013-2019	Annual statistical-area level data on Socio-Economic status and demographic variables

2.1 Transportation

Bus[25]

Bus data is merged from three separate datasets provided by the Israeli Ministry of Transportation. The BUS_RIDES dataset records real travel times for the universe of

[23] There are already several planned stations whose planned inauguration date is in 2040.
[24] More information on the uncertainty in the planning schedule can be found at Bank of Israel (2015)
[25] The 'Metronit' - a BRT system inaugurated in Haifa in august 2013 is also included in these datasets.

all regular bus rides for the years 2016-2019. Real travel time data exists only for the complete ride. A separate dataset, BUS_ROUTES, contains data on planned travel times for each transportation period,[26] including the route of each line, road distance, and planned travel times between all stations in the line. A third dataset, BUS_SCHEDULE, includes detailed schedule for each bus ride in 2013-2019. I use imputed real travel times, and planned schedule, to compute effective travel times between each two bus stations – see detailed explanation at appendix A.

<u>Trains & Light Rail</u>

The TRAIN_RIDES dataset contains raw data from Israel Railways Ltd. and includes the universe of all train rides between 2013 and 2019. The dataset contains, among other fields, planned and actual arrival and departure times for each station in each train ride during these years. This allows for direct computation of travel time between any two train stations at each time of day during the sample.

The LIGHT_RAIL dataset, composed by the Jerusalem Transport Master Plan Team, contains raw data of actual departure and arrival time of each ride in the Jerusalem Light Rail. I divide the total travel time to different segments using each segment's proportion in the planned travel time to obtain travel times between stations.

<u>Roads</u>

The ROADS_NETWORK dataset is part of the standard BENTAL dataset produced by Survey of Israel (mapi). ROADS_NETWORK includes quarterly GIS data of the entire Israeli road network. The dataset includes, among other fields, information about the number of lanes in each direction.

2.2 Rent ads

The RENTS dataset includes information on asked rents and apartment characteristics. RENTS is collected by a private firm by scarping rental ads from all popular sites in Israel. RENTS is used by the Israeli CBS, the Bank of Israel, and other public organizations and is available to me starting 2013. RENTS includes information on asked rent, address, dates when the ad was uploaded or updated, and

[26] Transportation periods are defined as the period between January 1st to the Jewish holiday Pesach, and the period starting from the end of Pesach until July 1st. I impute transportation data for the rest of the year (July-December) as the average value of the two adjacent periods.

various physical characteristics.[27] RENTS may contain multiples spans of the same ad. I use the last appearance of an ad to diminish noise from errors and idiosyncratic beliefs on prices.

I geo-reference apartments in RENTS using the ADDRESSES dataset, which contains coordinates for most addresses in Israel and keep only successfully geo-referenced ads.[28] I further cleanse RENTS by filtering out ads that have no access to public transportation,[29] or ads with missing, clearly wrong or unusual characteristics.[30] I preform finer filtering by comparing the price and size reported to the corresponding median value of the 100 geographically closest similar[31] apartments. I only keep ads where the ratio between ad and the median is within the 0.5-1.5 interval. This procedure results in a final dataset of 760,568 ads in 147,283 unique addresses.

2.3 Origin-Destination matrix

The OD_MAT dataset, received from the Israeli Ministry of Transportation, contains detailed data on journeys. OD_MAT is the product of a large-scale project continuously monitoring the location of roughly a half of all mobile phones in Israel. The dataset contains the 2018-2019 average flow in half-hour intervals between 1250 transportation polygons. OD_MAT is based on data from 3.77 million unique cellphones, and represents roughly 2.75 billion human days. After appropriate weighting, the data describes a total of 15.76 million rides in an average weekday – roughly 2.1 daily rides per person. [32]

Since OD_MAT is collected using cellular location data, the size of a transportation polygon is influenced by the density of cellular antennas in the area. Size ranging

[27] I use all physical attributes that are non-missing in more than 90% of the ads in the dataset: rent, size, number of rooms, floor, number of floors in the building, number of toilet rooms, and dummies for renovation status and the existence of: air conditioner, lift in the building, parking, balcony, security room, new kitchen, and barred windows.

[28] Only 2.7% of the observations were dropped due to unsuccessful geo-referencing. A small number of addresses missing from the ADDRESSES dataset were geo-referenced using either Google Maps or Open Street Map API's.

[29] I examine the effect in terms of elasticity, keeping ads without any access to public transportation would imply that a small improvement to service would show up as a huge change in log points.

[30] Apartments with less than 1, or more than 6.5 rooms, or apartments whose rent per square meter is not within the 10-200 NIS interval.

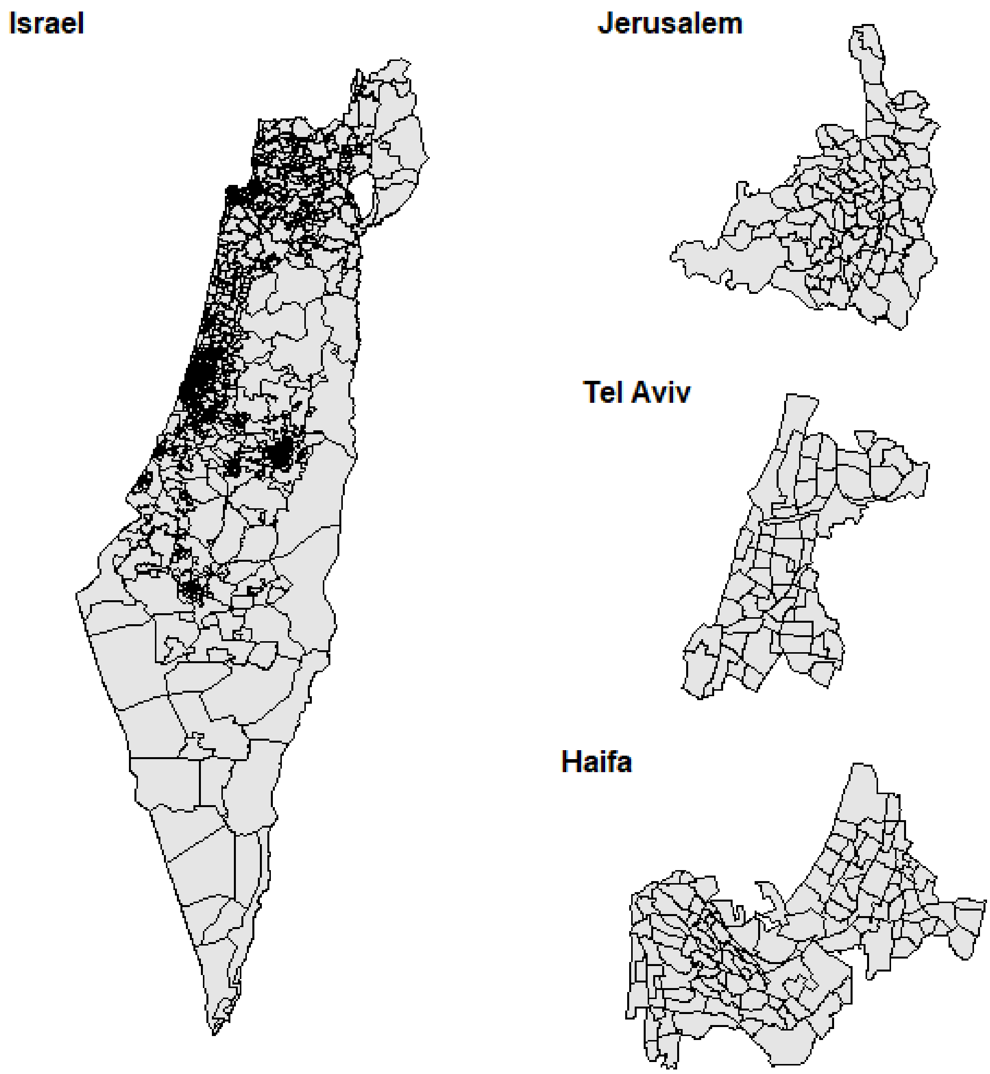
[31] Apartments that have roughly the same number of rooms (difference in the number of rooms isn't more than half) and were offered to rent roughly at the same time (during the 90 days period centered around the ad's publication date).

[32] A presentation of the project appears in Matat, (2019). The flows to, and from 40 transportation polygons were blanked due to confidentiality issues. These flows constitute less than 3% of all journeys in the dataset.

between 0.12 to 1,079 Square KM, with a median value of almost 2. Polygons in populated areas are much smaller than polygons in rural areas, as presented in figure 7. The cities Jerusalem, Tel-Aviv, and Haifa are divided to 83, 63 and 69 polygons accordingly. The median polygon contained 6,244 residents in 2018.

Figure 7

Transportation polygons in Israel



There is no direct way to reveal the purpose of rides, nor the complete round-trip journeys of individuals from the data. Therefore, one must choose which times of day are more likely to represent a residence-workplace commute. A narrow interval would result in a less accurate measure of the commuting flow and might bias the counts by selecting specific types of commuters. A wide interval might mix morning commutes with rides for different purposes: rides for leisure purposes, rides during the workday, or the commute back home for people departing early.

I define the residence-workplace commuting flow between every pair as the sum of all journeys between them starting between 6:30-9:30. According to the Israeli Population Census (2008), which is the most relevant dataset containing direct (self-reported) information about the distribution of commutes throughout the day, this interval covers two thirds of all regular commutes to the workplace. I use OD_MAT to extract another type of flows: journeys originating between 19:30-21:00. I view these journeys as largely consisting of rides to leisure activity such as restaurants, concerts etc. Thus, I view the sum of flows to a specific polygon originating in this time of day as proxying for amenities in the polygon.

2.4 Additional data

Most of the additional data is statistical area level data extracted from the Israeli Central Bureau of Statistics website, named CBS_DATA in table 2. I extract the following annual data for each statistical area: Population count, Socio-Economic status[33] share of non-Jews, ultra-orthodox, males and each of the following age groups in the area: 0-19,20-39,40-59, over 60.

Other data includes: The date of all bus tenders in Israel since the beginning of the reform, which is used to construct the instrumental variable later described, the dates of each transportation period used by the Israeli Ministry of Transportation for planning, and the CBS Labor Force and Social surveys used for stylized facts.

3. Commuter Market Access

3.1 Framework

I adopt the Commuter Market Access framework (CMA) developed in Tsivanidis (2019), to define accessibility by mode for each address in the dataset in each point in time. CMA for a spatial unit is given by Residential Commuter Market Access (RCMA), representing accessibility of a resident in the unit to possible employers, and Firm Commuter Market Access (FCMA) which represents the accessibility of a firm located in the unit to possible employees. Tsivanidis shows that in a wide class of

[33] This data exists for the years 2013, 2015 and 2017. I assume linear progress in the missing periods and impute 2017 level in 2018-2019.

quantitative urban models, CMA is a sufficient statistic summarizing the impact of commuting costs on economic equilibrium outcomes.

The Commuter Market Access is defined by the following set of equations:

$$(1)\ \ RCMA_o = \sum_d \frac{LF_d}{FCMA_d} \kappa_{od}$$

$$(2)\ \ FCMA_o = \sum_d \frac{LR_d}{RCMA_d} \kappa_{do}$$

LF_d is the number of workers in polygon d, LR_d is the number of residents in polygon d, κ_{od} is a measure of connectivity between the polygons discussed below. For N polygons, this system represents a set of 2N equations with 2N unknowns and has a unique solution. The contribution to Residential Commuter Market Access of polygon o from the connection to polygon d is higher if polygon d isn't easily accessible to workers, the number of workers in d is high, and the travel time to d is low.

