Bigger and not to far :
How the trade-off between purchasing power and transport connectivity shaped the evolution of the Paris area housing market

Pierre Vidal
PhD student, Université Cergy-Pontoise
Economist, MeilleursAgents.com

Thomas Lefebvre
Research fellow, Université Paris Dauphine
Scientific director, MeilleursAgents.com

Working paper

Abstract : In this study we try to understand how the trade-off faced by households between housing purchasing power and transport connectivity impact real estate prices evolution. We computed hedonic price indexes for every municipality of the Paris area between 2001 and 2011, then used them to build a housing purchasing power index for every typology of households living there. Aggregating timetables of metro and suburb trains, we computed a connectivity similarity index of the same municipalities. Merging them in an attractiveness score, we were able to explain part of the great diversity of price evolutions across the Paris area, over this ten-year period.

Introduction:

Fulgence Bienvenüe, civil engineer and conceptor of the Paris Metro, established general guidelines for his transportation network : lines are independent, trains always stay on the same line and stop at every station. Lines were designed so that no point in Paris was more than 400 meters away from a metro station and that no trip involved more than two connections. Those guidelines, which governed the route and the construction of the metro until the death of its conceptor, still shape the Parisian urban space. Indeed planning of public transportation infrastructure is regarded nowadays as one of the most important amenities of urban renewal. Appreciation of served neighborhoods, redevelopment of the public spaces and launch of urban operations are direct consequences of public transport planning. Thus the impact of transports on housing prices has been largely documented.

Already in 1826, Von Thünen shows spatial organization of agricultural land is a function of transport cost : farmer rent decreases as distance to the market increases correlatively with transport costs. Area of production is thus delimited by the distance where the rent get null. Landowners organize the different cultures spatially in order to maximize their rents.

Similar concepts have been later used in theoretical models of urban economics related to households settlement choice and real estate price formation. The housing market seems indeed affected by public transport. Cost of transport is key in households trade-off between their location and the land value they are ready to pay : increase in transport costs is immediately linked to a higher desirability for central locations (Alonso, 1964). Conversely improvement in transports decreases value of central properties, increase those of the periphery, open new areas for urbanization and expand the city limits.

In the residential real estate market households compete with each other for land use. This competition reveals itself in the housing prices. As they value access to transportation networks and especially their hubs, they are ready to outbid to get access to the locations close to their entry points (Ana, 1985).

In most of the studies and models about households locations, authors used long term static models. As all equilibrium situation is the output of a more or less long evolution, static description seems insufficient and should be completed with a dynamic approach that accounts for sustainability of infrastructures as well as households anticipation impacting their decisions (Boniver, 1979). Time and space should be incorporated in a continuous model.

This study aims to propose a new theoretical framework to housing prices dynamics, within the urban area, based on the trade-off households faces between the increase of their real estate purchasing power and the minimization of their commute time. We study more precisely the case of the Paris area and its housing prices evolution between 2001 and 2011. Along this period real estate prices have boomed, squeezing Parisians and commuters housing purchasing power that had to migrate to previously discarded locations. We argue that these internal migrations were directed by the public transport lines and caused asymmetric price evolution.

The rest of the paper is organized as follows. Section I presents the construction of hedonic prices indexes for 125 municipalities of the Paris area, including Paris arrondissements and cities of the "Petites Couronnes" (the three departments that surround Paris), and how we used these price indexes to construct housing purchasing power indexes. Section II explains how we aggregate public transport services data to build a commute time matrix for most of those municipalities. Our trade-off mechanism is presented in Section III. We then show, in section IV, how the attractiveness score designed thanks to this framework can help us forecast price evolution, confronting it to price variations across the Paris area during the 2001-2011 period. Section V concludes.

I. Hedonic price and housing purchasing power indexes

The first step of our analysis is to know how housing prices evolved in the Paris area over our period of interest. Precisely we focus on Paris and the so-called "Petite Couronne", the three departments that surround Paris (Haut-de-Seine, zip code 92; Seine-St-Denis, zip code 93; Val-de-Marne, zip code 94). Note that, as the Paris region is mostly a dense urban area, we focus only on the prices of apartments. In order to do so we use the price indexes of the French real estate internet platform MeilleursAgents.com. These indexes are based on a large database of sales references that go from early 2000s to present. They are based on a hedonic method that we describe completely in the following section.

Filters

The filters we used prior to any calculation have two goals : to get rid of incomplete references and to eliminate outliers. As we choose to use a hedonic method, the sale references we used have to describe the apartment precisely enough. The database being built by manual inputs, some variable might be missing or taking absurd value. To prevent our result from being perturbed by such mistake we apply filters that discard between 5% to 15% of the references, depending on the city. These filters are fully described in annex I.

