I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15025

Is There a Diminishing Value of Urban Amenities as a Result of the COVID-19 Pandemic?

Aico van Vuuren

JANUARY 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15025

Is There a Diminishing Value of Urban Amenities as a Result of the COVID-19 Pandemic?

Aico van Vuuren
University of Groningen, University of Gothenburg and IZA

JANUARY 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Is There a Diminishing Value of Urban Amenities as a Result of the COVID-19 Pandemic?

We investigate whether the COVID-19 pandemic decreased the willingness to pay for urban amenities such as restaurants, cinemas and theaters. We do this by using a hedonic pricing model in combination with a time-gradient difference-in-difference approach. We use a data set that contains virtually all apartments for sale in the larger Stockholm area. We use a very detailed and flexible definition of density of urban amenities based on the exact location of these amenities and the walking distance from the apartments to these amenities. We find a decrease of 1.9 percent of apartments that we label as amenity rich.

JEL Classification: R00, R23, R30

Keywords: COVID-19, urban economics, amenities

Corresponding author:
Aico van Vuuren
Department of Economics
University of Gothenburg
Vasagatan 1
SE-405 30 Gothenburg
Sweden

E-mail: aico.van.vuuren@economics.gu.se

1 Introduction

This paper uses the hedonic pricing approach in order to investigate whether the willingness to pay for urban amenities decreased during the period of the Covid-19 pandemic. We compare prices of apartments in locations that have relatively many of these *urban* amenities with the apartment prices in locations that do not have a high density of such amenities. Our investigation limits itself to the Stockholm county. We use modern geocoding and routing machines to obtain precise information of which apartments are located in amenity dense and which apartments are located in amenity thin areas. We use semi-parametric versions of difference-in-difference methods and use machine learning techniques for these estimations.

House prices vary a lot between different neighborhoods in larger metropolitan areas (Gibbons and Machin, 2008). Many reasons have been brought up for such a variation such as the quality of schools, the presence of historical buildings, and crime rates (Machin, 2020, Black and Machin, 2010, Ahlfeldt and Holman 2018, Koster et al., 2016 and Ihlanfeldt and Mayock, 2010). An additional, less investigated, reason for such a variation is the supply of within-walking-distance *urban* amenities such as the presence of bars, restaurants, theaters and cinemas.[1] Many of these amenities were not in service or had to reduce their business during the Covid-19 worldwide pandemic. In addition, it can be expected that the use of such amenities will be limited long after the pandemic has ended either by law or by individuals' preferences to avoid being infected (Nathan and Overman, 2020). Therefore, one might expect that the value of these amenities is also diminished and houses that are very close to several *urban* amenities that were affected by the pandemic may face lower price developments than houses that are located somewhat further from these amenities.

Our method compares houses within small areas (zip codes) inside a much larger

[1] An exception is Ahfeldt and Kavetsos (2014) who investigate the impact of the building of two stadiums in London.

metropolitan area. Hence, we only ask ourselves the question whether the houses very close to the urban amenities are facing less favorable price developments than houses in the same area but located somewhat further from (all) of these amenities. This implies that we do not intend to answer the more ambitious question whether metropolitan areas will become smaller (or at least change) as a result of a new equilibrium. As discussed in Ahfeldt, et al. (2020), such changes may take a long time and hence we would potentially not be able to observe that within the short time frame that, for obvious reasons, we are able to investigate. In addition, using a hedonic approach of house prices has the advantage that house prices are forward looking and hence adjust more quickly than the residential locations of city residents as well as the rents of apartments.

Many countries used very stringent rules with respect to the opening of bars and restaurants. In that sense, Sweden followed a rather modest approach with many recommendations but with relatively little rules. Nevertheless, it has been reported that restaurants and bars did receive much fewer customers (also due to some restrictions of how many people were allowed in these places).[2] In addition, many government buildings, such as museums and theaters, were closed during the pandemic. Still, it seems reasonable to expect that the access to most *urban* amenities was better in Sweden than in many other European and North-American places. This would suggest that if anything, our estimates can be regarded as a lower bound of the results one can expect from other countries. Also, as house prices are forward looking, individuals' expectations about how they are able to use urban amenities in the future may be more important than the strict rules that hold during the (short period of the) pandemic.

The rest of this paper is organized as follows. The next section discusses the literature and some background related to the country and city of analysis. Section

[2] See for example https://www.svt.se/nyheter/lokalt/stockholm/restauranger-tanker-nytt-for-att-forsoka-overleva

3 discusses the data set used in this paper together with a description of how we determine centrality of locations. Section 4 discusses the empirical implementation and Section 5 discusses our results. Some robustness checks are presented in Section 6 and our conclusions follow in Section 7.

2 Literature and background

2.1 Literature

Our paper contributes to a long tradition that adopts the first step of Rosen's hedonic pricing model to obtain the willingness to pay for commodities using prices of composite goods (Rosen, 1974 and Bajaree and Benkhard, 2005). Houses have many commodities such as size, number of rooms as well as their neighborhood (characteristics). Hence, using certain additional assumptions such as the possibility to divide these commodities into smaller quantities, a regression of these commodities on the house prices is able to reveal the willingness to pay of the commodities.

One of the first papers to recognize that spatial amenities can be important for the determination of house price differences was Chesire and Sheppard (1995). They estimate a standard hedonic pricing model but add some location specific characteristics, such as the availability of schools and access to public transport.

Gibbons and Machin (2008) present a very general overview of how to use the first step of Rosen's hedonic pricing model in order to obtain the willingness to pay for particular (dis)amenities. They recognize three potential problems with respect to the use of a standard hedonic pricing model as (1) the supply of the amenity could be affected by the composition of the home buyers (typically called sorting), (2) there could be other unobserved amenities which are correlated with the investigated amenities (the problem of confounders), (3) measurement problems with respect to how an amenity relates to which location. All these problems are problematic for the investigation of urban amenities. As the first issue is concerned, restaurants and bars may locate themselves in areas with high levels of income where citizens can spend high amounts on luxury goods. Also problem (2) is related to our case as many places with a lot of restaurants and bars are centrally located and hence have also high levels of other amenities such as historical amenities, high concentration of jobs and good hospitals. However, these are amenities that did not change a lot or even became more important after the outbreak of Covid-19. Hence, even though it is difficult to assess the value of the amenities, we are able to investigate the impact of the pandemic on the willingness to pay for these amenities.

Most studies have tried to take these problems into account using either (1) a regression discontinuity approach that uses borders between areas that are affected or not affected by particular amenities, or (2) using (extensions of) a difference-in-difference approach to explore (unexpected) changes with respect to supply of the amenities over time. Examples of a strict use of the first method are the use of school districts in the United States (Gibbons et al., 2013), the use of different building periods of areas to measure the price impact of historical buildings on nearby houses (Koster et al, 2008) and the use of so called preservation areas to measure the capitalized value of architectural design (Ahlfeldt et al., 2017). Examples of the second method are the impact of changes in public transport (Gibbons and Machin, 2005), the impact of wind farms (Gibbons, 2015), the impact of stadiums (Ahlfeldt and Kavetsos, 2014) as well as the impact of refugee housing (Van Vuuren et al., 2019). Also the use of panel data to look at the impact of changes in crime rates can be interpreted as an adoption of this method (Ihlanfeldt and Mayock, 2010). The method that we adopt in this paper is also an adoption of this second method.

There is a growing literature on the impact of Covid-19 on house prices and rents (Davis, Ghent, and Gregory, 2021, Ling, Wang, and Zhou, 2020, Rosenthal, Strange and Urrego, 2021, and Brueckner, Kahn, and Lin, 2021). Gupta, Mittal and Peeters (2021) investigate the change of the bid-rent function. That is the change in the relative house prices (or rents) in the urban core versus those house prices (or rents) in suburban areas. They find that the bid-rent function has become flatter, but that this is more the case for rents than for house prices. The authors conclude from this that the market expects the impact of Covid-19 to be temporary rather than permanent. The authors also investigate differences between MSAs and conclude that the change in the bid-rent function was larger for those cities for which the population has more opportunities for remote work and had more stringent rules. However, the final result is less robust and this makes them conclude that the main driver for changes in the bid-rent function comes from the fact that people work from home. Hence, other explanations, such as the presence of nearby urban amenities seem less important for the change in the bid-rent function. Our work is complementary to theirs. We do not estimate a bid-rent function but instead control for location by using detailed levels of zip code. Hence, houses with many urban amenities have about the same distance to the central business district(s) as houses that we compare them with. Note that our results could be affected in the case that the locations of urban amenities are almost perfectly correlated with those of jobs and when workers prefer to live very close to those jobs. Another problem arises when urban amenities are located at the same places as commuter train stations. We think that the first argument is not very plausible as commuting costs are generally considered only problematic when longer than 15 minutes (Redmond and Mokhtarian , 2001). We check for the second argument in a robustness check.

2.2 Background

Sweden's approach to the spread of the pandemic has been different from many other countries. Instead of applying lockdowns, the Swedish government has not actively closed down businesses but relied almost entirely on recommendations and statements concerning societal responsibility. Among these were the recommendation to keep a distance of at least two meters, to stay at home when having symptoms and working from home if possible. Nevertheless, the Swedish government did close many governmental buildings such as universities, museums, theaters and sport stadiums. The main reason for not closing businesses was that it was regarded against the Swedish constitution. In addition, the Swedish approach was to rely on the recommendations of experts from governmental institutions (in particular the Public Health Agency of Sweden or Folkshälsomyndigheten) instead of those from political leaders which has been the procedure in many other European countries. The responsible experts at the Public Health Agency regarded the use of lockdowns as ineffective for the two main goals set by the Public Health Agency, *i.e.* protecting the vulnerable (the elderly) in society from being infected and to flatten the curve in order not to over-demand the healthcare system. Another reason for not applying lockdowns was that it was regarded as impossible to keep these effective for the whole period of the pandemic.

Among the few regulations that were in place during the pandemic were a prohibition of gatherings of more than 50 individuals, the limitation of serving food and drinks to tables in bars and restaurants. Moreover, tables had to be placed further apart in order to guarantee at least a one meter distance between their clients. Local authorities had the opportunity to close restaurants and bars that did not comply to these rules although we have no information whether this actually happened. Further restrictions were imposed later during the year such as a maximum gathering of 8 individuals and restrictions of opening hours for places that serve alcoholic drinks.

The Swedish approach is generally regarded as not very successful and there has been a lot of national and international critique to this approach (Claeson and Hanson, 2021). Overall, infection and death rates were not lower in Sweden than in most other European countries while they were much higher than the other Scandinavian countries.