3.2 Definition of Connectivity

I follow Dingel & Tintelnot (2020) in the parametrization of travel times, as defined in Appendix B, to commuting costs. For any transportation mode, I parametrize commuting costs as:

$$(3)\ \ \delta_{od}^{m} \equiv \frac{H}{H - t_{od}^{m}}$$

m represents the mode of transportation and can take one of three values: PT for public transportation, car for private car, or all for a weighted average of those (described below). t_{od}^{m} is the total roundtrip travel time between polygons o and polygon d by transportation mode m.[34] H represents the total daily hours a worker dedicates to working and commuting. Hence, commuting cost is the inverse of the share of time a worker spends on travelling in his workday. The average number of daily working hours for a worker in a full-time position in Israel is 8.7, and his

[34] The sum of travel times from o to d during the morning rush hour, and from d to o during the afternoon rush hour. Since I use national data, there are some observations for which $t_{od}^{m} > H$. In those cases, I winsorize travel times to 539 minutes (1 minute less than 9 hours).

average one direction commute time is 30.7 minutes, leading to an empirical $H = 9.7$.[35] For consistency with prior research, I assume $H = 9$.[36]

Connectivity between o and d in mode m is defined as:

$$(3)\ \ \kappa_{od}^{m} \equiv [\delta_{od}^{m}]^{\epsilon}$$

Where ϵ is the elasticity of commuting with respect to commuting costs. Note that by construction κ_{ij}^{m} is bounded between 0 and 1. A travel time of 0 minutes implies a connectivity measure of 1.

3.3 Disaggregation by mode of travel

I estimate mode-unified Residential and Firm Commuter Market Access measures ($RCMA^{all}, FCMA^{all}$) for all polygons according to equations (1, 2), and using κ_{od}^{all} as the measure of connectivity. κ_{od}^{all} is defined in terms of a mode-unified travel time measure, t_{od}^{all}: the average of travel times by public transportation and private cars, weighted by the share of workers commuting with or without private cars:[37]

$$(4)\ \ t_{od}^{all} \equiv (car.share * t_{od}^{car} + (1 - car.share) * t_{od}^{PT})$$

Lastly, I use the mode-unified Firm Commuter Market Access measure, $FCMA_{d}^{all}$, to assign Residential Commuter Market Access by transportation mode m for each ad I that appear in the dataset at transportation period t, using the following equation:

$$(5)\ \ RCMA_{i}^{m} = \sum_{d} \frac{LF_{d}}{FCMA_{d}^{all}} \kappa_{idt}^{m}$$

Which is equivalent to equation (1), With κ_{idt}^{m} representing the connectivity from an apartment i to area d, at transportation period t, and by transportation mode m.[38] Note that $FCMA_{d}^{all}$ and LF_{d} remains constant across time and transportation modes. Hence, all variation across time and transportation modes in $RCMA_{i}^{m}$ for a given address is the result of changes in travel times and does not reflect changes in the attractiveness of the possible destinations.

[35] Average values from the Israeli Labor Force survey for 2018-2019. The commute time is relatively long compared to a rough OECD average of 20 minutes (OECD, 2011).
[36] Estimates where I assumed H=10 or H=8 yielded practically identical connectivity measures.
[37] I calculate this share using the Israeli Social Survey. I use the average 2014-2019 value: 67.7%
[38] The travel time considered is the minimal total travel time from the apartment to any of the public transportation stations in the destination polygon.

3.4 Estimation

I present the elasticities using travel times by different modes of transportation, ϵ^m, in table 3, and the implied connectivity measures $\kappa_{od}^m(t_{od}^m) = [\delta_{od}^m]^{\epsilon^m}$ in appendix figure A.1. The gravity equations are estimated using a Pseudo Poisson Maximum Likelihood estimator.[39] This estimator has been shown to outperform the OLS estimator when estimating gravity equations, especially in granular settings.[40]

Table 3

Commuting elasticity estimates

	Mode-Unified	PT	Car
Elasticity	-10.96*** (0.228)	-9.182*** (0.445)	-10.17*** (0.247)
Pseudo R^2	0.728	0.639	0.701
Location pairs	1,464,100		
Commuters	2,592,630		

Note: Standard errors shown in parentheses.

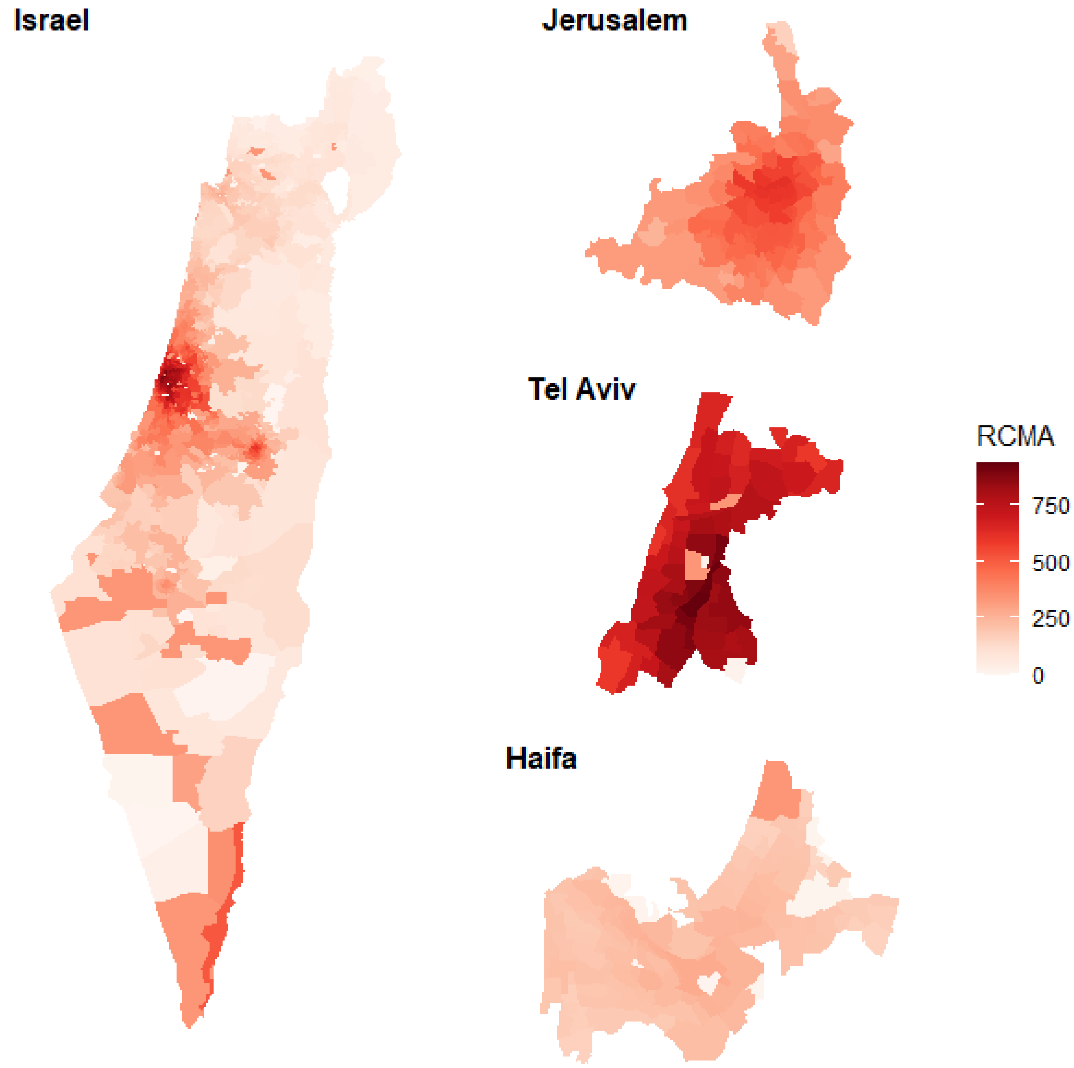
Since OD_MAT represents average 2018-2019 values, I use the average 2018-2019 travel times for estimation. The estimated elasticity is of similar magnitude to elasticities reported in papers using the same parametrization.[41] The gravity model estimated with mode-unified travel times presents the best goodness of fit, lending support to its construction.[42]

Figure 8 presents estimated $FCMA^{all}, RCMA^{all}$. As expected, both in the national and urban levels all measures of accessibility escalate near important economic centers. Also note a surprisingly high Firm Commuter Market Access in eastern Haifa, which might drive some of the results later presented concerning the proximity to the Metronit (Israel's only BRT system operating in Haifa).[43]

[39] Specifically, I use the PPMLHDFE command available in Stata (Correia et al, 2019).
[40] See Silva & Tenreyro (2006) for discussion of the shortcomings of estimating gravity equations with OLS, and Dingel & Tintelnot (2020) for a discussion on granular settings.
[41] Dingel & Tintelnot (2020) reports elasticities ranging between -7.99 to -19.81
[42] The elasticities by distinct modes of transportation are not further used in the paper but presented to demonstrate that the model specification uncertainty stemming from the fact that the data does not offer disaggregation by transportation mode is probably small.
[43] Also note a surprisingly high $RCMA^{all}$ near Eilat (an important tourism center in the southern end of Israel). This finding might be the result of leisure rides to the city of Eilat originating during morning rush hours, which are indistinguishable from commutes in my dataset. This should not affect results since there are almost no ads in areas that are relevant for a commute to Eilat in the RENTS dataset.

Figure 8.a

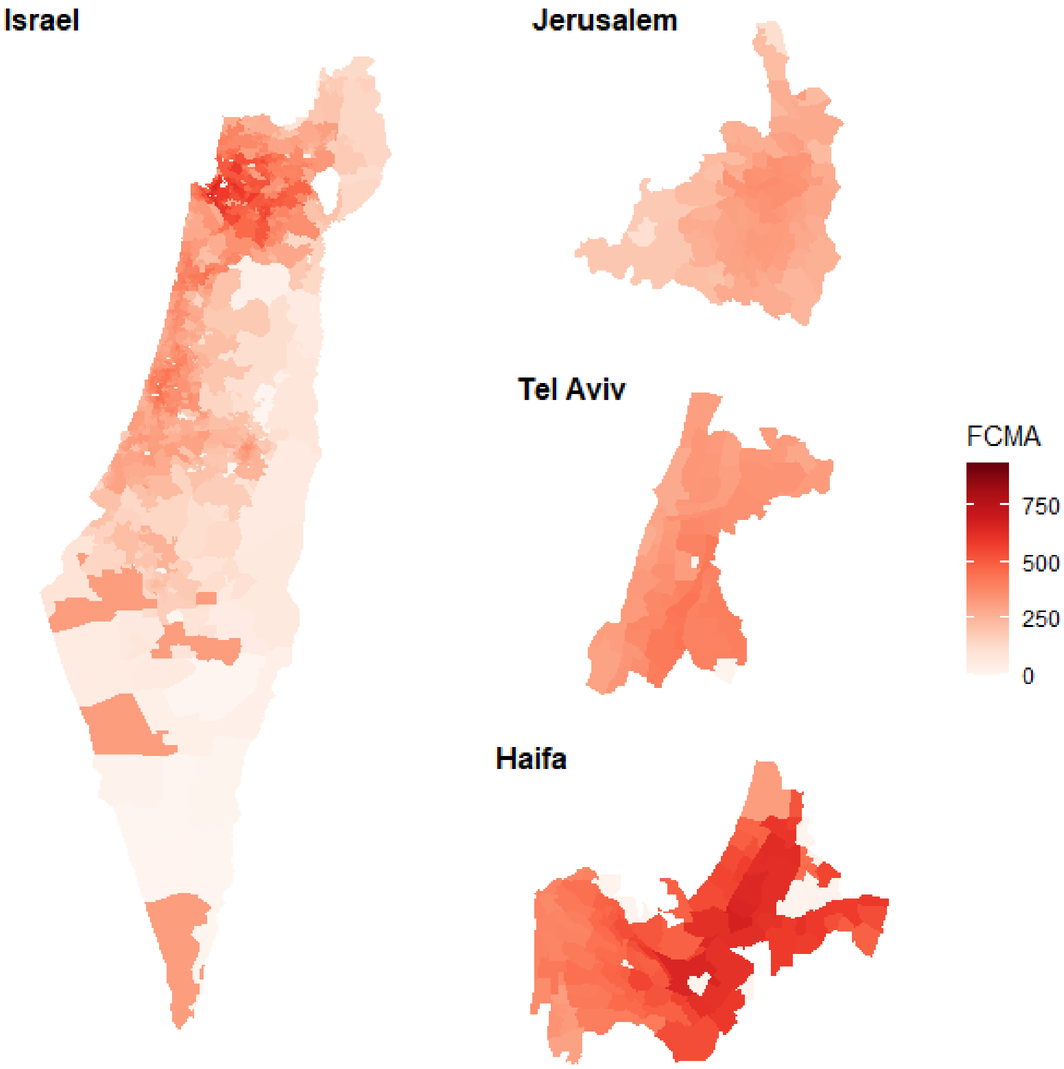
Estimated Residential Commuter Market Access



Note: No data was received for flows from and to 40 polygons due to confidentiality issues. These areas are plotted with the average value of Residential Commuter Market Access in their region.