Hedonic regression model

The hedonic model we regress study the price over meter square at which the apartment has been sold. We control for different features that impact the price of apartments such as the number of rooms, number of bathrooms, the floor the apartment is located at, etc… The complete list of variables is available in annex 2. Of course, as we intend to construct price index, we control for the year the transaction occurred at. Thus we measure the price evolution for a reference apartment of two rooms, one bathroom, with no parking spot, no balcony, no cellar and no secondary room attached, located between the first and the third floor of a building from the beginning of the 20th century, with an elevator.

$$pm^2 \sim \sum_i \alpha_i feature_i + \sum_j \beta_j year_j$$

Once we perform an OLS regression of the previous model against our data for every municipality in the Paris metropolitan area, to construct our prices indexes, we just add the coefficient corresponding to each year (the βs of the previous formula) to the intercept and divide by it to obtain a price index based in 2001. For example, for the city of Neuilly-sur-Seine, we obtain the following price index (the full result of the regression is available in the annex) :

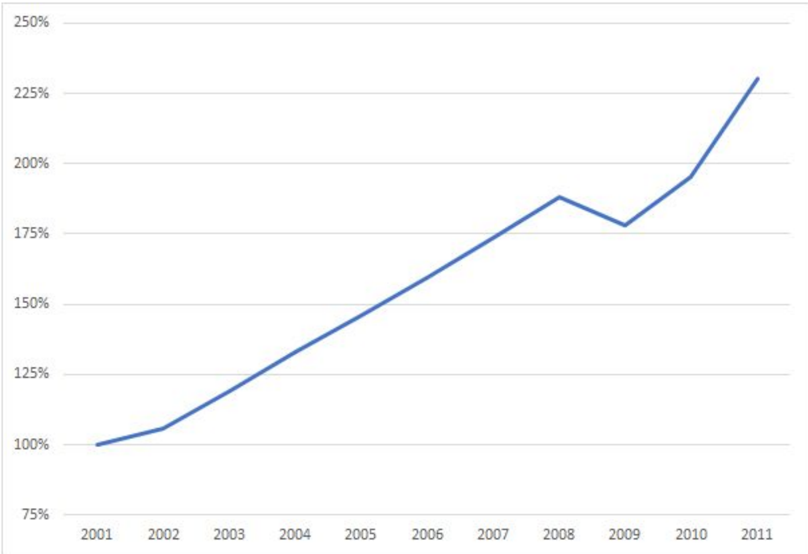

Fig. 1: Hedonic price index of Neuilly-sur-Seine between 2001-2011

This evolution is symptomatic of what most municipalities within the Paris area experience during the period we are interested in. First a constant and strong price surge from 2001 to 2007 or 2008 depending on the municipality. Then a sharp regression in 2008-2009 due to the global financial crisis. And finally two years of increased prices. The summary statistics for the price evolution across all municipalities, for every year and the whole period are regrouped in the table below.

Overall, prices have been multiplied by roughly x2.5 on average at an average pace just under +9%. Note that this price boom has been unequal across the area, as top performer, Montreuil, saw its housing prices progress by +230% when, at the other end of the spectrum, Clichy-sous-Bois prices only increased by +73% (still several times higher than the general inflation of +19% over the period according to the French bureau of statistics INSEE). This disparity is what interests us here, as we argue it can be explained by the trade-off between housing purchasing power and commuting time.

	Average	**Standard Dev**	**Median**	**Min**	**Max**
2001-2002	11,2%	4,2%	10,9%	1,7%	25%
2002-2003	14,5%	4,3%	14,7%	0,7%	29%
2003-2004	16,9%	4,4%	17,5%	7,6%	30%
2004-2005	17,4%	4,9%	17,1%	7,4%	33%
2005-2006	11,7%	3,4%	11,6%	6,0%	21%
2006-2007	7,0%	3,3%	6,9%	-0,8%	21%
2007-2008	0,0%	2,9%	-0,3%	-6,9%	10%
2008-2009	-2,7%	3,7%	-3,4%	-13,1%	7%
2009-2010	10,5%	5,9%	9,3%	-4,7%	24%
2010-2011	8,4%	4,8%	8,4%	-5,1%	20%
2001-2011	**143,4%**	**31,5%**	**141,3%**	**73,0%**	**230%**

Table 1 : Summary statistics of price evolutions of the Paris area between 2001-2011