Our research limits itself to the Stockholm county which is by far the largest metropolitan area in Sweden. By the end of 2019, it inhabited roughly 2.4 million citizens. Stockholm is also the economic center of the country with far above income and education levels. It has 26 municipalities of which the municipality of Stockholm is the largest (with around 1 million inhabitants). The population density is high for Scandinavian standards but it is still much lower than many other European metropolitan areas. For example, Île-de-France is about twice the size in surface of Stockholm county but it inhabits about 5 times as many citizens. Infection and death rates were much higher in Stockholm than in any other area in Sweden.

3 Data

We collected data for 271,817 transactions from the site Booli.se using their API to retrieve it (see also Van Vuuren et al., 2019 who use the same data source). These are all transactions within the Stockholm region. Booli is an independent search engine for private properties. The site collects publicly available data from most real-estate agencies' websites.[3] Apart from the houses that are currently for sale, Booli also presents houses that have been sold since 2012 and this is the information that we use for our analysis.

After deletion of transactions with missing observations for some of the variables used in our regression analysis, we end up with a sample of 208,586 transactions. We also use only those observations that are labeled as apartment. The reason for this choice is twofold. Central houses are almost always apartments and hence using only apartments for the decentral houses makes the two groups more comparable. Another reason is that, due to privacy reasons, houses that are *owned* are usually not presented in the Booli API. Instead, the Booli API has an almost full coverage of houses for which the buyer does not officially own the house but instead buys a "right to live" in the house. This implies that the buyer needs to pay a rent in

[3] One does not need to actively advertise on Booli for the property to be available on the website; it only needs to be available on the real-estate agencies' websites.

addition to the original payment for the house. These houses are almost exclusively apartments and there are very few apartments that do not fall in this category.[4] This leaves us with a total of 191,174 transactions and after deletion of those houses without a rent or floor number, we end up with 163,054 transactions.

For our main analysis, we only use those transactions in the period from one year before the start of the pandemic up till 26 weeks after the start of the pandemic. This leaves us with 43,118 transactions.

We use the Overpass API in order to determine the amenities within the Stockholm region. We use the following amenities: restaurants, bars, pubs, museums, cinemas, theaters, shopping malls and fitness centers.[5] For example, we were able to identify 2,349 restaurants, 102 museums and 36 cinemas within the Stockholm county. Maps of the location of these amenities are presented in Figure 1. We use the open source routing machine (OSRM) to calculate the walking distances between the apartments and the amenities. Using walking distances instead of using physical (Euclidean) distances has the advantage of accounting for the fact that there are many rivers, islands, forests and hills in Stockholm and even though two sites might be very close to each other based on geographic location, it may take quite some time to move between these locations.

Even though we use many different definitions of what constitutes a central location, we start with a rough definition in order to obtain some insights of the data. That is, we define an apartment to be centrally located if it has at least five restaurants within a ten minutes walking distance, at least one museum within half an hour walking distance and at least one cinema within half an hour walking distance. Decentral locations are those locations that have less than five restaurants within half an hour walking distance and no museum or cinema within a forty-five minutes walking distance. The distributions of our transactions with either central

[4]Before 2009, it was not allowed to own an apartment. Since then, only new dwellings can be sold as ownership apartments.

[5]See https://wiki.openstreetmap.org/wiki/Overpass_API

or decentral locations are in Figure 2. Of the 43,118 transactions as determined above, 18,872 are central, 14,205 are non-central and 10,041 are neither central nor decentral. In our main analysis, we only use those observations that are either central or decentral and hence this leaves us with 33,373 observations for our regression analysis.

Table 1 lists the descriptive statistics of our data set. Apartments sell for a little over 3.8 million SEK (or around 425 thousand US dollars) for this period and the average size of an apartment is a bit more than 600 squared feet. The most important observation that can be made from Table 1 is the fact that prices of apartments in central locations are on average 75 percent more expensive than houses located in decentral areas. Houses that are neither central or decentral have prices in between these two extremes. Note that central and decentral houses are also different in some other respects. For example, the rent in central locations is lower, the distance to the sea is shorter and the construction period is earlier. The first two characteristics definitely are in favor of central locations, while also the last characteristic is likely to be in favor as 1970s apartments are not in a high demand, especially so in Sweden. Nevertheless, these characteristics cannot fully explain the price difference between central and decentral locations. Hence, the most likely explanation is that the houses are differently priced from each other simply because they are located in different areas with different types of amenities (including the ones discussed above but also schools, crime levels etc.). Unfortunately, these price differences may be somewhat problematic for our empirical implementation. We come back to this issue in Section 4.

4 Empirical implementation

We use two different strategies to investigate the impact of the Covid-19 pandemic on the hedonic value of urban amenities. The first model is represented as (see also

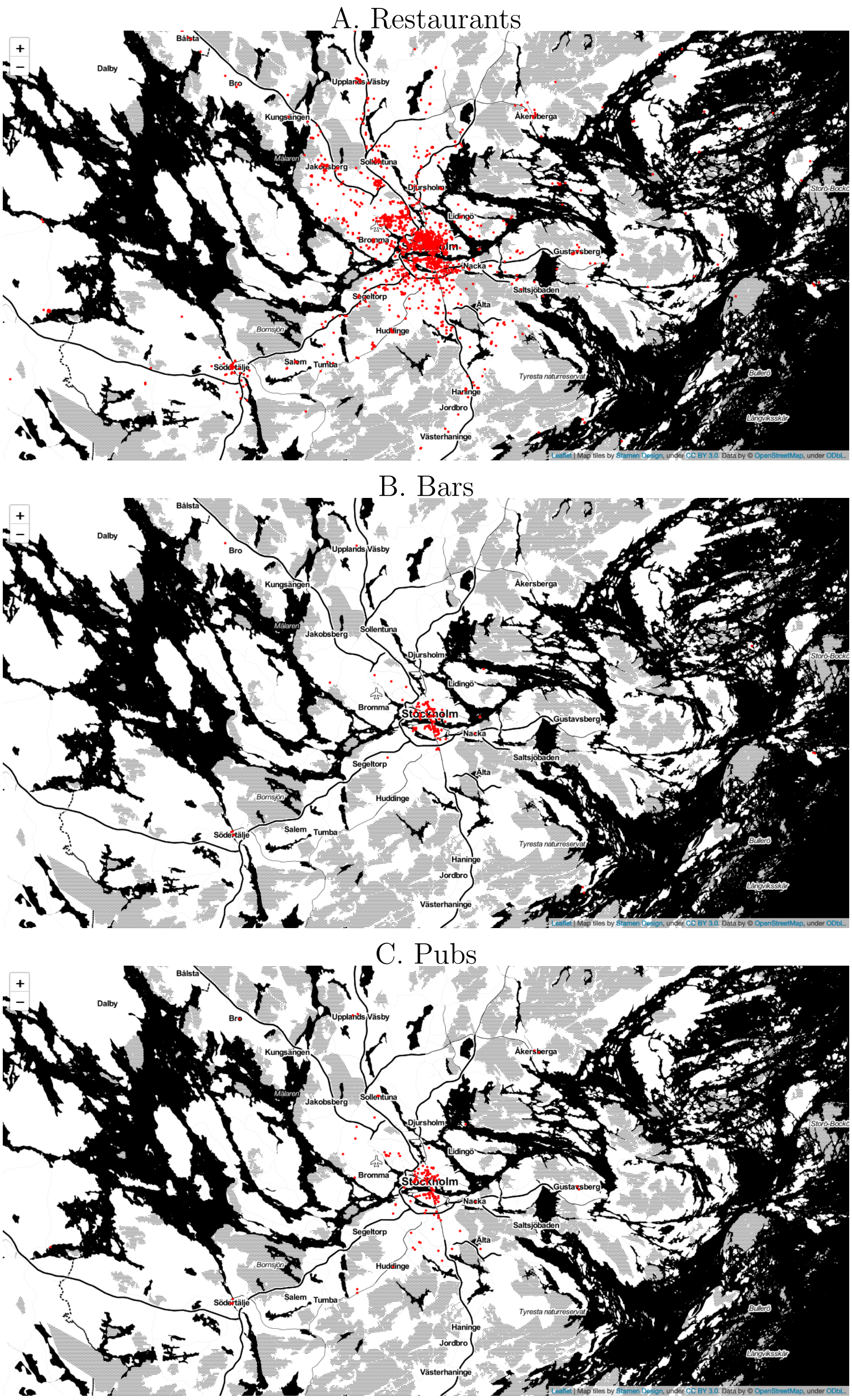
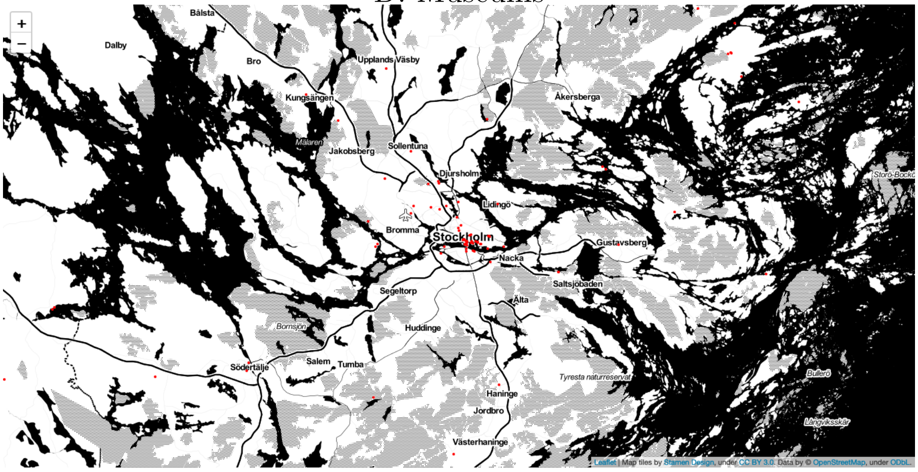
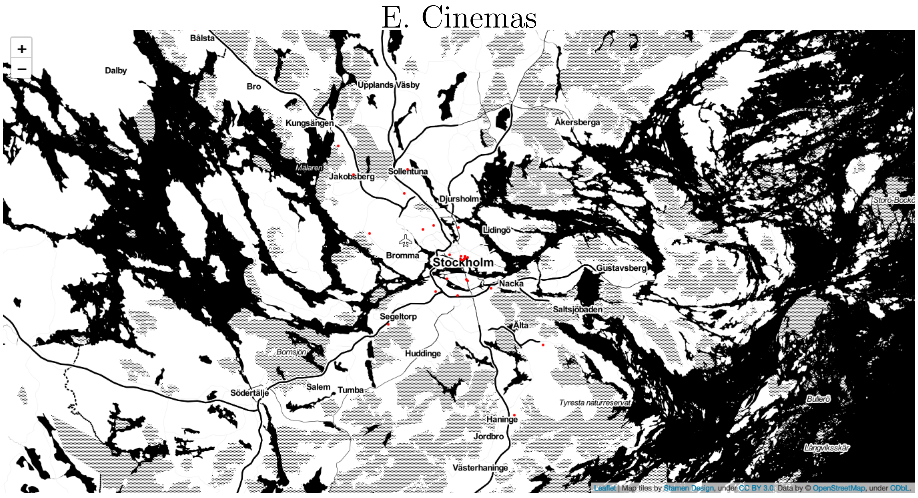


Figure 1: Location of amenities.