Figure 8.b

Estimated Firm Commuter Market Access



Note: No data was received for flows from and to 40 polygons due to confidentiality issues. These areas are plotted with the average value of Firm Commuter Market Access in their region.

4. Methodology

In all models presented here the dependent variable is the log of asked rent, and the treatment variable is the log of $RCMA_i^{PT}$. Hence, the estimated effect can be interpreted as the elasticity of rent with respect to accessibility by public transportation. I focus on the effect on rents instead of sales price to mitigate threats to identification arising from anticipation. An observation in the dataset is an ad i, located at address j, in region r, whose last appearance in the dataset is at year t.

4.1. Baseline Model

The baseline specification relies on variance of the level of service and prices within addresses over time, and conditional on the district-specific trend, to identify a causal effect. Specifically, I estimate:

$$(6)\ \log(rent)_{ijrt} = \beta_0 + \beta_1 * \log(RCMA^{PT})_{jt} + \mu_j + \psi_{rt} + \beta X_{ijrt} + \upsilon_{ijrt}$$

Where μ_j is address fixed effects, ψ_{rt} represents a set of district-year dummies, and X_{ijrt} a set of apartment specific features including $\log\left(RCMA_{jt}^{car}\right)$, population density, Socio-Economic status, the floor of the apartment, number of floors in the building, number of rooms and toilet rooms, the size of the apartment in square meters, the ratio of its size to the size of similar nearby apartments, and dummies for: a new kitchen, air conditioner, parking, barred windows, balcony, security room and renovation status. υ_{ijrt} is an ad-specific error term.

The choice of controls and functional forms to partial-out apartment specific idiosyncrasies is not trivial. Misspecification of functional forms might pose a threat in my context since the price of an apartment could be a non-trivial function of its features, and there might be correlations between changes in public transit services and certain apartment characteristics.

To address this issue, I estimate the model using two approaches: (1) In the Baseline specification I rely on a best-linear-approximation argument[44] and control for apartment specific characteristics linearly. In practice, I estimate the model with all apartment-specific, and time-variant spatial features controlled for linearly. (2) In A second approach, denoted LASSO, I augment the dataset with all possible two-way interactions between apartment-specific, and spatial time-variant features in the data (the vector denoted X), and estimate the model with automatic selection of controls using the Double-Selection LASSO method developed in Belloni et al (2014).[45]

[44] See Angrist & Pischke (2008).

[45] Briefly, the procedure has three stages. Denote, Y the dependent variable (log asked rents), and W the treatment variable (log $RCMA^{PT}$). (1) Run a LASSO model of Y on a vector X of possible controls, denote the group of selected variables $G(Y,X)$. (2) Run a LASSO model of W on a vector X of possible controls, denote the group of selected variables $G(W,X)$. (3) Run a linear regression of Y on W and the union: $G(Y,X) \cup G(W,X)$ as controls. District-year and address fixed effects are included in all regressions.

Inclusion of the address fixed effects partials out all time-invariant variation in the spatial features of the apartment. Since the research period is relatively short, and spatial reorganization typically takes a long time, I only estimate the short-term effect of transit on rents. This effect represents direct utility to individuals from residing in accessible areas internalized into asked rents. The estimated effect is not likely to include utility stemming from changes in zoning, densification, gentrification, or other slowly changing long-term effects of public transportation projects.

4.2. IV Model

The estimated Instrumental Variable model is identical to the baseline model except for $RCMA_I^{PT}$ being instrumented for using information on major transportation events. The instrument is a dummy variable indicating if the apartment is in an area where bus services were already tendered since the beginning of the research period,[46] or a new train station was inaugurated up to 1 kilometer away from the apartment. As in the baseline model, I estimate IV models both linearly and with automatic selection of controls, augmenting the dataset with two-way interactions and using the triple selection algorithm developed at Chernozhukov et al (2015).[47]

This approach addresses concerns of endogeneity in the allocation of public transit. If allocation is correlated with factors affecting rents, the baseline estimation would be biased. The IV approach exploits only variation within addresses, stemming from major transportation events: a bus tender or inauguration of a new train station. The estimated effect is the Local Average Treatment Effect. Tenders during the research period took place in many different urban contexts, but not in Tel Aviv or Jerusalem who compose a large share of the ads in the dataset. This might create a difference between the estimations in the baseline and IV models. I have argued in the empirical context section that the timing of major transportation events (hence, within-address variation in $RCMA^{PT}$) is exogenous, and the exclusion restriction is satisfied.

The Rank condition depends on the correlation between the conditional instrument and treatment. Figure 6 in section 3.1 above visualizes this correlation. There is no

[46] Specifically, for each ad i, I mark all bus stations within a one-kilometer radius from the apartment. I sum the number of bus stops-at-station in those marked stations that are part of clusters that were or were not tendered since the beginning of the research period. The instrument gets a value of 1 if the share of already tendered stops-at-station relevant to the apartment exceeds 50%.
[47] This algorithm applies on similar reasoning to the double selection algorithm described above.

clear trend in the activity[48] before tenders, but there is a large spike, followed by an upward trend in activity post tender. This finding supports the strength of the instrument. More formally, first stage regressions for the IV, and IV-LASSO models indicate a strong relation between the instrument and changes in $RCMA^{PT}$, with F test statistics exceeding 1500 (First stage results reported in appendix table A.2).

4.3 Causal Forest Models

The effect of public transportation services on rents is theoretically expected to depend on the geographic, urban and demographic context of the apartment and its area. This dependence could make the estimation of an average treatment effect in a specific sample uninformative in other empirical contexts. Estimation of a heterogeneous treatment effect in a large sample consisting of apartments in different areas might reveal patterns of dependence of the treatment effect in characteristics of the urban context. These patterns are likely to have better external validity than the average treatment effect, and can better inform planners, researchers and policy makers considering different alternative allocations of public transportation resources.

In my context heterogeneity is difficult to uncover with traditional methods and data. Linear regressions, the exclusive workhorse in the literature, allow only shallow exploration of heterogeneity across a small number of predetermined dimensions. To better explore heterogeneity in the treatment effect I estimate Causal Forests,[49] a standardized machine learning model designed for estimation of heterogenous treatment effects. A brief introduction to Causal Forests is supplied in appendix C.

I estimate the model with a set of spatial time-invariant features[50] and the same time-variant features described in the linear models above. I apply a newly developed procedure to incorporation of fixed effects to the model. The procedure aims to incorporate information about location and district-dependent trends when partialling-out confounders, while maintaining the ability to estimate the role of time-invariant

[48] Number of times a bus stops at the station in a regular weekday.
[49] Athey & Wager (2018).
[50] All spatial features are defined as the average value of the feature in circles with radiuses of 500, 1500 and 5000 meters centered around the apartment. The features originate either from the OD_MAT dataset (2018-2019 average, time invariant) or official statistical area level statistics from CBS_DATA (annual, time variant). OD_MAT based features are the density of morning in and out-commutes and evening in-commuters, proxying for population density, worker density, and recreational amenities accordingly. CBS_DATA based features are: population density, share non-Jewish, share male, share ultra-orthodox, Socio-Economic Status index, share in age groups 0-19, 20-39, 40-59, 60 plus, and distance to closest shore.

features in the determination of heterogeneity. The procedure is a trivial extension to the semi-parametric difference-in-differences estimator presented in Abadie (2005)[51] for data with multiple periods and a continuous treatment.

Estimation Procedure:

Denote X the set of controls, Y the dependent variable (log asked rents), and W the treatment variable ($\log(RCMA^{PT})$).

1. Divide covariate matrix X to time-variant and time-invariant features, X^{tv} and $X^{constant}$ accordingly.
2. Demean X^{tv}, Y, W by address id and time-district group membership.[52]
3. Estimate with separate regression forests the demeaned dependent and treatment variables, conditional on X using only information on the demeaned time-variant part of the controls:
   $\hat{Y}_i^{demeaned} = f\left(X_i^{tv,demeaned}\right), \hat{W}_i^{demeaned} = g(X_i^{tv,demeaned})$
4. Estimate a causal forest using the demeaned original and predicted dependent ($Y_i^{demeaned}, \widehat{Y_i}^{demeanded}$) and treatment ($W_i^{demeaned}, \widehat{W_i}^{demeanded}$) variables, the complete, not demeaned, covariate matrix X.

This procedure offers semi-parametric estimation of heterogenous treatment effects. Information from address and year-district features enters the model linearly when estimating the dependent and treatment variables. Estimation of the effect of time-variant features on the dependent and treatment variables, and of all features when estimating heterogeneity is then preformed a-parametrically as in regular Causal Forests. In addition, I recognize that addresses can entail information on heterogeneity by considering address clusters in the sampling and estimation procedures of the causal forest. I denote this model CF-FE.

I estimate a second version of the Causal Forest model which incorporates address information using a more standard fashion. In this version, I estimate a standard causal forest,[53] incorporating address information with a sufficient representation to

[51] See application of the method in Causal Forest estimation in Wang (2019).
[52] I use the implemented procedure available at the fixest R package (Berge, 2018).
[53] I deviate from standard implementation by estimating $\widehat{W} = g(X)$ with LASSO. This is done to avoid overfitting, which is an acute problem in predicting transit allocation in data that contains multiple spans of the same address or close addresses at the same time. This issue does not rise in the fixed-effects version of the model, since in this version the prediction is done on within-address variation.

group (address) membership using the approach presented in Johannemann et al (2019).[54] This approach offers a representation which is sufficient in the sense that no predictive power is lost, while the required number of features is considerably smaller than the number of groups of the categorical variable.

The approach relies on the sufficient latent state assumption: Group membership does not directly affect the dependent variable but is determined by a latent state which also affects observed variables. In my context, the assumption implies that location doesn't directly affect asked rent, but the latent state associated with location affects observed variables included in the model. For example, location can affect the probability of observing different values of the apartment' size, or residential density which is relevant to prediction. I denote this version of the Causal Forest CF-SR.

4.4. Difficulties in estimation

I will now present possible counter arguments to my identification strategy and argue that these difficulties don't pose a major threat in my context. I will focus on endogeneity and anticipation.[55]

Omitted Variable Bias

Omitted Variable Bias results from omission of a feature in the 'data generating process' affecting both the dependent and treatment variables. An example might be an opening of a new major occupation center who could increase rents at adjacent localities and cause planners to allocate more resources towards the same localities.

The allocation process described in the empirical context section supports the notion that planners have no possibility to effectively time major allocations such that they will correspond to other events. The date when a new bus operation agreement takes place is predetermined over a decade before the tender is being formulated. The argument for rail services lies on observed schedule overruns.

A fixed effects approach, utilizing within address variation and partialling-out district-level trends reasonably suffice to account for most omitted variables. This method

[54]In practice, for a model with k features, the approach requires plugging in another k features, each including the average by group (address id) value of the original feature.

[55] I do not explicitly argue for the model's strength against measurement error, which can be seen as a type of endogeneity and is always present to some degree, but I will point out that this paper relies on a more accurate and higher resolution data than prevalent in the literature.

does not rule out minor changes in the network corresponding to other events not observed in my data. The estimated Instrumental Variable model accounts for this possible bias as well by exploiting only variance stemming only from timing of major transportation events. The LASSO and causal forest models address concerns of misspecification due to non-trivial functional forms of included variables.

Simultaneity

An example of codetermination of the treatment and dependent variables in my context could be an ultra-orthodox community settling in a locality. In the Israeli context, It's plausible to claim that such a community can affect residential costs and transit allocation in their area. These communities can affect specific issues in the network troubling them: increasing frequency or rerouting specific lines. But, as discussed above, cannot change the timing of major transportation events. The IV approach, exploiting only variation stemming from such events addresses this issue.

Anticipation

The housing market can react to expected changes in transit supply years before they occur.[56] I address this issue by estimating the effect on rents instead of sales price. Tenants gain no extra utility from living near a still inactive transportation project. Thus, they will not be willing to pay more for apartments in these areas, compared to areas without planned improvements. This choice largely mitigates, though does not eliminate the problem. There might still be some anticipation effect due to an increase in house prices resulting in tougher negotiation by landlords, lowering of rents in apartments adjacent to large projects due to the disutility from living near a construction site,[57] or households looking to settle in a new area expecting improvement in allocation, and willing to absorb lack of services in the early period.