Construction of the real estate purchasing power indexes

The key value in our trade-off model is not the housing prices but a more complex quantity : the real estate purchasing power. The real estate purchasing power index we compute represents the surface per person, in square meter, a household can afford dedicating one third of its revenues to the payment of a twenty years maturity loan contracted to pay the apartment, with no down payment. The formula, for household *i*, in city *A*, at time *t* and for a maturity *M* equal to 20 years :

$$HPP_{i,A,t} = \frac{1/3\ Revenu_{i,t}\ pm^2_{A,t} \tau_t}{1-(1+\tau_t)^M}$$

We used the price index presented above (note that here we used the indexes in real value and not in points based in 2001), the average rate for housing loan recorded by the Banque de France and the revenue as computed by INSEE. The real estate purchasing power is calculated every year for different typologies of households. INSEE recorded for the period the deciles of revenue for households of different sizes, from one person to five and more, every year and for every city and arrondissement. Thus we can easily compute the purchasing power per person (we consider that five and more persons as only five persons) of every typology of household in their own city but also in all the others, as we are interested in the migration across the metropolitan area.

The following table represents summary statistics for a two-person households across the area, in their own municipalities, along 11 years. As one can see the price boom describe above had a major impact on the housing purchasing power of the Paris area inhabitants, as on average it falls from -43%. According to the French administration (INSEE) a housing unit is considered as overcrowded if it has less than 18 m² per person. In 2001 the 1st quartile was above this level, in 2011 the median is equal to this value.

Year	Min	1st Quart	Median	Average	3rd Quart	Max
2001	6	22	32	35	45	92
2002	6	21	31	33	44	88
2003	5	19	29	31	41	82
2004	5	17	26	28	36	71
2005	5	16	23	25	33	68
2006	4	14	21	22	29	63
2007	4	13	19	20	27	57
2008	4	13	19	20	27	56
2009	4	14	22	23	31	63
2010	3	13	20	22	29	59
2011	3	12	18	20	26	54

Table 2 : Summary statistics of housing purchasing power per person of households two persons in the Paris area between 2001-2011

II. Connectivity to the public transportation network

As we want to understand the trade-off between maximization purchasing power and the minimization of commute time, we need to construct a metric that shows how well a place is connected to the rest of the metropolitan area. In simple terms we need to know how long it takes for a person to go from one place to all the others. This commute time can be split in two: the walking time from home to a public transportation station and the travel time inside the public transportation network.

Let us start with the latest. The Paris area urban transportation is operated by two state-owned companies : the Régie Autonome des Transports Parisien (or RATP) and the Société Nationale des Chemins de Fer (or SNCF). They offer a dense multimodal network of fourteen metropolitan lines that operate mostly in Paris and the closest suburbs, fifteen suburb train lines (RER or Transilien) which connect the suburbs to Parisian hubs, nine tramways and more than a thousand bus lines which cover more densely the most of the urban spaces. As they are the one used by "franciliens" to commute on longer distance (city to city or from one arrondissement to another) we focus only on the rail transportation network (Metro, RER and Transilien) which is mostly radial and increasingly dense from the suburb to the center of Paris.

Both companies, RATP and SNCF, make their timetables for the next two weeks and connection time within their stations available at the GTFS format. We merged the two datasets and matched the common stations based on their GPS coordinates to create a complete timetable of all the rail transportation network. As we are interested in the impact of the transportation to the real estate market, we restricted our dataset to the trains operating from 7:30 to 9:00 on a weekday to simulate the home-to-work trip. Based on this dataset we created a directed graph by applying the following rules :

- every stop is a vertex
- stops within a same station are connected by an edge
- if a train stops in stop A at time t then in stop B at time t+Δt, then an edge of weight Δt connects A to B
- the edges that represent transfers between stops of a same company within a station are weighted by the walking time indicated by the companies' datasets plus 120 seconds of waiting time, as a train stop every 4 minutes on average
- the edges that represent transfers between stops of different companies within a station are weighted by an arbitrary 240 seconds time, as we have no indication of transfer time for those, plus 120 seconds of waiting time, as a train stop every 4 minutes on average

Once the graph is built we used the Dijkstra's algorithm (of the R *distances* method of the *igraph* package) to compute the shortest paths from every stop to all the others, thus obtaining a distance matrix of all the stops in the Paris metro area.
The distance from one stop to another is only a part of the story. To compare one city or arrondissement connectivity to another, we must take into account its density in stops and how long its inhabitants have to walk from their home to the station. To compute such a time we used the geographical shape of the cities of OpenStreetMap. For every city we located the stations that are in the city or just nearby, less than 300 meters from the city boundaries. We draw 1,000 random points within the city limit and measure the Euclidean distance of each point to all the stations of the city. Given a walking speed of 3 km/h, as we take into account the Euclidean distance and not the longer actual walk path along the streets, we compute the time a person that lives at one of the points needs to reach each station. Note here that our method excludes *de facto* municipalities without any metro or RER station (20 municipalities over the 143 of the studied area).