D. Museums



E. Cinemas



F. Theaters

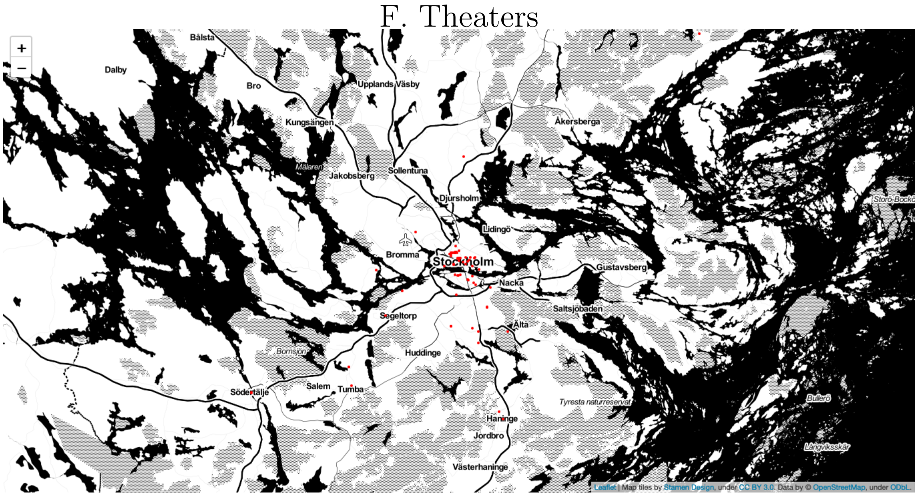
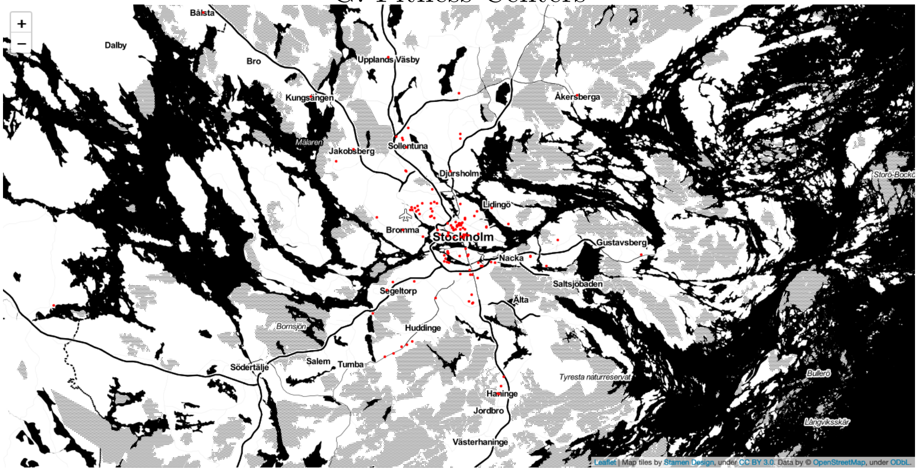


Figure 1: Location of amenities (*continued*).

G. Fitness Centers



H. Malls

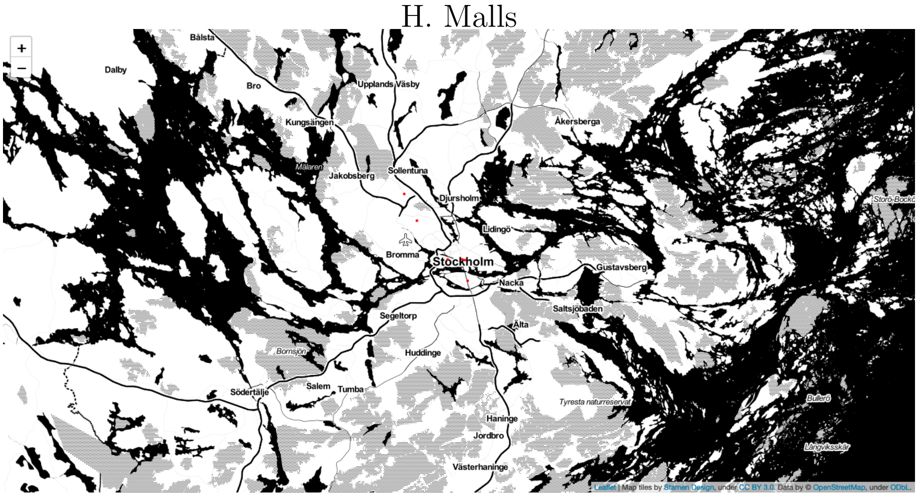


Figure 1: Location of amenities (*continued*).

	All			Central			Decentral			None		
	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Min.
Sold price (in 1,000 SEK)	3832	40000	450	4994	40000	630	2666	13500	450	3296	17650	795
Living Area	62	244	20	61	244	20	65	211	21	62	214	21
Rent	3407	15000	125	3016	12945	132	3831	12308	125	3544	15000	440
Number of rooms	2.41	9	1	2.32	7	1	2.51	8	1	2.45	9	1
Floor number	2.81	10	0	2.99	10	0	2.49	10	0	2.91	10	0
Distance to sea	3322	13507	1	2158	11186	49	4483	13507	1	3868	11594	2
Year of construction	1958	2020	1880	1943	2020	1880	1970	2020	1880	1971	2020	1880
Index of decentralization	0.48	0.99	0.12	0.23	0.91	0.12	0.8	0.99	0.23	0.47	0.95	0.21
Number of observations	43118			18872			14205			10041		

Table 1: Descriptive statistics of the data set.

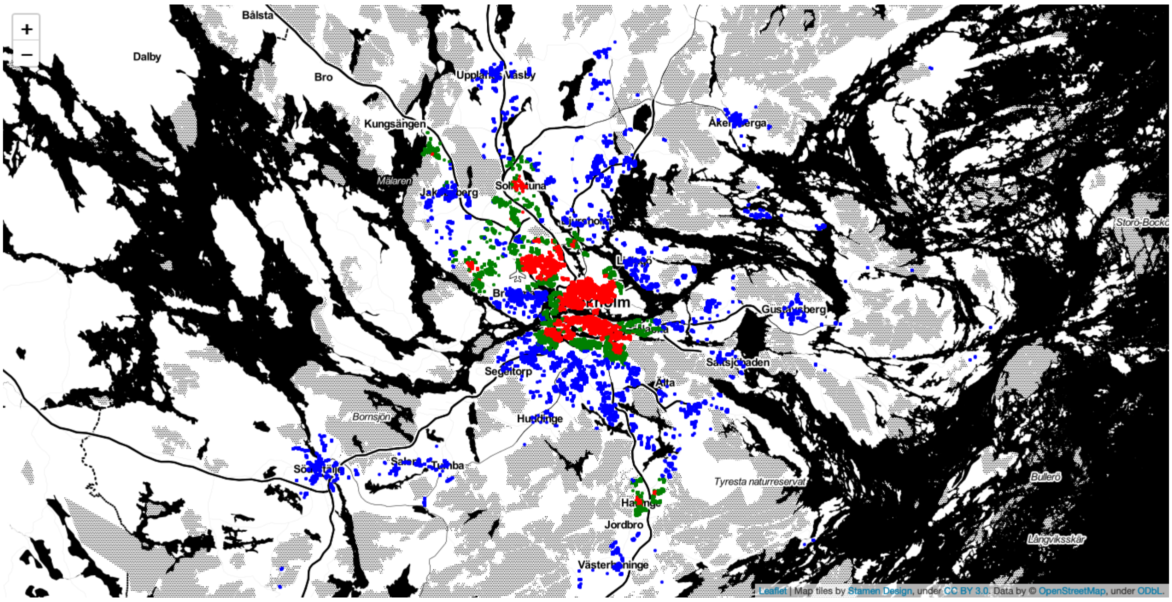


Figure 2: Location of apartments transacted in our data set for the different zones. Apartments in central location are in red. Apartments in decentral locations are in blue. Apartments in neither central or decentral locations in green.

Ahlfeldt et al. (2017) and Van Vuuren et al. (2019)),

$$\log P_i = D_i^{\text{central}} \Phi_1(T_i) + D_i^{\text{central}} D_{T_i}^{\text{post}} \Phi_2(T_i) + \beta' X_i + \alpha_{J_i} + \lambda_{T_i} + U_i, \tag{4.1}$$

where $\log P_i$ is the log price per square feet of transaction i. D_i^{central} is a dummy for the fact that transaction i concerns a house in a central location, while $D_{T_i}^{\text{post}}$ is a dummy that indicates that the transaction was after the start of the pandemic. The vector X_i is a set of apartment characteristics related to the apartment of transaction i. We use a quartic of the size of the apartment in square feet, the number of rooms (10 dummy variables with the highest indicating 10 or more rooms), the floor on which the apartment is situated (10 dummy variables with the highest indicating 10$^{\text{th}}$ floor or higher), construction period (14 dummy variables for decades of construction since 1890), rent as a second-order polynomial and distance to the ocean as a second order polynomial. The coefficients α_{J_i} form a full set of dummy variables of zip codes (3 digits, 85 different zip codes). The coefficients of λ_{T_i} is a full set of week dummies for all weeks considered (77 dummy variables).

The most important part of our analysis is related to the functions $\Phi_1(T_i)$ and $\Phi_2(T_i)$. The first function is a polynomial that indicates the difference in trend between central and decentral areas in Stockholm. It can be interpreted as the difference in trend for our whole period in the case that the pandemic would not have taken place. The function $\Phi_2(T_i)$ is the additional difference in trend in the period after the start of the pandemic. Under certain assumptions, to be discussed below, this function can be interpreted as the impact of the pandemic in terms of relative house prices in central locations.