5. Results

5.1 The Average Treatment Effect

Table 4 presents the average treatment effect estimated using the different models described in the methodology section, and table 5 presents specifications of the

[56] See for example Yiu & Wong (2005), Agostini & Palmucci (2008), Liang et al (2018).

[57] This argument is relevant only to rail projects and should not pose a major problem since most new rail stations inaugurated are located at the outskirts of the urban area, not manifesting major disturbances during construction.

baseline model with different choices of time and geographic definitions for region specific trends. The Average Treatment Effect is always of an economically negligible magnitude. To make this point more explicit, the average 2013-2019 log difference in $RCMA^{PT}$ for addresses appearing in the dataset in both years is 0.19. Applying the baseline elasticity, the negative effect on the average rent quantifies to roughly 0.2% of the its total cost.

Table 4

The Average treatment effect of Residential Commuter Market Access on rents

	Baseline	LASSO	IV	LASSO-IV	CF-FE	CF-SR
Average Treatment Effect	-0.001 (0.003)	-0.002 (0.003)	0.022 (0.061)	-0.055 (0.06)	0.009 (0.006)	-0.009*** (0.003)
R^2 (Within, adjusted)	0.583	0.600	0.583	0.600		
N - observations	760,568					
N - unique addresses	147,283					

Note: Models described in text, standard errors clustered by address id shown in parentheses.

Table 5

$RCMA^{PT}$ coefficients with different specifications of time-geographic trends

Geo\Time	year	transportation period	month
locality	-0.009** (0.004)	-0.022*** (0.004)	-0.023*** (0.004)
Natural area	-0.003 (0.003)	-0.015*** (0.003)	-0.015*** (0.003)
Sub district	0.000 (0.003)	-0.008*** (0.003)	-0.008*** (0.003)
District	-0.001 (0.003)	-0.010*** (0.003)	-0.010*** (0.003)
none	-0.007*** (0.003)	-0.018*** (0.003)	-0.018*** (0.003)

Note: Standard errors clustered by address id shown in parentheses

5.2 Heterogeneity in the effect

The estimated zero magnitude of the average treatment effect doesn't imply that there is no important heterogeneity. I estimate heterogeneity in the effect for several groups of interest: Apartments located near Mass Transit systems,[58] or in areas with high

[58] Proximity is defined as location up to 1,000 meters from an active station, consistent with standard practice in the literature (see in Ingvardson & Nielsen, 2018).

density, accessibility, or high Socio-Economic status residents. Estimated models are variants of the baseline model including an interaction term between treatment and group membership. Specifically, I estimate:

$$(7)\ \log(rent)_{ijrt} = \beta_0 + \beta_1 * \log(RCMA^{PT})_{jt} + \beta_2 * \left(\log(RCMA^{PT})_{jt} * \xi_i\right) + \mu_j + \psi_{rt} + \beta X_{ijrt} + \upsilon_{ijrt}$$

Where ξ_i is a dummy representing group membership of ad i. Table 6 presents the results. Apartments located in areas with a relatively high density of workers experience a positive economically meaningful effect, and apartments located in areas with high socio-economic residents are responsible for the negative (though negligible) average treatment effect.

Table 6

Heterogeneity in the effect of Residential Commuter Market Access on rents - Specified subgroups

Heterogeneity group	Baseline	Near Train	Near Light rail	Near BRT	Population density	Workers' density	Socio Economic Status	RCMA^{Car}	RCMA^{PT}
Definition	All	0-1000m	0-1000m	0-1000m	Top Quartile	Top Quartile	Top Quartile	Top Quartile	Top Quartile
Interaction term		-0.000 (0.001)	0.036* (0.021)	0.086*** (0.008)	0.002*** (0)	0.069*** (0.008)	-0.086*** (0.006)	-0.000*** (0)	-0.000*** (0)
Average Treatment Effect	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)	-0.009*** (0.003)	-0.001 (0.003)	-0.004* (0.003)	0.020*** (0.003)	-0.000 (0.003)	-0.001 (0.003)
R^2 (Within, adjusted)	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.583
N - in interaction group		105,758	22,144	65,800	190,142	190,147	190,146	190,142	190,142
N - observations	760,568								
N - unique addresses	147,283								

Note: Standard errors clustered by address id shown in parentheses. The choice of 1000 meters cutoff for proximity to stations is based on standard practice in the literature (See Ingavrdson & Nielsen, 2018)

I further explore these patterns by examining the effect in deciles of Socio-Economic status and workers' density. The effects are estimated with a variant of equation (7) including 9 interactions of $RCMA^{PT}$ with membership in each decile of the interacted variable. A separate model is estimated for each variable. Figures 9 and 10 present the results. The effect is decreasing in Socio-Economic status. This might be the result of residents in high Socio-Economic status areas, who typically use less transit, experiencing mainly transit's negative externalities.[59] The effect remains fairly constant across the distribution of workers' density, with a hike in the effect in the top

[59] For example, noise, or less available parking resulting from transit services.

decile, this might the result of better ability to rely on public transit for residents in areas with a lot of job opportunities.

Figure 9

Estimated treatment effect by Socio-Economic decile

with 95% confidence intervals

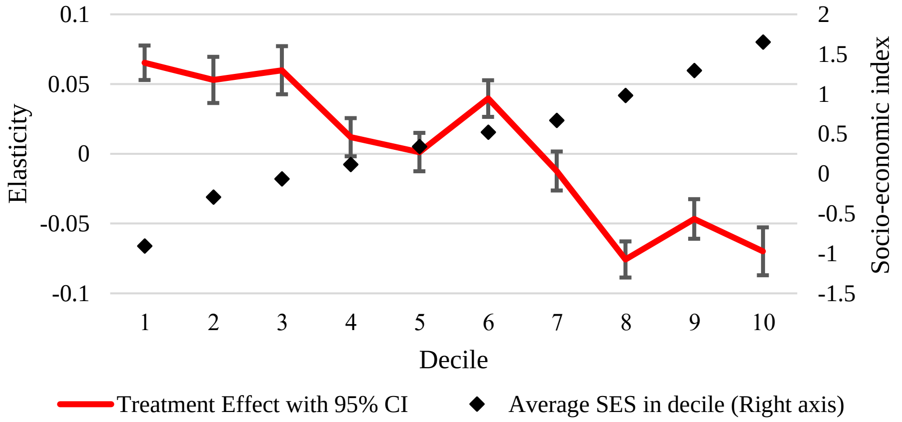


Figure 10

Estimated treatment effect by workers' density decile

with 95% confidence intervals

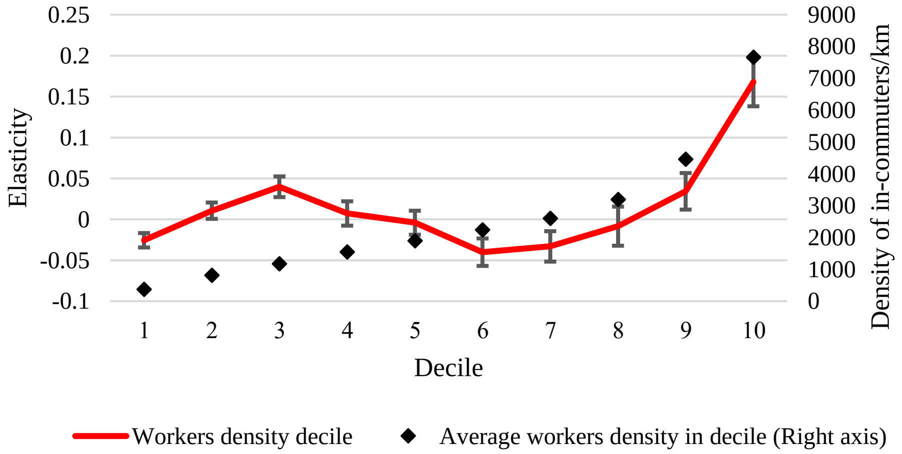


Table 6 also presented a significant positive effect to apartments located near the BRT system in Haifa (Metronit), and to a lesser extent, near the Jerusalem Light rail. To compliment this finding, I estimate a model explicitly designed to examining benefits from proximity to a train station.[60] Specifically, I limit the sample to apartments located up to 3 kilometers away from one of the 15 stations inaugurated during the research period. An observation in this model is an ad i, located at address j, within a 3 kilometers radius of station s, last appearing in the dataset at year t. I estimate:

[60] A similar model for other modes of transportation (BRT, Light rail) cannot be estimated with my data since no new light rail stations were inaugurated during the research period, and the BRT system started its full operation in the beginning of the research period.

$$(8)\quad \log(rent)_{ijst} = \beta_0 + \beta_1 * post_{st} + \beta_2 * [proximity_{js} * post_{st}] + \mu_j + \lambda_t + \beta X_{ij} + v_{ijst}$$

With 'post' representing a dummy variable for whether the relevant station is active, proximity is a dummy indicating whether the ad is in the inner or outer parts of the circle surrounding the station, X is a vector of apartment specific features identical to the baseline model, and μ and λ represents address and year effects accordingly. The estimation relies on the difference between the post-after difference in apartments located close to the station and apartments in the outer parts of the circle surrounding the station. Note that this estimation does not rely on the Commuter Market Access concept guiding the rest of the estimation in this paper. Table 7 presents the results.

Table 7
The effect of proximity to train stations on rents

	Constant effect	Heterogeneity by distance				
Interaction group (distance in meters from station)	0-1000	0-200	200-400	400-600	600-800	800-1000
Difference in Differences	0.012*** (0.004)	-0.006 (0.06)	0.022** (0.01)	0.014* (0.009)	0.011 (0.007)	0.009* (0.006)
R^2 (Within, adjusted)	0.603	0.603				
N - observations	47,837	47,837				
N - unique addresses	10,779	10,779				
N - observations in treatment group	10,006	62	1,076	1,833	3,044	3,991

Note: Standard errors clustered by address shown in parentheses. The control group is always defined as observations located 1000-3000 meters from stations.

There is a small positive effect, monotonically decreasing with distance from the station.[61] This result implies that, like the positive effect for the BRT and Light rail reported at table 6, the rents market internalizes in the short-term utility from living near a Mass Transit system. The higher effect for these systems can be explained by higher visibility, mode preferences or other considerations not embodied in my specified commuting cost function, expressed only in terms of travel time.

The reason the positive effect was not found for trains in table 6 can be due to train stations affecting the rents market through channels other than accessibility. The different effect can also result from the different comparison group. Namely, focusing on variance between the core and the periphery of the new stations' catchment areas

[61]There is exception to the monotonicity in the closest proximity group (0-200), this might be the result of negative externalities in the immediate surrounding of a train station, or a spurious result stemming from the small number of apartments in this proximity group in the sample.

emphasizes patterns of re-organization.[62] Another possible explanation is the different characteristics of apartments near new train stations, compared to apartments near existing stations. The new inaugurated stations are mostly spread across peripheral and suburban regions, and mostly at the outskirts of the urban area. Most existing stations are located at central regions and within cities.

I now turn to a more systematic examination of the patterns of heterogeneity using the Causal Forest model. Table 8 presents summary statistics for both estimated Causal Forests.[63] I assess the different models' fit using the omnibus test developed at Chernozhukov et al (2018).[64] The test results reported at table 8 show that both models capture the average treatment effect and heterogeneity in the underlying signal quite well, but there is a clear advantage to the CF-FE model. Thus, I choose the CF-FE model for further exploration of heterogeneity.

Table 8

Summary statistics for causal forest models

		CF-FE	CF-SR
Results	Average Treatment Effect	0.009 (0.006)	-0.009*** (0.003)
	Share with positive effect	51.0%	43.4%
Omnibus Calibration Test	Mean Forest Prediction	1.226*** (0.514)	1.351*** (0.159)
	Differential Forest Prediction	1.003*** (0.027)	1.592*** (0.048)
Parameters	sample.fraction	0.5	0.441
	mtry	28	18
	min.node.size	5	10
	alpha	0.05	0.05
	imbalance.penalty	0	0
	honesty.fraction	0.50	0.50
	honesty.prune.leaves	TRUE	TRUE
Data	Number of observations	760,568	760,568
	Number of unique addresses	147,283	147,283

Note: Models described in text, standard errors clustered by address id shown in parentheses.

[62] A thorough discussion of Growth Versus Re-Organization in the effect of transportation on economic phenomena appears in Redding & Turner (2015).
[63]The model's parameters were chosen by automatic parameter tuning, using 5-fold cross-validation, or the default when the process failed to find a better alternative. The decision rule for the 'best' parameters was developed at Nie & Wager (2021) and is readily implemented in the grf R package.
[64] This test is discussed specifically for Causal Forests in Athey & Wager (2019). Presentation of the test based on the description at Athey & Wager (2019) appears in Appendix C.

Figure 11 visualizes the heterogeneity of the treatment effect.[65] The effect is centered around zero, but there is considerable variation in the estimated effect - as expected considering the variation in urban contexts in the dataset. Though there is considerable variation, the magnitude of the effect is usually small. Only 14.3% of the observations' point estimates are of absolute elasticity larger than 0.25.[66]

Figure 11

Distribution of the estimated treatment effect

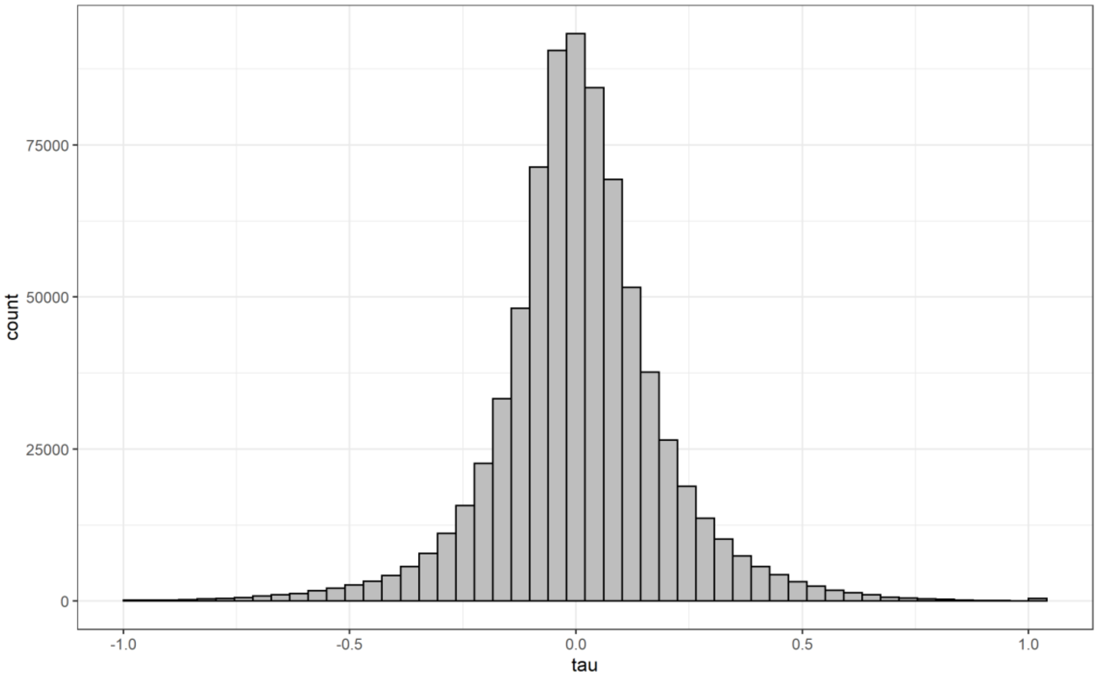
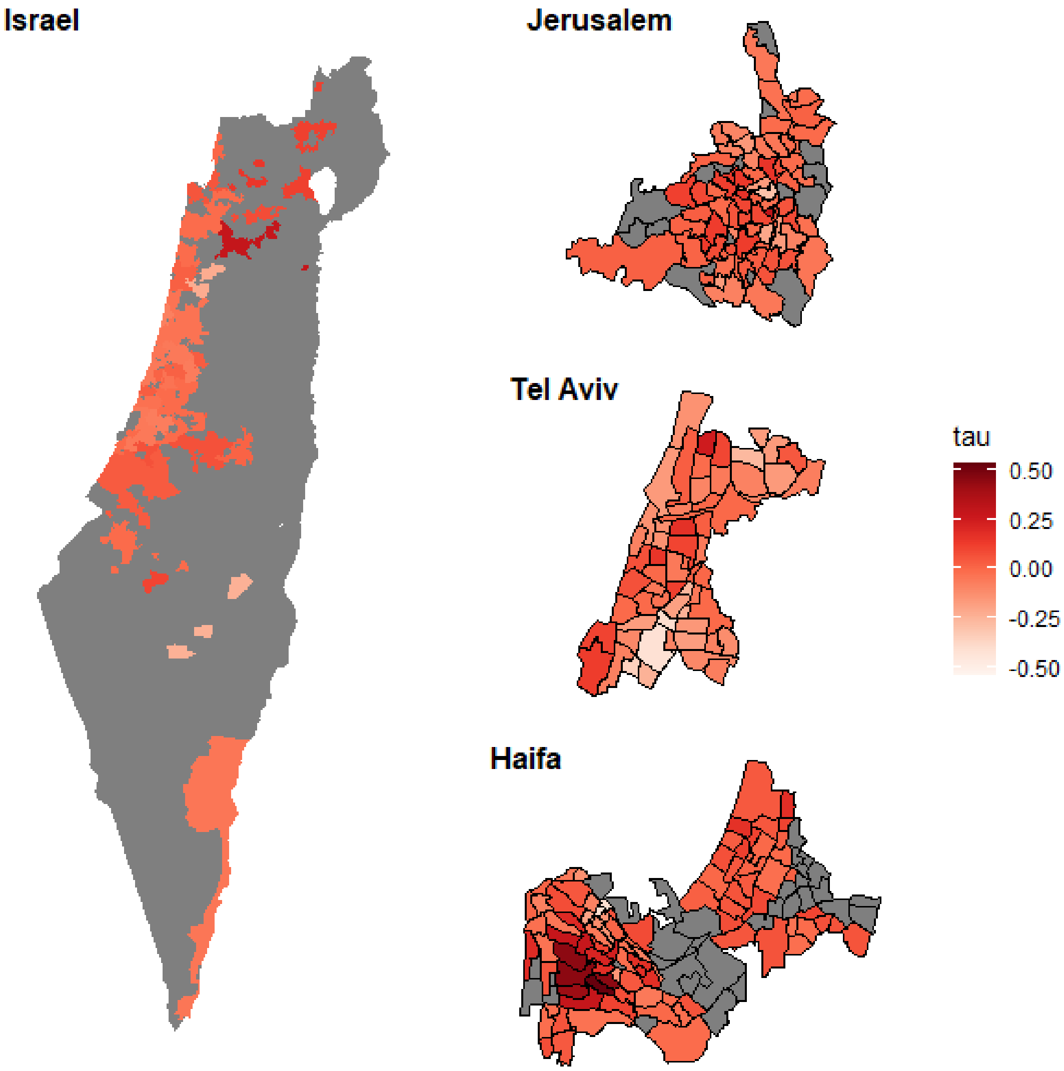


Figure 12 displays the average treatment effect by transportation polygons. The effect seems to be especially high in the Jezreel Valley, Jerusalem, and Western Negev. Main contribution to accessibility in these areas during the research period is the inauguration of new rails. This finding supports results from the difference-in-differences estimation for newly inaugurated train stations. Examining within-city variation, the effect is relatively strong in neighborhoods in Jerusalem adjacent to its core, and the western slopes of the Carmel Mountain in Haifa: areas characterized by mixed-use zoning.[67]

[65] For illustrative purposes, the estimated effect is winsorized to an absolute value of 1.
[66] This cutoff implies roughly a 5% change in rents for the average 2013-2019 $RCMA^{PT}$ difference.
[67] The large effect throughout Haifa, and especially in the western slopes of the Carmel west Haifa also lends support to the unconditional results for apartments close to the Metornit reported in table 6.

Figure 12

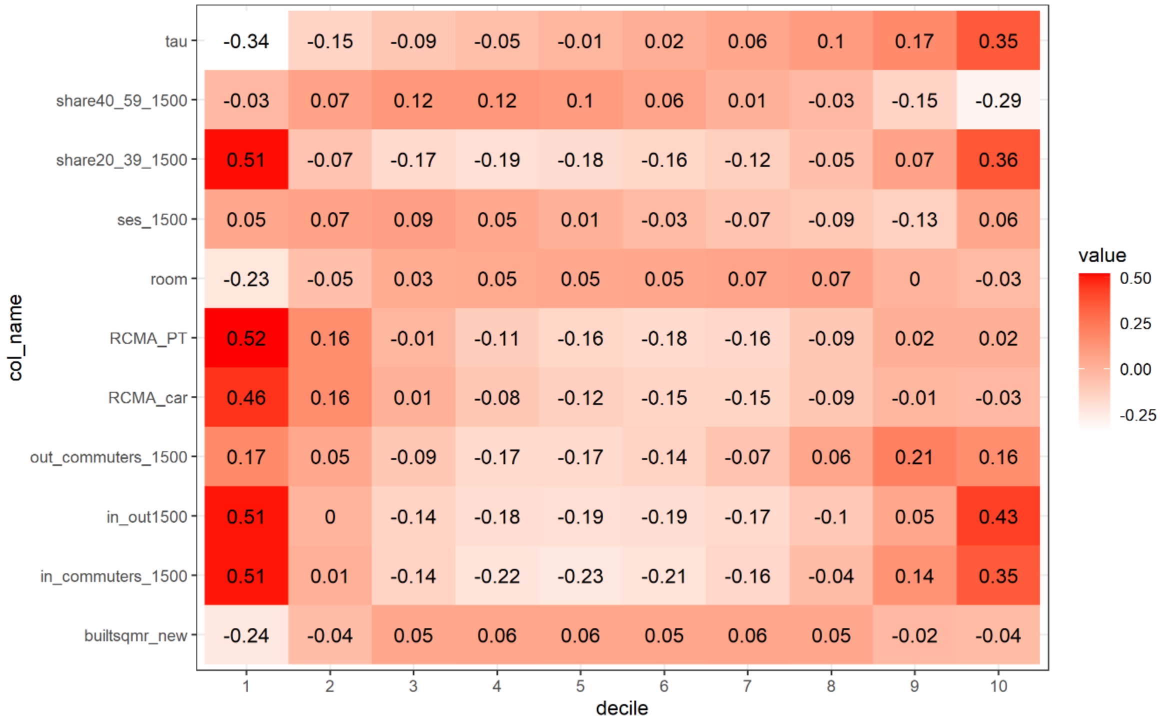
Spatial Distribution of the Treatment Effect



I explore patterns in the heterogeneity by examining average apartment characteristics in deciles of the estimated Effect. Figure 13 presents average normalized values of features corresponding to speculated determinants of heterogeneity, and the average treatment effect in each decile. Apartments with a low treatment effect are relatively small and characterized by very high accessibility, worker and residents' density, and worker-resident ratio. Surprisingly, let alone size, apartments in the higher end of the distribution have similar characteristics. These apartments are also characterized by a high share of 40-59 aged population. Note that like the interpretation of an interaction with the treatment variable in a linear regression, different treatment effects along the covariate space are not necessarily caused by the features examined.

Figure 13

Normalized ad characteristics along the deciles of the treatment effect



To better understand determinants of heterogeneity I estimate the Best Linear Projection of the model's covariates on the treatment effect using a doubly robust estimation. The interpretation of the coefficients is similar to the interpretation of a linear regression of the estimated idiosyncratic treatment effect on the covariates. To reduce collision, I omit spatial variables not defined by the 1500 meters radius. I also add: $\log(RCMA^{PT})$, ratio between in and out commuters in the area (representing variation in the degree of Mixed-use zoning), and dummies for addresses located less than a kilometer from any Mass Transit System's station. All variables are standardized to conduct meaningful comparison of the magnitude of the effects. Top 15 variables by absolute coefficient 's magnitude are presented in table 9.