As is standard in the evaluation literature, we are (ideally) interested in the estimation of the Average Treatment Effect on the Treated at time t (ATT_t)

$$\mathbb{E}(\log P(1)|D^{\text{central}} = 1, D_T^{\text{post}} = 1, T = t) - \mathbb{E}(\log P(0)|D^{\text{central}} = 1, D_T^{\text{post}} = 1, T = t),$$

where, for this case, state 1 is the state in which the Covid-19 pandemic already took place in a central area. State 0 is somewhat more difficult to describe. It is either the state before Covid-19 or the state in which the price development in a central area after Covid-19 is not differently affected than any other area. That is ATT_t is the impact on the central areas in the case they would not be affected. This effect is somewhat different from the standard evaluation literature as we cannot rule out that prices in non-central areas are not affected by Covid-19. It implies that we intend to measure whether the presence of urban amenities in an area affects the house prices in such an area differently from other areas.

Our model is similar to a regression discontinuity design in the case that there is a time running variable (Ahlfeldt et al., 2017). Ahfeldt et al. (2017) call the model a time-gradient difference-in-difference method. The assumptions made by this model depends on the exact specification of Φ_1. For example, it reduces to the standard difference-in-difference method in the case of a constant specification of $\Phi_1(T_i)$ and $\Phi_2(T_i)$ (*i.e.* $\Phi_j(T_i) = \phi_j; j = 1, 2$). Such a model assumes that the impact of Covid-19 is immediate and does not change over time and that there is no difference in trends before the pandemic between central and decentral locations. Both of these assumptions are not very plausible as central and decentral locations could have different trends and the impact could be either increasing or decreasing over the period of the pandemic (based on changing expectations of the apartment buyers). A linear specification assumes that the difference in trends, rather than the differences in outcomes, should be constant over time between the control and treatment group. That is, we need the following assumption to hold for $t_0 \leq 0 < t_1$

$$\begin{aligned}
&\mathbb{E}(\log P(0)|T = t, S = 1) - \mathbb{E}(\log P(0)|T = t_0, S = 1) = \\
&= \mathbb{E}(\log P(0)|T = t_1, S = 0) - \mathbb{E}(\log P(0)|T = t_0, S = 0) \\
&+ \{\mathbb{E}(\log P(0)|T = 0, S = 1) - \mathbb{E}(\log P(0)|T = -1, S = 1) \\
&- [\mathbb{E}(\log P(0)|T = 0, S = 0) - \mathbb{E}(\log P(0)|T = -1, S = 0)]\} \times (t_1 - t_0).
\end{aligned} \tag{4.2}$$

The first line of the left-hand side is the famous parallel trend assumption for standard difference-in-difference estimation. The second part of the left-hand side is additional. Of course, the question whether an assumption about the difference in trends is plausible, even in the case that an assumption about the differences in outcomes is not plausible, depends on the particular application. A quadratic specification further relaxes the assumptions about the trends.

The variable U_i is an error term that captures all unobserved aspects related to the transaction price. In order to obtain the causal interpretation of $\Phi_2(T_i)$ as presented above, we have to assume that U_i is uncorrelated with all the right-hand side variables presented in (4.1). This implies first of all that the unobserved characteristics that affect the house prices should not be affected by the timing of the pandemic other than the fact that the houses are either central or decentral. That is, it should not be the case that, for example, very badly maintained central apartments entered the market just after the start of the pandemic. Also, it should not be the case that the difference in trends changed just around the start of the pandemic (other than caused by the pandemic). Given the fact that we use a rather short period before and after and the fact that we correct for a polynomial, we expect this not to be a problematic assumption.

An additional assumption that we have to make is that the β coefficients of (4.1) are identical for central and decentral locations. This is necessary in order to predict the house price in, for example, a central location would it have been in a decentral location. Applied econometricians typically try to fulfil this requirement by using comparable houses in central and decentral locations. For our analysis, we do this by using only apartments and transactions in a short period of time as well as in a relatively small region. Still, we find that the prices between centrally located apartments are quite different from those of decentrally located apartments. Note that this is not necessarily problematic. That is, house prices per squared meter could be only higher for centrally located apartments, while the relative impact of the other characteristics may be identical. However, such an assumption could be still somewhat restrictive. For example, the impact of construction period could be different for centrally located houses. One way to solve for this problem is to only use those apartments that are (almost) identical based on observed characteristics. This method, generally referred to as matching, is very common in the literature. However, as we stated in the data section, it is unlikely that this will make the prices exactly identical. Therefore, we prefer to restrict the area of investigation. We discuss this in our robustness checks, see Section 6.1.

Note that the identification of the parameters in $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ is based on the assumption that there are zip codes for which there are both centrally and decentrally located apartments. Fortunately, our definition of a central location is very locally determined and hence it is possible that some zip codes have such a variation in central and decentral locations. In particular, in the case that we only use central and decentral locations in our baseline as defined above, then there are

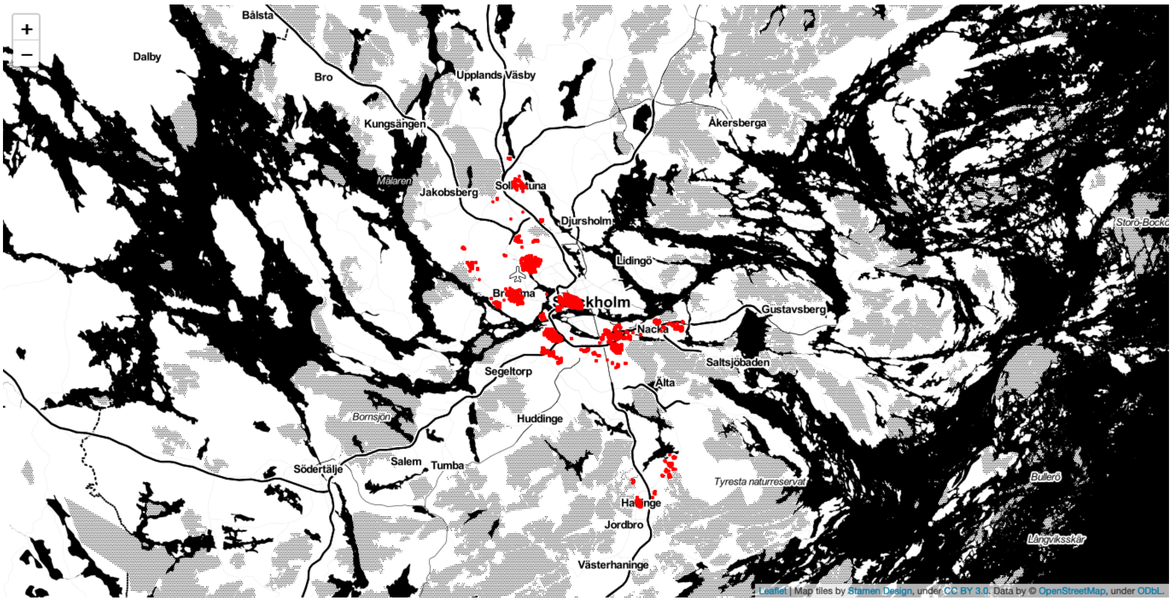


Figure 3: Location of apartments transacted in our data set in zip codes that have both centrally and decentrally located apartments.

7670 apartment sales that are in zip codes with both types of locations. 2553 of these transactions took place after March, 12$^{\text{th}}$ 2020. The location of these apartments is presented in Figure 3.

As an alternative, we can also use an index for (de)centrality instead of a simple dummy variable. We define this index by Z_i and define the construction of such an index below. Then (4.1) can be rewritten as

$$\log P_i = \Phi_1(Z_i, T_i) + D_{T_i}^{\text{post}} \Phi_2(Z_i, T_i) + \beta' X_i + \alpha_{J_i} + \lambda_{T_i} + U_i, \tag{4.3}$$

where Φ_1 and Φ_2 are now polynomials of both Z_i and T_i. The most simple specification is to use a linear term for Z_i. However, we will also introduce higher-order terms of Z_i and D_i as well as interaction terms. Note that the interpretation of our results is not as easy as in specification (4.1) as the impact of the pandemic on the value of urban amenities depends on the specification of the polynomials as well as on the construction of Z_i. However, we can use predictions with respect to fixed values of Z_i, for example, at respective quantiles.

An important benefit of using an index is that it weakens the identification restrictions. Remember from above that the identification restriction for the polynomial Φ_1 and Φ_2 was that there are 3-digit zip code areas with both the most centrally located apartments and the least centrally located apartments. This requires a large variation in the access to amenities within these relatively small areas. Using an index implies that we only need that it has variation within an area. Since our index has continuous variation, this identification restriction is surely satisfied.

The construction of Z_i is as follows: for all the different amenities, we calculate the relative distance as the rank in the distribution of distances.[6] This gives eight different values between 0 and 1 for every apartment. Then, we add these values up to one index and use again the ranks of this index for Z_i. The reason for adding up the ranks is that we do not want some of the amenities to become more important than others. For example, given the low number of shopping malls, the time it takes to walk to the nearest shopping mall may be very high for any location far from such a mall. Hence, simply adding up all walking times would give a bias towards the walking distance of the shopping malls. Of course, our index is somewhat ad-hoc and hard to interpret. Therefore, we also look at the relative times separately for all different amenities.

4.1 Semiparametric estimation using machine-learning methods

Although specification (4.1) is one of the most flexible specifications estimated these days in applied econometrics, it still has many restrictions. For example, assumption (4.2) needs to be true for the total population and is not allowed to be different for different subpopulations such as for differently sized houses. This is especially

[6] We still use differences in terms of amenities. For example, for restaurants we look at the distance to the fifth closest restaurant, while for museums we look at the distance to the closest museum.

problematic in the case that the control and treatment areas are different in terms of observed characteristics. Note that (4.1) makes even more restrictive assumptions such as that treatment and time have the same impact for all houses irrespective of their characteristics. Therefore, instead of restricting our analysis to (4.1), we also use the following semi-parametric framework for our analysis

$$\begin{aligned} Y_i = \psi_0(X_i) + \psi_1(X_i)S_i + \psi_2(X_i)S_iT_i + \lambda_{T_i}(X_i) + \delta_0 S_i \mathbf{1}(T_i > 0) \\ + \delta_1 S_i \mathbf{1}(T_i > 0) + U_i, \end{aligned} \tag{4.4}$$

Note that this model is linear in time and therefore only an extension of the linear model. Higher order terms could be included at the expense of complicating the analysis substantially. Nevertheless, this specification is more general as it no longer assumes the coefficients to be constant with X_i. As in Nie et al. (2021), we can write this specification in its *Robinson* transformation (Robinson, 1988)

$$\log P_i = \mu(X_i) + (S_i - s(X_i))\zeta(X_i) + A(X_i, S_i, T_i) \sum_{t=\underline{t}, t\neq 0}^{t=\bar{t}} \nu_t(X_i) +$$

$$B(X_i, S_i, T_i) \sum_{t=\underline{t}, t\neq 0}^{t=\bar{t}} \eta_t(X_i) + C(X_i, S_i, T_i)\delta_0 + D(X_i, S_i, T_i)\delta_1 + U_i,$$

where

$$s(x) = \mathbb{P}(S_i = s|X_i = x),$$

$$p_t(x) = \mathbb{P}(T_i = t|X_i = x),$$

$$\mu(X_i) = \mathbb{E}(Y_i|X_i = x),$$

$$\zeta(X_i) = \mathbb{E}(Y_i|X_i = x, S_i = 1) - \mathbb{E}(Y_i|X_i = x, S_i = 0),$$

$$\nu_t(X_i) = \mathbb{E}(Y_i|X_{it} = x, T_i = t) - \mathbb{E}(Y_i|X_i = x, T_i = 0),$$

and

$$\eta_t(X_i) = \mathbb{E}(Y_i|X_i = x, T_i = t, S_i = 1) - \mathbb{E}(Y_i|X_i = x, T_i = 0, S_i = 1).$$

The functions $A(\cdot), B(\cdot), C(\cdot)$ and $D(\cdot)$ are functions that do not only depend directly on the data but also indirectly through the probabilities $s(\cdot)$ and $p_t(\cdot)$. Finally $\underline{t}$ and $\overline{t}$ denote the start and end of the sample period. In contrast to Nie et al. (2021), but in line with, for example Abadie (2005), we assume treatment S_i to be independent over time *i.e.* $S_i|X_i \perp\!\!\!\perp T_i|X_i$. For our case this implies that conditional on X we should not see a difference in the number of central houses being sold over time. Even though Nie, et al. (2021) claim that this standard assumption in the evaluation literature may not always hold in cross-section data, we think that it is plausible for our case.