The most important determinants of the treatment effect are the density parameters. The rents market internalize utility to residents in areas that have many possible users: A one standard deviation increase in residential density[68] implies a 0.11 increase in the elasticity of rents to $RCMA^{PT}$. The estimation also displays a large positive effect for the ratio between workers and residents' density, and a negative effect for workers'

[68] Proxied for using the number of individuals leaving the area for their morning commute.

density. To further explore this relationship, I present binned scatterplots of the relation between the in-out commuters' ratio and the estimated treatment effect in figure 14. I present the relation both in its raw form, and in residualized values.[69]

Table 9

Best Linear Projection of the treatment effect on CF-FE covariates, Top 15 features by absolute magnitude of coefficient

	Coefficient	Robust Standard Error
out_commuters_1500	0.113***	(0.036)
in_commuters_1500	-0.081***	(0.028)
in_out1500	0.064***	(0.018)
ln_RCMA_PT	-0.06***	(0.013)
ses_1500	0.045***	(0.01)
room	0.037*	(0.02)
share40_59_1500	-0.035***	(0.01)
ln_RCMA_car	-0.028**	(0.012)
male_share_1500	-0.028**	(0.011)
dist_to_metronit_0_1000	-0.025***	(0.007)
evening_commuters_1500	-0.023	(0.042)
dist_to_light_rail_0_1000	0.019**	(0.008)
builtsqmr_new	-0.015	(0.022)
share20_39_1500	0.014	(0.015)
ultra_orthodox_1500	0.01	(0.011)

Note: Doubly Robust estimation, all variables standardized to have a mean of zero and variance 1

There is an optimum level for the effect of in-out commuters' ratio on the tendency of public transportation to increase rents. This finding emphasizes the importance of mixed-use zoning in public transportation effectiveness. The density findings imply a significantly lower utility to residents from public transit services in areas with separate-use zoning such as suburbs who are characterized by low residential density and very low workers' density, or employment hubs with low residential density.

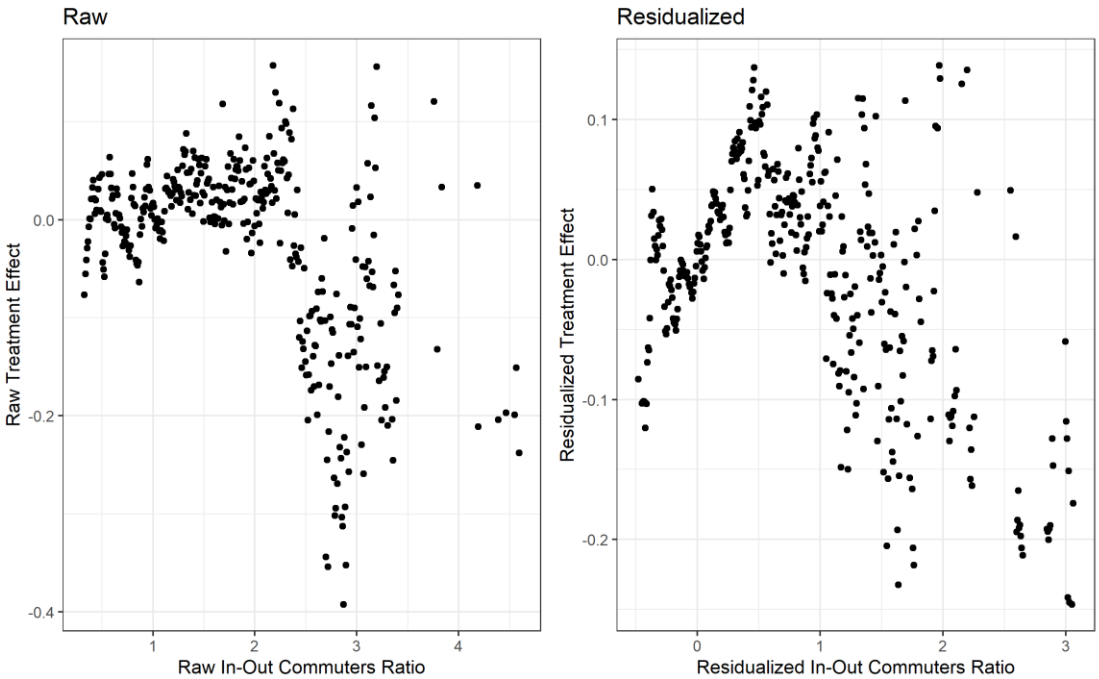
Another finding reported in table 9 is that the treatment effect is declining with accessibility. This finding, along with the distribution presented in figure 13, might hint at the existence of an upper bound for the level of accessibility which still influences rents. Residents in areas with accessibility levels above this bound still

[69] Residualization is performed by a linear regression of all variables used for the best linear projection (table 9) on both the in-out commuters' ratio and the treatment effect.

suffer from adverse effects of large volumes of public transportation, e.g., noise, pollution, and crowdedness, which can explain the negative direction of the effect.

Figure 14

The relationship between in-out commuters' ratio to the treatment effect



Note: The plots are based on all (760,568) observation in the dataset, binned to 500 dots based on their in-out commuters' ratio. The residualized plot presents the relationship after residualization done by linear regressions of the in-out commuters' ratio and of the treatment effect on the same variables used for the Best Linear Projection (table 9).

Socio-Economic Status seems to be an important demographic feature, positively impacting the effect. This contrasts the decreasing gradient estimated in the linear model, and the somewhat similar pattern presented in figure 13. The difference is probably due to the linear projection conditioning the socio-economic status with other features of the area, implying a ceteris paribus interpretation of the effect.

Other demographic features are consistent with the hypothesis that effective public transportation requires high density of potential users. Areas with a larger share of the population not prone to using public transportation (aged 40-59, males) experience a lower effect.[70] The opposite signs on apartment size and number of rooms might also

[70] The positive, but statistically insignificant sign on the share of population aged 20-39 and the share of ultra-orthodox point at the same direction.

point towards that direction. A larger effect for households residing in apartments with a large number of small rooms: young singles, or young families – both typically more frequent public transportation users.

The higher effect for apartments located near the Metronit reported at table 6 doesn't seem to be caused by proximity, but rather by other characteristics of this area. In contrast, accessibility stemming from the Jerusalem Light Rail does seem to contribute to the size of the treatment effect over and above the reduction in travel times. The effect of proximity to a train station is small, hence not presented here, consistent with results from the linear model reported in table 6.

Conclusion

Within a hedonic framework, utility to individuals from public transit in their residential area would be internalized by the rents market. I utilize high resolution data, a theoretically grounded measure of accessibility and both causal machine learning and standard econometric methods to explore this effect. I find that, on average, there is no short-term effect. I do find important patterns of heterogeneity: the effect is larger in areas hosting a large pool of possible users (higher residential density, and demographic composition characterizing transit users). The effect is further enhanced in areas characterized with mixed-use zoning.

I also find evidence that the rents market internalize utility stemming from proximity to Mass Transit stations over and above the utility stemming from reduction in travel times. This finding especially holds for train stations inaugurated during the research period, and for the Jerusalem Light Rail. Possible explanations are high visibility of Mass Transit projects, compared to a lower visual signal of a well-developed bus network. Other possible explanations are mode preference, amenities from urban development around stations, or better protection from possible future planning decisions re-routing transit lines serving the area.

An important avenue for future research is the examination of patters of heterogeneity in the long-term effects of public transit on rents. The long-term effect includes both utility from increased accessibility, and from the role public transportation might play in shaping the spatial distribution of economic activity.

Bibliography

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica*, *83*(6), 2127-2189.

Ahlfeldt, G. M., & Feddersen, A. (2017). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, *18*(2), 355-390.

Albouy, D., & Lue, B. (2015). Driving to opportunity: Local rents, wages, commuting, and sub-metropolitan quality of life. *Journal of Urban Economics*, *89*, 74-92.

Amior, M., & Manning, A. (2019). Commuting, migration and local joblessness. *Discussion papers (No. 1623), Centre for Economic Performance.*

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton university press.

Arestis, P., & Gonzalez-Martinez, R. A. (2017). Housing market in Israel: Is there a bubble?. *Panoeconomicus*, *64*(1), 1-16.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353-7360.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, *47*(2), 1148-1178.

Athey, S., & Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational Studies*, *5*(2), 37-51.

Banerjee, A., Duflo, E., & Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, *145*, 102442.

Bank of Israel (2010). Chapter 6. *Bank of Israel Annual report – 2009,* Jerusalem.

Bank of Israel (2015). Chapter 6. *Bank of Israel Annual report – 2014,* Jerusalem.

Barak, A. (2019). The effect of public transportation on employment in Arab society. *discussion papers (No. 2019.03)*, Bank of Israel.

Baum-Snow, N. (2007). Did highways cause suburbanization?. *The Quarterly Journal of Economics*, *122*(2), 775-805.

Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960-2000. *American Economic Review*, *100*(2), 378-82.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, *99*(3), 435-448.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608-650.

Bleikh, H. (2018). Back and Forth: Commuting for Work in Israel. *Policy papers (No. 2018.05)*, Taub Center for Social Policy Studies in Israel.

Brennan, M., Olaru, D., & Smith, B. (2014). Are exclusion factors capitalised in housing prices?. *Case Studies on Transport Policy*, *2*(2), 50-60.

Büchel, K., Puga, D., Viladecans-Marsal, E., & von Ehrlich, M. (2019). Calling from the outside: The role of networks in residential mobility.

Cao, K., Diao, M., & Wu, B. (2019). A big data–based geographically weighted regression model for public housing prices: A case study in Singapore. *Annals of the American Association of Geographers*, *109*(1), 173-186.

Caspi, I. (2016). Testing for a housing bubble at the national and regional level: the case of Israel. *Empirical Economics*, *51*(2), 483-516.

Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, *30*(4), 457-490.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018). Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized

Experiments, with an Application to Immunization in India, *(No. w24678)*. National Bureau of Economic Research.

Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Post-selection and post-regularization inference in linear models with many controls and instruments. *American Economic Review*, *105*(5), 486-90.

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.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, *1*(1), 269-271.

Dingel, J. I., & Tintelnot, F. (2020). *Spatial economics for granular settings* (No. w27287). National Bureau of Economic Research.

Dorantes, L. M., Paez, A., & Vassallo, J. M. (2011). Analysis of house prices to assess economic impacts of new public transport infrastructure: Madrid Metro Line 12. *Transportation research record*, *2245*(1), 131-139.

Dovman, P., Ribon, S., & Yakhin, Y. (2012). The Housing Market in Israel 2008-2010: Are House Prices a 'Bubble'?. *Israel Economic Review*, *10*(1).

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616-52.

Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, *79*(4), 1407-1440.

Duranton, G., Morrow, P. M., & Turner, M. A. (2014). Roads and Trade: Evidence from the US. *Review of Economic Studies*, *81*(2), 681-724.

Falkov, I. (1982). Trains in Israel: Past, Present, Future. *Israeli Railways Ltd. Press, Lod.*

Friedman, Y. (2019), Private transportation in Israel: An analysis of developments in the past two decades. *Selected Research and Policy Analysis Notes 2019(1)*, 50-61, *Bank of Israel*, Jerusalem.

Frisch, R. & Tzur, S. (2010), Investment in transport infrastructure, commuting and wages. *Seker Bank Israel, 83*, 7-34.

Greenwald, D., Grossman G., & Levi, A. (2018). Does greater public transit access increase employment for the Israeli-Arab Population? A Preliminary Analysis. *M-RCBG Associate Working Paper Series*, No. 95

Ingvardson, J. B., & Nielsen, O. A. (2018). Effects of new bus and rail rapid transit systems–an international review. *Transport Reviews*, *38*(1), 96-116.

Ida, Y., & Talit, G. (2018). What we can learn 17 years after the reform in public bus transportation in Israel. *Case Studies on Transport Policy, 6*(4), 510-517.

Israel, E., & Cohen-Blankshtain, G. (2010). Testing the decentralization effects of rail systems: Empirical findings from Israel. *Transportation Research Part A: Policy and Practice*, *44*(7), 523-536.

Leck, E., Bekhor, S., & Gat, D. (2008). Equity impacts of transportation improvements on core and peripheral cities. *Journal of Transport and Land Use*, *1*(2), 153-182.

Li, H., Wei, Y. D., Wu, Y., & Tian, G. (2019). Analyzing housing prices in Shanghai with open data: Amenity, accessibility and urban structure. *Cities*, *91*, 165-179.