We can now estimate the parameters δ_0 and δ_1 using the following algorithm.

Algorithm 1

1. *Split the data into K equal folds $\mathcal{I}_1, \ldots, \mathcal{I}_K$.*
2. *For every fold $k \in \{1, \ldots K\}$ estimate $\mu^k(x)$, $p_t^k(x)$, $s^k(x)$ using any kind of prediction method not using the observations in the k-th fold.*
3. *Estimate ζ^k, ν_t^k and η_t^k by using non- or semiparametric estimation methods of heterogeneous treatment effects (*i.e. *treatment S_i, T_i and T_i for $S_i = 1$) and by not using the observations in the k-th fold.*
4. *Use the estimators in (2) and (3) to predict $\widehat{\mu}^k(X_i)$, $\widehat{p}_t^k(X_i)$, $\widehat{s}^k(X_i), \widehat{\zeta}^k(X_i)$, $\widehat{\nu}_t^k(X_i)$ and $\widehat{\eta}_t^k(X_i)$ for the observation i in fold k.*
5. *Construct the estimators of $\widehat{A}^k(X_i, S_i, T_i)$, $\widehat{B}^k(X_i, S_i, T_i)$, $\widehat{C}^k(X_i, S_i, T_i)$ and $\widehat{D}^k(X_i, S_i, T_i)$ for the k-th fold.*

6. *Define*

$$\widehat{\Delta \log P}_i = \log P_i - \widehat{\mu}^k(X_i) - (S_i - \widehat{s}^k(X_i))\widehat{\zeta}^k(X_i) - \widehat{A}^k(X_i, S_i, T_i) \sum_{t=\underline{t}, t \neq 0}^{t=\bar{t}} \nu_t(X)$$
$$- \widehat{B}^k(X_i, S_i, T_i) \sum_{t=\underline{t}, t \neq 0}^{t=\bar{t}} \eta_t(X)$$

and

$$\widehat{\Delta}(X_i, S_i, T_i) = \left(\widehat{C}(X_i, S_i, T_i), \widehat{D}(X_i, S_i, T_i)\right)$$

for the observations in the k-th fold.

7. *Run a regression of* $\widehat{\Delta \log P}_i$ *on* $\widehat{\Delta}(X_i, S_i, T_i)$ *to obtain the estimators* $\widehat{\delta}_0^k$ *and* $\widehat{\delta}_1^k$

8. *Do steps 2-7 for all K folds.*

9. *Estimate* δ_0 *and* δ_1 *as the average over* $\widehat{\delta}_0^k$ *and* $\widehat{\delta}_1^k$*, respectively.*

As prediction methods in 2, we can use any kind of machine learning method. We have chosen here to use Ridge regression as it is fast and generally has a high performance. Also, as the method is global, it usually has a relatively low variance. The tuning of the meta parameters of these methods is done independently of the algorithm, implying that they are set identical for all different folds.

Note that this method can be regarded as a semiparametric version of the famous regression discontinuity approach. Although we realize that our result is in line with Nie et al. (2021) who investigate difference-in-difference, this is an additional contribution of this paper.

4.2 Synthetic control groups

One way to make sure that the standard assumption of parallel trends is satisfied is to use a synthetic control group. That is, use a weighted average of observations in the control group in such a way that we obtain by definition a group that has exactly the same pre-trend as the treatment group. In panel data, the weights are based on individuals. This is not possible in a cross section, but we can still base the weights on observed characteristics. We use zip codes in our analysis. In words, this implies that the observations from the treatment group are not compared directly to those of the control group, but instead we use a weighed average of the control group. The weighting is done in such a way that at least the weighted control group has a pre-trend which is identical to the treatment group.

5 Results

5.1 Main results

The baseline results of our analysis are presented in Table 2. We estimate three specifications for our analysis. The first is a specification in which Φ_1 and Φ_2 are both constants. The results are relatively easy to interpret for this case by looking at the table. We find that apartments in central areas are around 8.4 percent more expensive even after controlling for 3 digit zip codes and all other characteristics used in our analysis. The result is highly significant but, as we discussed earlier in the paper, this cannot be interpreted as a direct evidence that amenities close to an apartment make those apartments more expensive. We find for the period after the start of the pandemic that the apartment prices are even higher than the period before. Note that this is a comparison to the other houses as we already take an overall trend into account. The coefficient is small, only 0.56 percent, and it is only significant on a 10 percent significance level. Nevertheless, the result implies that, if anything, the pandemic did increase apartment prices in central, urban amenities dense locations instead of reducing them. However, this result is based on the assumption that there was no difference in the trend of apartment prices in the year before the start of the pandemic.

The second column of Table 2 shows that this assumption is not correct. Apartment prices have been increasing 0.08 percent more per week in the year before the pandemic. This may not be a large surprise: Sweden has seen high levels of urbanization in the past decades and especially Stockholm has increased in terms of the number of inhabitants (Bjerke and Mellander, 2017). Such urbanization, especially of higher educated young individuals, typically increases the apartment prices of centrally located areas as young professionals have a high preference to live close to where the jobs are (Van Vuuren, 2018). Correcting for this trend, we find a quite different result for the period after the start of the pandemic. That is, we find a small initial decrease of 0.63 percent. This result is not significant. In addition, we find a significant decrease of 0.13 percent per week of apartment prices for every week after the start of the pandemic. This result is illustrated in the left panel of Figure 4.

The results in the third column of Table 2 are more difficult to interpret due to the use of a second order polynomial. Therefore, we focus here only on the trend line as reported in the right panel of Figure 4. This result is in line with the linear specification presented in the left panel, but the impact is a little higher. Nevertheless, the standard errors are a lot higher and hence the impact is only different from zero between 6 and 19 weeks after the start of the pandemic.

Table 3 lists the results for the case in which we do not only use decentral locations as control group observations but also those observations that were defined earlier as neither central nor decentral. Figure 5 is the corresponding illustration of these results. Using these observations has the benefit of obtaining more observations in zip code areas that have both observations in the control and treatment group. The total number of observations in such zip codes increases from 7,670 to 24,334. Again, about one third of these observations come from the period after the start of the pandemic. However, the caveat of using these observations is that

	Constant	Linear	Quadratic
Whole period			
Constant	0.0837	0.103	0.1063
	(0.0058)	(0.0065)	(0.0073)
Weeks (× 100)		0.077	0.1181
		(0.012)	(0.0004)
Weeks^2 (× 100)			0.0008
			(0.0009)
Period after			
Constant	0.0056	-0.0063	-0.0042
	(0.0032)	(0.0065)	(0.0102)
Weeks (× 100)		-0.1268	-0.2751
		(0.0355)	(0.1549)
Weeks^2 (× 100)			0.003
			(0.0052)
R^2-within	0.5405	0.5412	0.5412
Number of observations	33373	33373	33373

Table 2: Results of the main analysis. Heterokedasticity robust standard errors between parentheses. R^2-within is the relative explained variation within zip code areas.

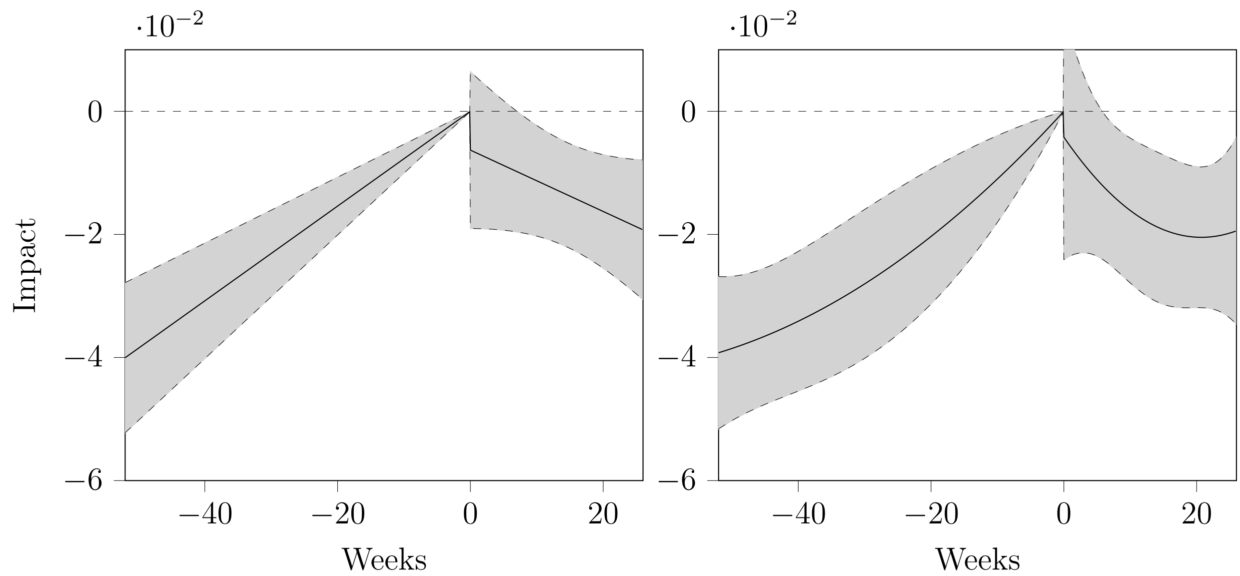


Figure 4: Results of main analysis

	Constant	Linear	Quadratic
Whole period			
Constant	0.0628	0.0797	0.0832
	(0.0027)	(0.0036)	(0.0046)
Weeks (× 100)		0.0682	0.1123
		(0.0101)	(0.0004)
Weeks^2 (× 100)			0.0009
			(0.0007)
Period after			
Constant	0.0059	-0.0062	-0.0069
	(0.0027)	(0.0055)	(0.0087)
Weeks (× 100)		-0.1012	-0.2009
		(0.0304)	(0.1317)
Weeks^2 (× 100)			0.0011
			(0.0044)
R^2-within	0.4437	0.5649	0.565
Number of observations	43445	43445	43445

Table 3: Results of the main analysis using all non-central locations in the control group. Heterokedasticity robust standard errors between parentheses. R^2-within is the relative explained variation within zip code areas.

some apartments may be situated in locations that are not very different from the centrally located apartments. We conclude that our results are not affected by a large extent to the use of these different treatment groups. Note that it is somewhat surprising that the standard errors are not decreasing by a large extent especially those of the coefficient after the start of the pandemic.

Table 4 lists the results using our overall index as introduced in Section 4. Note that for these regressions, a higher index means that the apartment is less centrally located in terms of access to urban amenities. Again, the estimates of the period after are the ones which are most interesting although they are harder to interpret. Still, the index, as the dummy variable, can only take values between zero and one. Hence, the impact can be seen as the extreme case of going from the most centrally located apartment (a zero) to the most decentrally located apartment (a one). Comparing these results with the results presented in Tables 2 and 3, we can draw the conclusion that there is no qualitative difference between them.

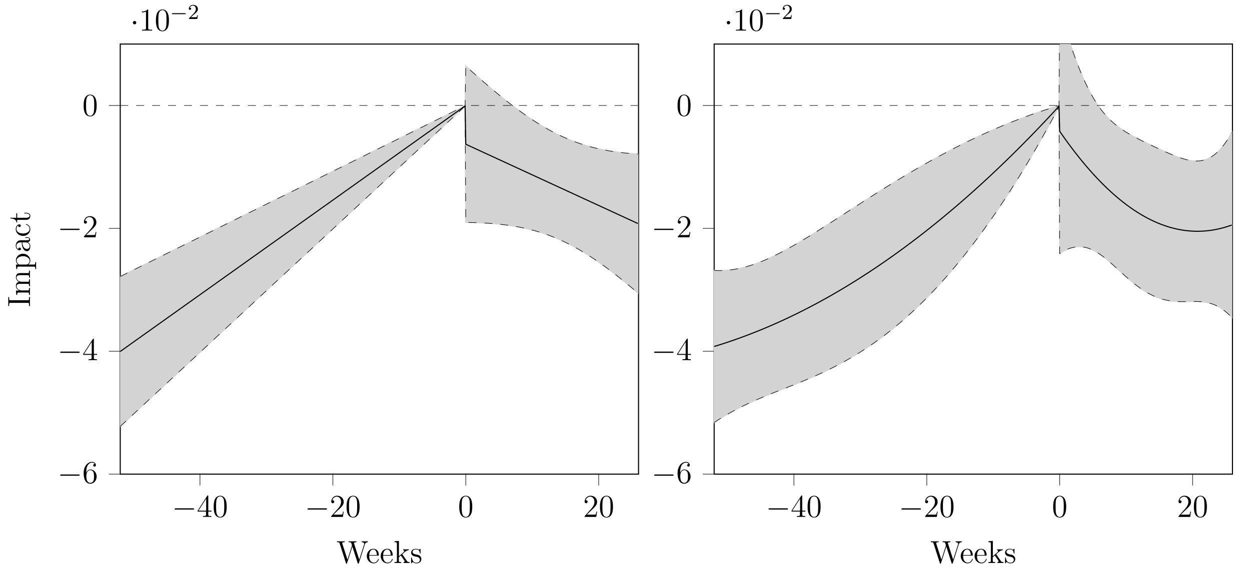


Figure 5: Results of main analysis using all observations for control group.

Nevertheless, we find that the constant terms for the case in Table 4 are much larger in absolute values, while the standard errors are, at least in comparison to Table 2, comparable. On the other hand, the values of the linear term are, if included, much smaller. Still, the coefficient for the linear term is significantly different from zero in the linear case.

We present the results for the different amenities using the index in Table 5. We have only looked here at the specification with a linear term. We do not find huge qualitative differences between the different amenities. However, some amenities have large constants and relatively low slope coefficients (such as bars) while others have lower constants but larger slope coefficients (such as shopping malls).

5.2 Results of the semiparametric estimation technique

Figure 6 presents the results of our semiparametric estimation technique. In order to omit a very complex model, we only use 8 different time periods for this analysis. These are: (1) 52-36 weeks before start, (2) 36-27 weeks before start, (3) 27-18 weeks before start, (4), 18-9 weeks before start, (5) 9-0 weeks before start, (6) 0-9 weeks

	Constant	Linear	Quadratic
Whole period			
Constant	-0.2948	-0.3193	-0.3216
	(0.0114)	(0.0123)	(0.0128)
Weeks		-0.001	-0.0013
		(0.0002)	(0.0004)
Weeks^2 ($\times$ 100)			0.0
			(0.0)
Period after			
Constant	-0.0112	0.0186	0.0228
	(0.0047)	(0.0084)	(0.0102)
Weeks ($\times$100)		0.0038	-0.0017
		(0.0018)	(0.0024)
Weeks^2 ($\times$ 100)			0.0051
			(0.0012)
R^2	0.5662	0.5665	0.5667
Number of observations	43445	43445	43445

Table 4: Results for the analysis using an index for the different amenities. R^2-within is the relative explained variation within zip code areas.

after start, (7) 9-18 weeks after start, (8) 18-26 weeks after start. We use Ridge regression in order to fit the regression line $\mu(X_i)$. For this Ridge regression, we use a polynomial of size 50 for the living area in squared feet, rent and distance to the sea. We use dummy variables for the floor, construction year, number of rooms and zip code. We interact zip code dummy variables with living area and distance to the sea and we also interact the construction year dummy variables with living area. This leaves us with 414 different regressors or features. For the other regressions, we only use the living area, the rent and distance to the sea as regressors. We use 10 folds to find the tuning parameter of Ridge regression and K is set equal to 25. We calculate the standard errors based on bootstraps.

Note that the pre-trend now depends on $\Psi_2(X_i)$ as defined in (4.4). Apart from the fact that this depends on the regressors included in X_i, it is not estimated in our estimation method as described in Section 5.2. Hence, we do not report it here. Figure 6 reports the results of our estimation method. We find a somewhat higher immediate impact of the pandemic. It equals 0.0193 rather than the 0.0066

	Restaurant	Museum	Cinema	Bar	Pub	Theater	Fitness	Mall
Whole period								
Constant	-0.2023	-0.1765	-0.1152	-0.1938	-0.1406	-0.2588	-0.0608	-0.4832
	(0.013)	(0.0127)	(0.0127)	(0.0111)	(0.0104)	(0.0097)	(0.0107)	(0.0167)
Weeks	-0.1736	-0.0976	-0.0913	-0.0871	-0.0729	-0.0906	-0.1116	-0.1018
	(0.0357)	(0.0244)	(0.0223)	(0.0165)	(0.0184)	(0.0189)	(0.0257)	(0.0188)
Period after								
Constant	0.0309	0.0042	0.0112	0.0195	0.0157	0.0212	0.0138	0.015
	(0.0171)	(0.0116)	(0.0106)	(0.0078)	(0.0088)	(0.009)	(0.0124)	(0.009)
Weeks	0.0051	0.0065	0.0069	0.002	0.0015	0.0021	0.0046	0.0047
	(0.0037)	(0.0025)	(0.0023)	(0.0017)	(0.0019)	(0.0019)	(0.0026)	(0.0019)
R^2-within	0.5623	0.5604	0.5581	0.5613	0.5595	0.5679	0.5576	0.5691

Table 5: Results for the analysis using an index for the different amenities

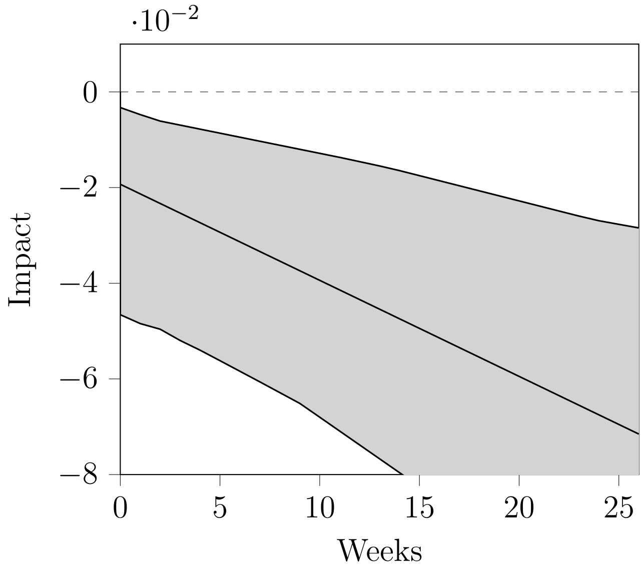


Figure 6: Results of main analysis

in our main analysis and, based on our bootstrap samples, it is significantly different from zero. Also the impact per period is higher. It equals 0.0179 but cannot be compared directly with the previous results as it is based on the change in 9 weeks rather than only 1 week. Calculating it in terms of weeks results in a coefficient equal to 0.001976. This coefficient is also significantly different from zero. Note that the total impact does become significantly different from zero already after the first period of 9 weeks when looking at Figure 6.

5.3 Results on synthetic control groups

Table 6 lists the results of the synthetic control groups. We only present here the results of the period after the start of the pandemic as the results before the start of the pandemic are always zero as a result of the estimation procedure (apart from the constant). We present the standard errors of these results based on the weighted bootstrap using weights of the exponential distribution. The results for the constant are negative but very small. The coefficient is not significantly different from zero. The linear model has both a negative constant as well as slope coefficient and also here the impact is very small. Neither of these coefficients are significantly different

	Constant	Linear	Quadratic
Period after			
Constant	-0.0060	-0.0037	0.0057
	(0.0051)	(0.0077)	(0.0111)
Weeks (× 100)		-0.0243	-0.1904
		(0.0430)	(0.1771)
Weeks^2 (× 100)			0.0058
			(0.0061)

Table 6: Results of the model in equation (4.1) using synthetic control groups.