Liang, X., Liu, Y., Qiu, T., Jing, Y., & Fang, F. (2018). The effects of locational factors on the housing prices of residential communities: The case of Ningbo, China. *Habitat International*, *81*, 1-11.

Mayer, T., & Trevien, C. (2017). The impact of urban public transportation evidence from the Paris region. *Journal of Urban Economics*, *102,* 1-21.

Mohammad, S. I., Graham, D. J., Melo, P. C., & Anderson, R. J. (2013). A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice*, *50*, 158-170.

Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, *108*(12), 3855-90.

Nie, X., & Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, *108*(2), 299-319.

Organisation for Economic Co-operation and Development. (2011). *How's life?: measuring well-being*. Paris: Oecd.

Portnov, B., Genkin, B., & Barzilay, B. (2009). Investigating the effect of train proximity on apartment prices: Haifa, Israel as a case study. *Journal of Real Estate Research*, *31*(4), 371-395.

Raz-Dror, O. (2019). The changes in rent in Israel during the years of the housing crisis 2008–2015. *Israel Economic Review*, *17*(1).

Redding, S. J., & Sturm, D. M. (2008). The costs of remoteness: Evidence from German division and reunification. *American Economic Review*, *98*(5), 1766-97.

Redding, S. J., & Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of regional and urban economics* (Vol. 5, pp. 1339-1398). Elsevier.

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 931-954.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, *90*(6), 1257-1278.

Rodriguez, D. A., & Targa, F. (2004). Value of accessibility to Bogotá's bus rapid transit system. Transport Reviews, 24(5), 587-610.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), 34-55.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74-104.

Severen, C. (2021). Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification. *Working papers (No. 18-14), Federal Reserve Bank of Philadelphia.*

Soffer, Y & Suhoy, T. (2019) Getting to Work in Israel: Locality and Individual Effects. *discussion papers (No. 2019.02).* Bank of Israel.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), 1228-1242.

Wang, G. (2019). The Effect of Medicaid Expansion on Wait Time in the Emergency Department. *Management Science*.

Wardrip, K. (2011). Public transit's impact on housing costs: a review of the literature, *Insights from Housing Policy Research*, Center for Housing Policy.

Wittowsky, D., Hoekveld, J., Welsch, J., & Steier, M. (2020). Residential housing prices: impact of housing characteristics, accessibility and neighbouring apartments–a case study of Dortmund, Germany. *Urban, Planning and Transport Research*, *8*(1), 44-70.

Yakhin, Y & Gamrasni, I. (2021) The housing market in Israel: Long-Run Equilibrium and Short-Run Dynamics. *discussion papers (No. 2021.08).* Bank of Israel.

Yiu, C. Y., & Wong, S. K. (2005). The effects of expected transport improvements on housing prices. *Urban studies*, *42*(1), 113-125.

Zhang, M., & Yen, B. T. (2020). The impact of Bus Rapid Transit (BRT) on land and property values: A meta-analysis. *Land Use Policy*, *96*, 104684.

<u>Government decisions</u>

Israeli Government Decision No. 1301, 1997. Opening the Public Transportation Sector to Competition. Jerusalem.

Israeli Government Decision No. 1539, 2010. The five-year plan for economic development in 13 localities of minorities. Jerusalem

Israeli Government Decision No. 3988, 2011. Establishment of a National Authority for Public Transportation and Metropolitan Transport Authorities. Jerusalem

Israeli Government Decision No. 922, 2015. Five Year Economic Development Plan for Arab Society. Jerusalem

Stata Packages

Ben Jann, 2005. "MOREMATA: Stata module (Mata) to provide various functions," *Statistical Software Components S455001*, Boston College Department of Economics, revised 16 Aug 2021.

Sergio Correia & Paulo Guimaraes & Thomas Zylkin, 2019. "PPMLHDFE: Stata module for Poisson pseudo-likelihood regression with multiple levels of fixed effects," *Statistical Software Components S458622*, Boston College Department of Economics, revised 18 Nov 2019.

R packages

Berge, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. CREA Discussion Papers.

Julie Tibshirani, Susan Athey, Erik Sverdrup and Stefan Wager (2021). grf: Generalized Random Forests. R package version 2.0.2. https://CRAN.R-project.org/package=grf

Larmet V (2019). “cppRouting: Fast Implementation of Dijkstra Algorithm in R.” <URL:https://github.com/vlarmet/cppRouting>.

Mark Padgham (2019) dodgr: An R package for network flow aggregation Transport Findings, 2(14). URL https://doi.org/10.32866/6945

Tianqi Chen, Tong He, Michael Belesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mithell, Ignacio Cano, Tianyi Zhou, Mu Li, Junyuan Xie, Min Lin, Yifeng Geng and Yutian Li (2021). xgboost: Extreme Gradient Boosting. R package version 1.4.1.1. https://CRAN.R-project.org/package=xgboost

Appendix tables and figures

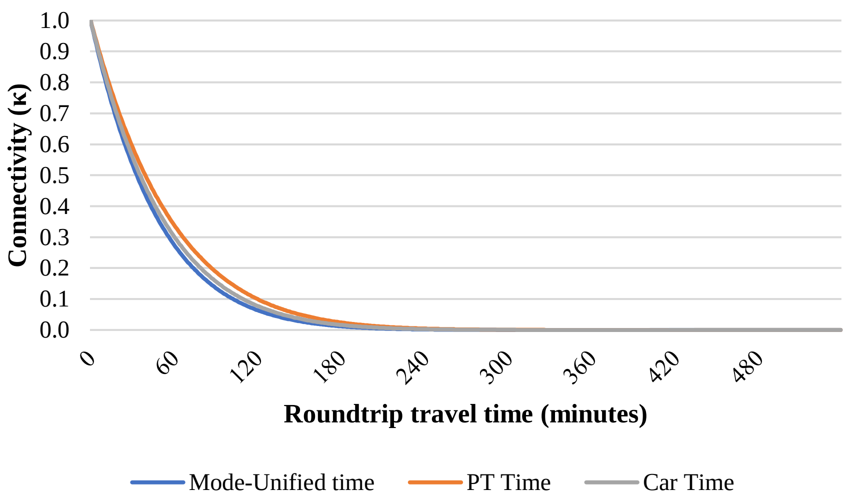
Table A.1

The effect of Residential Commuter Market Access on rents - First stage results

	IV	LASSO-IV
After Tender IV	0.01427*** (0.00121)	0.01434*** (0.0012)
R^2 (Within)	0.0125	0.0172
F test	1,563.70	1,582.00
Number of observations	760,568	

Figure A.1

Implied connectivity by different elasticities



Appendix A: Imputation of bus time-in-ride

The real bus travel time dataset, BUS_RIDES, used in this research only covers the years 2016-2019, raising the need to impute travel times for the earlier period. To do this, I construct a new dataset in which each row represents a distinct bus line in each direction, year, period, and hour of departure.[71] For each row I calculate characteristics including the average planned ride time in each half hour interval, total distance travelled, number of stops by activity,[72] all taken from BUS_ROUTES, and the median real travel time calculated from BUS_RIDES.[73]

To further improve predictive ability, I divide each ride to the separate edges composing it. I characterize each edge by its length, planned speed, and importance in the network.[74] I divide each of these characteristics to eight bins, and the edge is classified to one of the categories resulting from the interactions between the bins. I then sum the distance each line travels in each of these categories.

The prediction itself is done using a Stochastic Gradient Boosting Machine algorithm, as implemented in R's XGBoost package.[75] The target variable is the difference between the real and planned travel times. I use the difference instead of real travel times to maintain any line-specific knowledge known to the transportation planners but unknown to me. I train the model on data from the second transportation period of 2016 to the end of 2019 and test it on data from the first period of 2016. All model parameters are hyper tuned using 5-fold cross-validation.

Post estimation I sum the planned travel time with the predicted real-planned difference. Lastly, I impose that all predicted travel times will lie in the 10-120 kmh interval. Table A.1 presents Goodness of fit measures on the test set both in terms of minutes, and log minutes which can be interpreted as the deviation in percent.

[71] Departure groups are defined in half hour intervals during rush hours, and specific groups for rides departing before morning rush hour, between rush-hours and post afternoon rush hour.
[72] Drop-off only, Pick-up only, Both, and long refreshment stops.
[73] The median is calculated in two steps. I calculate it on the raw data, drop all observations whose ride time is either shorter than half, or longer than double the raw median. These observations contain obvious errors such as negative or close to zero ride times and unique events such as extreme congestion due to accidents or other extraordinary events. Finally, I calculate the median travel time of the subset of remaining observations. Lastly, I impose all median times to be in the 10-120 kmh interval.
[74] Defined as the share of all bus rides in the transportation period travelling in the same edge.
[75] Chen et al (2021)

Table A.2

Goodness of fit measures of bus times, imputation on test set

	minutes	log(minutes)
Mean Absolute Error	2.82	0.0629
Root Mean Squared Error	4.27	0.0932
R^2	0.982	0.977
N - train set	262,306	
N - test set	30,076	

To translate the imputed travel time from the entire ride to distinct edges, I multiply the planned travel time for each edge in the ratio between the estimated and planned travel times.

Appendix B: Definition of travel times

B.1. Public Transportation

I start by defining direct effective travel times between all public transportation stations in Israel.[76] I aim to calculate the travel time of a typical commute; hence I define travel time between area o and area d as the roundtrip journey: the sum of total travel time from o to d in the morning commute, and from d to o in the afternoon commute. I choose 6:30-9:30 as the relevant interval for the morning commute, and 14:30-17:30 as the relevant interval for the ride back.

$$(A1)\ \ Travel\ time_{od} = Travel\ time_{od}^{morning} + Travel\ time_{do}^{afternoon}$$

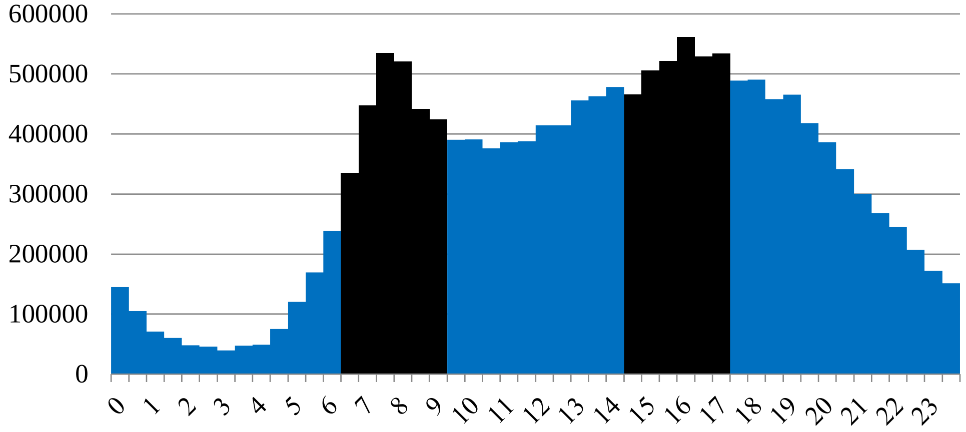
Travel between stations can occur by all modes of public transportation including walking. The actual travel time consists of both the waiting time (according to planned schedule for bus rides and the Metronit taken from BUS_SCHEDULE, actual schedule for other modes of transportation, taken from the LIGHT_RAIL and TRAIN_RIDES datasets), and the time in ride (median value for that period and time of day half hour interval, imputation process for buses in the early period appears in appendix A).

I allow walking between each two points (apartment to station, or station to station) up to one kilometer. The imputed time includes a constant of 2 minutes which is included to account for the need to arrive early due to some buses arriving early, and to increase the penalty on complicated rides where the replacement occurs between close stations. This constant is added to a function of the aerial distance which is defined as a walking speed of 4kmh in the first 400 meters, 3kmh in the 400-600 meters interval, 2kmh in the 600-800 meters interval and 1 kmh in the 800-1000 meters interval. Thus, the maximal walking journey is one kilometer long, which takes 30 minutes to complete. The gradual decrease in speed is designed to penalize, but not cancel out, accessibility which relies on long walks. A major advantage of the gradual slowdown approach is that it diminishes the phenomena of sharp discontinuity of the accessibility measure between close apartments.

[76] There are 34,652 such stations that were active in at least one point of time during the research period.