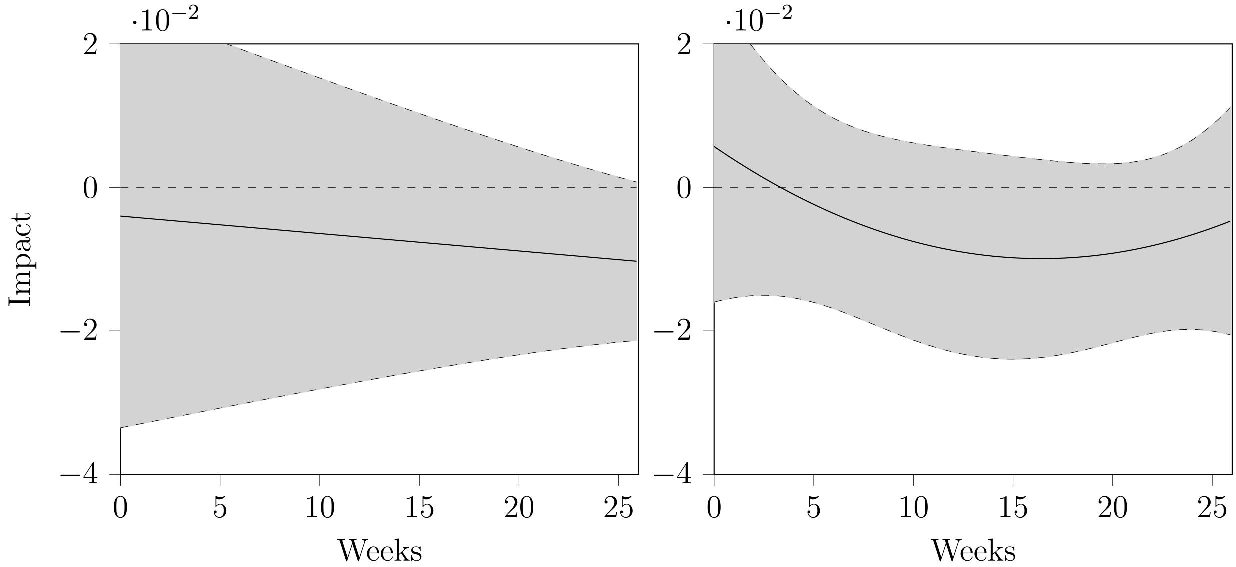


Figure 7: Results of the analysis using synthetic control groups.

from zero. Finally, the quadratic model has a positive and small constant, a negative and large linear coefficient and very small quadratic coefficient. Again, none of these are significantly different from zero. Figure 7 presents the results of the model using the synthetic control groups. Also this picture does not give any results which are significantly different from zero.

6 Robustness checks

6.1 Area of investigation

As discussed earlier, one of the problems of our analysis is that centrally located houses are around 75 percent more expensive on average than houses that are the least centrally located. One way to solve this is to use only those houses that are located in the zip codes that have both centrally and the least centrally located houses. This has the caveat of having very few houses for every week of observation. Nevertheless, we can check whether our results presented in Section 5 are robust and do not change signs. Table 7 lists the results of (4.1) using only those observations that are either the most centrally located or least centrally located (*i.e.* columns 2 and 3 in Table 1). Table 7 lists the results in the case that we use all observations.

By comparing Table 7 with Table 2, it is possible to conclude that the results are qualitatively similar. That is, the negative impact for central locations becomes more important over time. Looking at the second column of that table, the impact of the number of weeks is almost twice as large as before, but the standard error is much higher. Another important difference is the fact that the constant is now positive and quite large but insignificant. Comparing Tables 8 with 3 also results in the conclusion that these tables are qualitatively similar, but here we find a lower slope coefficient for the weeks in the linear model. Combined with a higher standard error, we conclude that this coefficient is only significant at the 10 percent level.

6.2 Using clustered standard errors

As a robustness check, we also provide the standard errors based on clustering of the postal codes. Note that the use of such clustered standard errors is only correct in the case that there is substantial heterogeneity in terms of the impact of the building sites between the postal codes (see Abadie, Athey and Wooldridge, 2017). In fact,

	Constant	Linear	Quadratic
Whole period			
Constant	0.0911	0.1025	0.1248
	(0.0077)	(0.0108)	(0.0129)
Weeks ($\times$ 100)		0.0443	0.3329
		(0.028)	(0.001)
Weeks^2 ($\times$ 100)			0.0058
			(0.0019)
Period after			
Constant	0.0109	0.027	0.0076
	(0.0079)	(0.0157)	(0.0234)
Weeks ($\times$ 100)		-0.2333	-0.5847
		(0.088)	(0.3651)
Weeks^2 ($\times$ 100)			-0.0036
			(0.0126)
R^2-within	0.5782	0.5787	0.5792
Number of observations	7916	7916	7916

Table 7: Robustness check using only those observations in zip codes with both central and decentral locations. Heterokedasticity robust standard errors between parentheses.

	Constant	Linear	Quadratic
Whole period			
Constant	0.0572	0.0691	0.0757
	(0.0029)	(0.0042)	(0.0055)
Weeks ($\times$ 100)		0.0478	0.1326
		(0.0129)	(0.0005)
Weeks^2 ($\times$ 100)			0.0017
			(0.0009)
Period after			
Constant	0.0079	-0.0011	-0.0042
	(0.0035)	(0.0072)	(0.0114)
Weeks ($\times$ 100)		-0.0677	-0.2236
		(0.0404)	(0.1729)
Weeks^2 ($\times$ 100)			0.0008
			(0.0058)
R^2-within	0.5683	0.5685	0.5686
Number of observations	25025	25025	25025

Table 8: Robustness check using only those observations in zip codes with both central and decentral locations – all observations included. Heterokedasticity robust standard errors between parentheses.

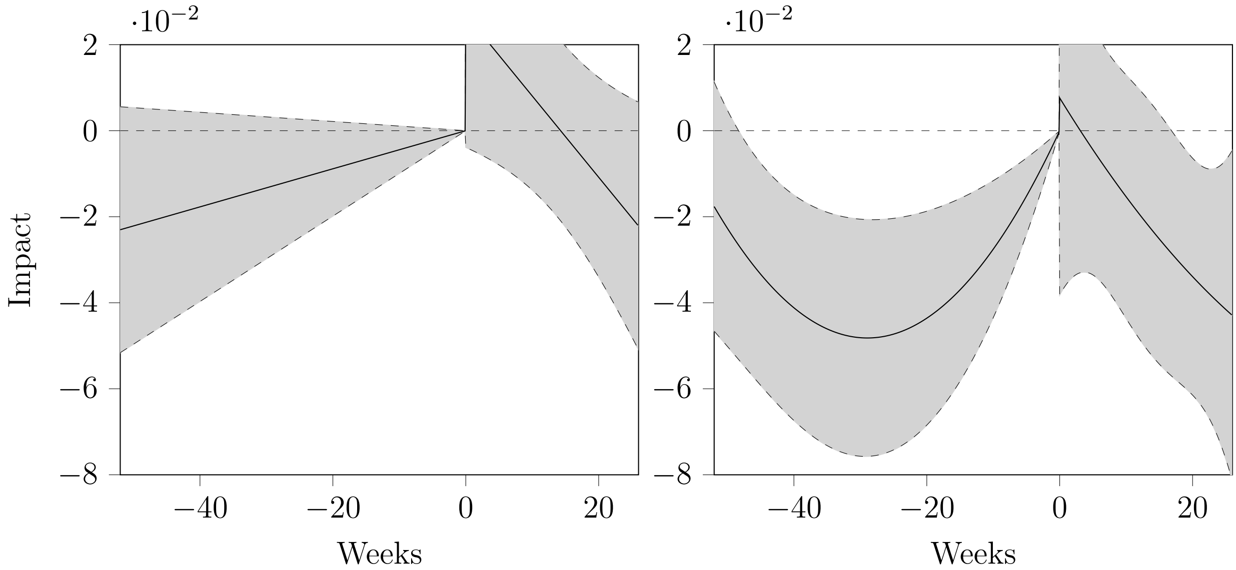


Figure 8: Robustness check using only those observations that are located in the zip codes that have both centrally and less centrally located houses. Only observations that are the most centrally or the least centrally located.

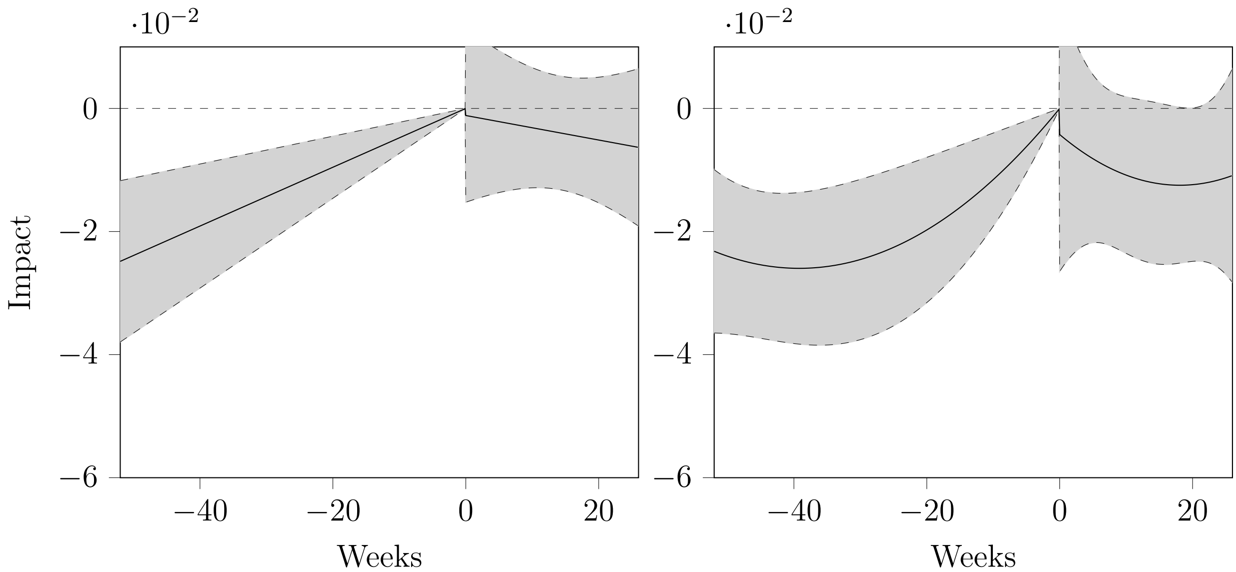


Figure 9: Robustness check using only those observations that are located in the zip codes that have both centrally and less centrally located houses. All observations included.