I calculate the total direct travel time in each mode of transportation between each two stations in one day per week,[77] once every three minutes during morning and evening rush hours in this day. I define the daily effective travel time for each rush hour as the average of those sampled travel times, weighted by the share of cellular departures in the corresponding half-hour interval out of all the departures in the morning or afternoon rush hour. Figure A.2 displays the number of journeys originating in each half hour interval. The black bars represent morning and afternoon rush hours.

Figure A.2
Average daily number of departures by time of day, 2018-2019



Note: Defined morning and evening rush hours are colored black.
Source: OD_MAT dataset, Israeli Ministry of Transportation

To get direct travel times for each transportation period from the daily times, for each period and each origin-destination pair I choose the median value of the daily calculated travel time.

Lastly, I apply Dijkstra's algorithm[78] to the direct travel time for each transportation period in the morning and afternoon rush hours to get total travel times between stations. I define total travel times between areas o and d as:

$$(A2)\quad t_{od} \equiv \min_{\mathrm{a,b}}\{t_{ab}^{morning}\} + \min_{\mathrm{a,b}}\{t_{ba}^{afternoon}\}, a \in area_o \;\&\; b \in area_d$$

[77] On Tuesdays only, according to a recommendation from the Israeli Ministry of Transportation. This is done to eliminate any unique day of the week effects. For example, a large part of the public transit system doesn't operate in weekends. Another example is increased service in some parts of the system that is targeted at getting soldiers to their base or back home on Sundays and Thursdays.
[78] Dijkstra (1959), as implemented in the R package CppRouting.

That is the sum of the minimal travel time between any station in area o and any station in area d in the morning rush hour, and the minimal travel time in the opposite direction between any (possibly other) stations in these areas in the afternoon.

B.2 Private car

There is no direct data available on travel times by private car in Israel. I apply a two-staged procedure to compute travel times: (1) Estimation of travel speed in each road segment in Israel (road segments taken from the ROADS_NETWORK dataset), and (2) find the shortest path between all transportation polygons in Israel.

<u>Estimation of travel speeds for road segments</u>

As mentioned before, there is no available universal data on travel speeds in road segments in Israel. I estimate those speeds using the travel speed of buses travelling nearby. Optimally I would have used buses travelling through the specific road segment, but buses do not ride in the entire road network, and I have data on the location and order of the stations for each bus line, but no knowledge on which road segment the bus travelled between those stations.

To estimate road speeds, I apply the following procedure:

1. Compute ride speeds for each origin-destination station pair. The outcome is a 'ray' which represents the straight line between the two stations in the pair, and the travel speed in this ray.

1.A. Compute median real total travel time (As described in appendix A) for each bus line-half hour interval.

1.B. Divide each ride to all possible edges (origin station-destination station pair) in the ride, assign travel time by the edge's proportion in planned travel time.

1.C. Compute travel speed using the value from (1.B) and the road distance between stations from the BUS_ROUTES dataset.

1.D. Filter out extreme or problematic data: kmh lower than 10 or higher than 120.

1.E. For each possible half hour interval-edge combination, assign the maximal speed.

1.F. For each edge in each transportation period and separately for morning and afternoon rush hours assign final speed value: weighted average of the speed in all half hour intervals (as described in appendix A).

2. Match public transportation 'rays' to road segments.

2.A. For each road segment: find the closest 5 public transportation 'rays'.

The distance calculated is the distance between two lines: the road segment and the public transportation ray. The two prominent distance concepts between lines are the Frechet and Hausdorff distances. I prefer the Frechet distance due its dependence on the direction one traverses on the line which is an important feature in this context.

2.B. For each road segment: assign travel speed: average of 5 closest 'rays'.

2.C. Calculate cost for each road segment using travel speed and road distance.

The main assumption required to accept this procedure is that the ratio of public transportation travel speed and private car travel speed remains fairly constant across time and space. A constant ratio which is different than 1 poses no problem for the analysis since it is equivalent to a linear transformation of the travel cost, which makes no difference to the rest of the analysis. A violation of this assumption might distort the path choices in the Dijkstra algorithm and the estimations relying on this procedure.

The result of the procedure up to this point is a GIS database of all roads in Israel with the travel time in each direction and in each road segment in the network for every transportation period and separately for morning and afternoon rush hours.

<u>Find the shortest path</u>

To find the shortest path I apply the following procedure separately for each transportation period and morning or afternoon rush hour.

3. prepare the dataset.

3.A. Transform roads network GIS object to a weighted graph: I preform this task using the weight_streetnet function from the dodgr package in R.[79]

[79] Padgham (2019)

3.B. For each transportation polygon (address) define the center as the point on the graph closest to its geometric centroid. This point will usually be an intersection of two roads or a turn within a road segment.

3.C. Simplify the graph (using cpp_simplify function from the cppRouting R).[80]

4. Apply Dijkstra's algorithm as implemented in cppRouting package in R.

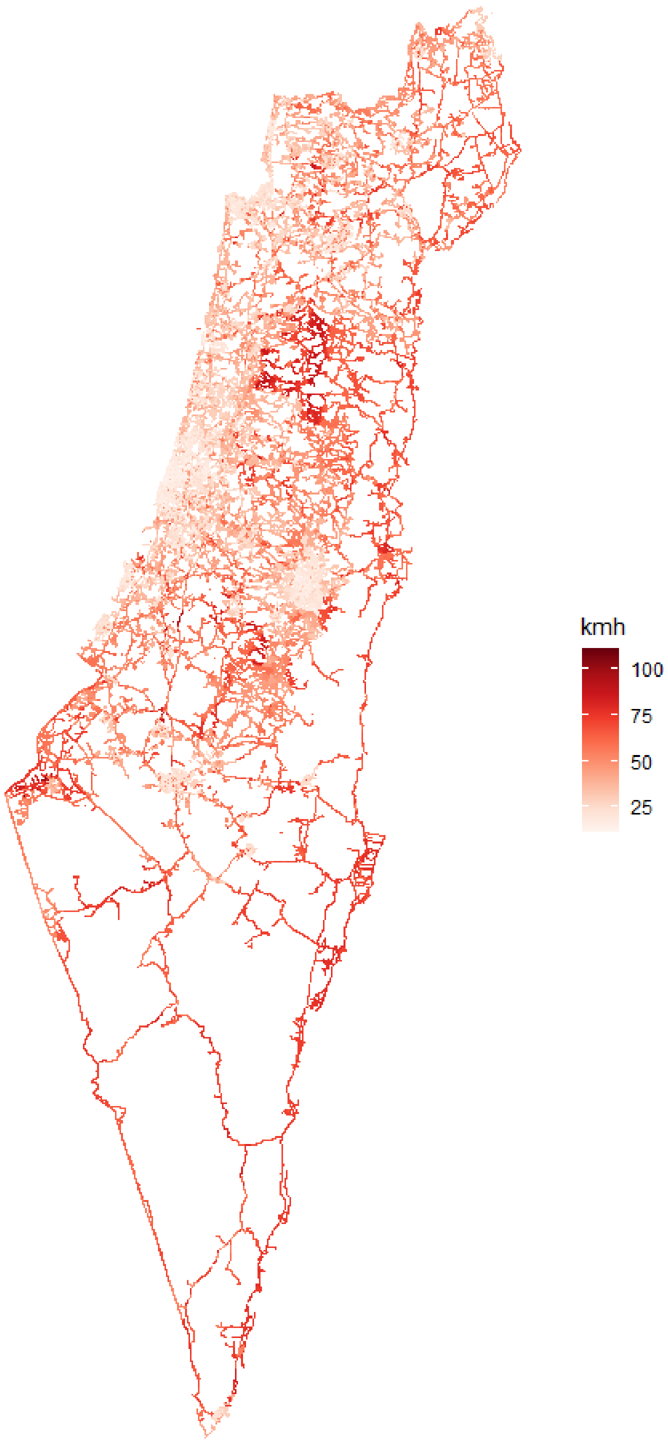
5. Compute travel time between all areas. Lastly, the computed travel time between points o and d in transportation period t is sum of the travel time between o and d in transportation period t in the morning rush hour, and the travel time between d and o in transportation period t in the afternoon rush hour.

The estimated speed for each road segment in Israel is presented in figure A.2. One can note that, as expected, the estimated speed is high in peripheral areas and highways, and rapidly declines when approaching the large metropolis.

[80] Larmet (2019)

Figure A.3

Estimated road speeds in Israel, 2019



Appendix C: The Causal Forest model: Brief Introduction

The Causal Forest model, developed by Wager & Athey (2018), is a standardized causal machine learning method to estimate heterogenous treatment effects, potentially enabling the researcher to uncover complex patterns of heterogeneity. This appendix briefly describes the model: the algorithm to grow a single causal tree, the concept of honesty in causal trees, and the ensemble of multiple trees to a causal forest, and the omnibus test for forest calibration.[81]

Causal Tree

The causal tree was developed in Athey & Imbens (2016).[82] The first step in estimation is orthogonalization using Robinson's transformation (Robinson, 1988). For each observation two statistics are estimated $y(X_i) \equiv E[Y|X = X_i]$ and $w(X_i) \equiv E[W|X = X_i]$. These vectors can in principle be estimated using any prediction method, in practice they are estimated using regression forests with Out of Bag predictions. The tree is grown on the residualized response and treatment variables: $\tilde{Y}_i \equiv Y_i - y(X_i)$, $\widetilde{W}_i \equiv W_i - w(X_i)$.

A causal tree represents recursive partitioning of the sample, where the split at each node is chosen to maximize heterogeneity in the estimated treatment effect within the tree. The estimated treatment effect within each terminal leaf is $\frac{Cov(\tilde{Y},\widetilde{W})}{Var(\widetilde{W})}$. The estimated treatment effect for observation i in tree j, $\hat{\tau}_i^j$, is the estimated treatment effect within the terminal leaf it belongs to.

Honesty

An honest causal tree is defined as a tree where for each observation i, the response, Y_i, is used either to estimate within-leaf treatment effect or to decide where to place the splits, but not both. The motivation for honesty is to avoid over-fitting the model, and bias reduction. In practice, the procedure relies on dividing the data to two subsamples. The splits of the tree are chosen using only information from one

[81] For simplicity, I describe the model as presented in Wager & Athey (2018). The implementation used in this paper, from R's grf package (Tibshirani et al, 2021), is slightly different, mostly due to computational reasons.
[82] I present a very simple version of the tree. In practice, several other factors, aimed to improve stability, affect the splits of a tree. See Athey & Imbens (2016) for a more complete description.

subsample, and the treatment effect within each leaf is estimated using information only from the second subsample.

Causal Forest

A causal forest is an ensemble of causal trees, where each individual tree in the forest is trained using only a subsample of the data and a subset of the features. The estimated treatment effect for each observation is the average of estimated treatment effects for this observation in the trees composing the forest and for which observation i was excluded from the sample: $\hat{\tau}_i \equiv \frac{\sum_{j=1}^{J} \hat{\tau}_i^j}{J}$.

The causal forest can also be estimated using an Instrumental Variable approach, as developed in Athey et al (2019). I avoid using method since it typically requires stronger instruments than the one available in my context.

Omnibus test for forest calibration

The omnibus test was developed at Chernozhukov et al (2018) and discussed specifically for causal forest models at Athey & Wager (2019).[83] This test seeks to fit the Conditional Average Treatment Effect as a linear function of out-of-bag causal forest estimates. Formally, define:

$$(A3)\ \ mean.forest.prediction_i \equiv \bar{\tau}\left(W_i - \widehat{w}^{(-i)}(X_i)\right)$$

$$(A4)\ \ differential.forest.prediction_i \equiv \left(\hat{\tau}^{(-i)}(X_i) - \bar{\tau}\right)(W_i - \hat{e}^{(-i)}(X_i)$$

Where out of bag estimates are denoted with superscript (-i), and $\bar{\tau}$ is the average of out of bag treatment effects estimates. The test is a regression of $\tilde{Y}_i$ on mean.forest.prediction and differential.forest.prediction. A coefficient of 1 for mean.forest.prediction implies that the model correctly estimates the average treatment effect, and a coefficient of 1 for differential.forest.prediction implies that the model captures the heterogeneity in the underlying signal.

[83] The presentation of the test here heavily relies on section 2.2 At Athey & Wager (2019).