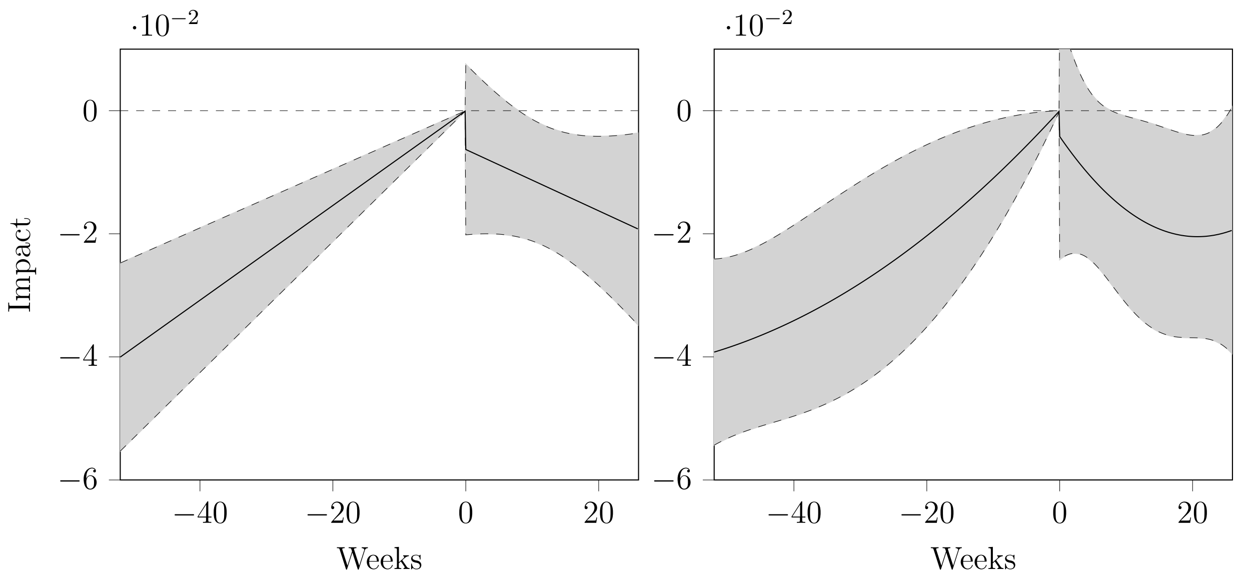


Table 9: Robustness check using clustered standard errors (based on zip codes).

the impact of the building sites is modelled to be homogeneous, *i.e.* specification (4.1), which makes the use of the standard heteroskedasticity robust standard errors more appropriate. Nevertheless, given the fact that we use three different building sites, which implies that the postal codes that are affected can be quite far from each other, it is possible that there is heterogeneity. As can be seen from the second column of Table 9, the clustered standard error of the impact is somewhat higher than the heteroskedasticity robust standard error, but the impact is still significantly different from zero (based on a significance level of 5 percent).

6.3 Investigating the impact of commuter train and underground stations

As discussed earlier, one of the reasons for our obtained results may be the presence of commuter train and underground stations near the urban amenities. As the use of these train stations also diminished during the Covid-19 pandemic, it is possible that our results are driven by the reduced value of places nearby those stations. We investigate this by looking at all the commuter train stations (Pendeltåg) and underground stations (Tunnelbanna). We estimate the following version of our baseline model

$$\begin{aligned}\log P_i = D_i^{\text{central}}\Phi_1(T_i) + D_i^{\text{central}} D_{T_i}^{\text{post}}\Phi_2(T_i) + \\ D_i^{\text{commuter}}\Phi_3(T_i) + D_i^{\text{commuter}} D_{T_i}^{\text{post}}\Phi_4(T_i) + \beta' X_i + \alpha_{J_i} + \lambda_{T_i} + U_i,\end{aligned} \tag{6.1}$$

where D_i^{commuter} represents a dummy for whether the location has a nearby train station. We take a walking distance of 5 minutes into account. Table 10 lists the results of our exercise. The results for our central locations are remarkably comparable with the ones presented in Table 2. Moreover, we find locations close to commuter train stations slightly more expensive than locations not located close to such stations. Nevertheless, there is no evidence of the fact that the pandemic had any impact on the price of houses close to commuter train stations.

7 Conclusions

This paper looked at the potential diminishing willingness to pay for urban amenities such as restaurants, bars, pubs and museums in the first half year after the outbreak of the Covid-19 pandemic. We investigated this based on the walking distance from transacted apartments to these urban amenities in combination with the hedonic pricing model. We find some decline in (the trend of the) apartment prices close to these amenities in comparison to apartments that are further from these amenities.

One potential thread for our analysis is the fact that apartment prices in the attractive central neighborhoods dropped during the Covid-19 period due to the economic crisis following the pandemic. We expected this not to be an important issue as apartment prices increased substantially during the period of analysis. In addition, the economic situation in Sweden during the crisis has not been so bad.

	Constant	Linear	Quadratic
Whole period			
Central locations Constant	0.0824	0.1020	0.1068
	(0.006)	(0.007)	(0.007)
Weeks (× 100)		0.0008	-0.0014
		(0.0101)	(0.0004)
Weeks^2 (× 100)			0.0012
			(0.001)
Commuter locations			
Constant	0.0141	0.0130	0.0069
	(0.002)	(0.004)	(0.005)
Weeks (× 100)		0.0000	-0.0008
		(0.0101)	(0.001)
Weeks^2 (× 100)			-0.0016
			(0.001)
Period after			
Central locations			
Constant	0.0063	-0.008	-0.0069
	(0.003)	(0.007)	(0.0087)
Weeks (× 100)		-0.1142	-0.2898
		(0.037)	(0.162)
Weeks^2 (× 100)			0.0029
			(0.005)
Commuter locations			
Constant	-0.0027	0.005	0.0092
	(0.004)	(0.007)	(0.012)
Weeks (× 100)		-0.0447	0.0800
		(0.040)	(0.175)
Weeks^2 (× 100)			0.0000
			(0.006)
R^2	0.9116	0.9117	0.9117
Number of observations	43445	43445	43445

Table 10: Robustness check taking account of the location of train stations.

References

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

[2] Abadie, A. S. Athey, G. Imbens and J. Wooldridge (2017), "When should you adjust standard errors for clustering?", working paper, Massachusetts Institute of Technology, Cambridge (MA)

[3] Ahlfeldt, G.M., F. Bald, D. Roth and T. Seidel (2020), "Quality of life in a dynamic spatial model.", working paper, London School of Economics.

[4] Ahlfeldt, G.M., K. Moeller, S. Waights and N. Wendland (2017), "Game of zones: the political economy of conservation areas.", *Economic Journal* **127**, F421–45.

[5] Ahlfeldt, G.M. and G. Kavetsos (2014), "Form or function?: the effect of new sports stadia on property prices in London.", *Journal of the Royal Statistical Society: Series A* **177**, 169–90.

[6] Ahlfeldt, G.M. and N. Holman (2018), "Distinctively different: a new approach to valuing architectural amenities.", *Economic Journal* **128**, 1–33.

[7] Banjaree, P. and C.L. Benkhard (2005), "Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach.", *Journal of Political Economy* **113**, 1239-76.

[8] Bjerke, L. and C. Mellander (2017), "Moving home again? Never! The locational choices of graduates in Sweden.", *Annals of Regional Science* **58**, 707–29.

[9] Black, S.E. and S. Machin (2010), "Housing valuations of school performance.", *in Handbook of the Economics of Education* **3**, 486–519.

[10] Brueckner, J., M.E. Kahn and G.C. Lin (2021), "A new spatial hedonic equilibrium in the emerging work-from-home economy?", working paper, University of California at Irvine.

[11] Cheshire, P. and S. Shepard (1995), "On the price of land and the value of amenities.", *Economica* **62**, 247–67.

[12] Claeson, M. and S. Hanson (2021), "Covid-19 and the Swedish enigma.", *The Lancet*, 259–61.

[13] Davis, M.A., A.C. Ghent and J.M. Gregory (2021), "The work-from-home technology boon and its consequences.", working paper, Rutgers University, Newark.

[14] Gibbons, S. (2015), "Gone with the wind: Valuing the visual impacts of wind turbines through house prices.", *Journal of Environmental Economics and Management* **72**, 177-96.

[15] Gibbons, S. and S. Machin (2005), "Valuing rail access using transport innovations."*Journal of Urban Economics* **57**, 148–169.

[16] Gibbons, S. and S. Machin (2008), "Valuing school quality, better transport, and lower crime: evidence from house prices.", *Oxford Review of Economic Policy* **24**, 99-119.

[17] Gibbons, S., S. Machin and O. Silva (2013), "Valuing school quality using boundary discontinuities.", *Journal of Urban Economics* **75**, 15–28.

[18] Ihlanfeldt, K. and T. Mayock (2010), "Panel data estimates of the effects of different types of crime on housing prices.", *Regional Science and Urban Economics* **40**, 161-72.

[19] Koster, H.R.A., J.N. van Ommeren and P. Rietveld (2016), "Historic amenities, income and sorting of households.", *Journal of Economic Geography* **16**, 203–36.

[20] Ling, D.C., C. Wang and T. Zhou (2020), "A first look at the impact of COVID19 on commercial real estate prices: asset level evidence.", working paper, University of Florida, Gainesville.

[21] Machin, S. (2011), "Houses and schools: valuation of school quality through the housing market.", *Labour economics* **18**, 723-29.

[22] Nathan, M. and H. Overman (2020), "Will coronavirus cause a big city exodus.", *Environment and Planning B: Urban Analytics and City Science* **47**, 1537–42.

[23] Nie, X, Lu, C. and S. Wager (2021), "Nonparametric heterogeneous treatment effect estimation in repeated cross sectional designs.", working paper, Stanford University.

[24] Redmond, L.S. and P.L. Mokhtarian (2001), "The positive utility of the commute: modeling ideal commute time and relative desired commute amount". *Transportation*, **28**, 179–205.

[25] Robinson, P.M. (1988), "Root-N-consistent semiparametric regression.", *Econometrica*, 931–54.

[26] Rosen, S. (1974), "Hedonic prices and implicit markets: product differentiation in pure competition.", *Journal of Political Economy* **82**, 34—55.

[27] Rosenthal, S.S, W.C. Strange, and J.A. Urrego (2021) "Are city centers losing their appeal? Commercial real estate, urban spatial structure, and COVID-19.", *Journal of Urban Economics*, forthcoming.

[28] Van Vuuren, A.P., J. Kjellander, V. Nilsson (2019), “ Refugees and apartment prices: A case study to investigate the attitudes of home buyers.”, *Regional Science and Urban Economics* **77**, 20–37.