Finally, we want to compute how long a person that lives in a given city, take to reach any station of the network. To do this we computed the shortest journeys of all the 1,000 random points to all the stations, as they first walk to a station within the city limit then used the public transportation network to reach their destination. Then we average to get the mean time for an inhabitant of a given to every station within the Paris area.

Of course all stations are not equally interesting. People would rather be able to reach Châtelet-les-Halles, the main hub of the network, quickly than Jouy-en-Josas, a small town close to Versailles. That is why we reduce the list of stations we considered to the main hubs as they allow, through interconnections, to reach all other stations. The hubs we selected are the ones detected by Gleyze (2003). He clustered those hubs under the name "Concorde" type of stations and we list them below :

- Châtelet les Halles (1st arrondissement)
- Concorde (1st arrondissement)
- Etoile (8th arrondissement)
- Gare du Nord (10th arrondissement)
- Madeleine (8th arrondissement)
- Montparnasse (14th arrondissement)
- Nation (11th arrondissement)
- Opéra (9th arrondissement)
- République (10th arrondissement)

We took the liberty to add the station La Défense (Puteaux), which serves the business district of the same name, which also offers multiple connexions but was excluded from the Gleyze study because located outside of Paris.

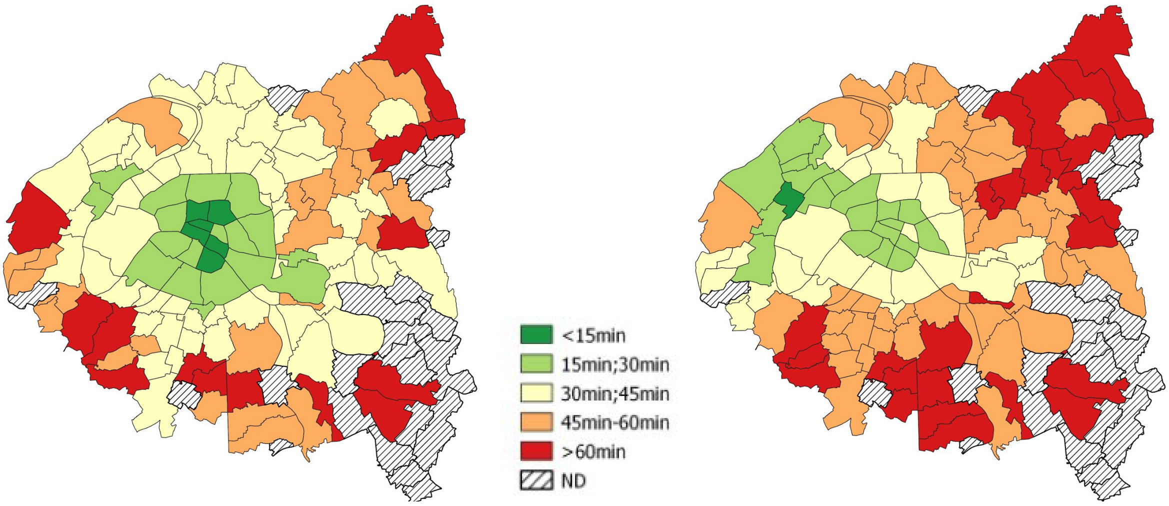

Fig. 2: Average time to reach Châtelet-les-Halles (left) and La Défense (right)

Our interest for the commute time is that it affects the housing market. We argue that the decision to move from one place to another is directed by how close those two places are to the same place. Thus we compute a similarity loss/gain matrix of the sum of the signed quadratic difference between the vector representing the transport time from a city to our list of hubs and the vector of all the other cities. Thus the similarity between city *i* and *j* is :

$$similarity[i, j] = \sqrt{sgn\left(\sum_{k \,\varepsilon\, hubs} sgn(\Delta t_{i,j->k})(\Delta t_{i,j->k})^2\right) \sum_{k \,\varepsilon\, hubs} (\Delta t_{i,j->k})^2}$$

with : $\Delta t_{i,j->k} = time_{j->k} - time_{i->k}$

A positive (resp. negative) *similarity[i, j]* means that on average, moving from city *i* to *j* will increase (resp. decreased) your commute time to the main hubs. Note that the use of a quadratic difference allow us to better take into account the similarity of the access and not just the average time. Indeed, as it is reasonable to think that people tend to choose to live where it is convenient for them, it makes sense to suppose they would rather move a little further away from the city center in the same direction than completely go across the city to reach a place that, on average as the same commute time.

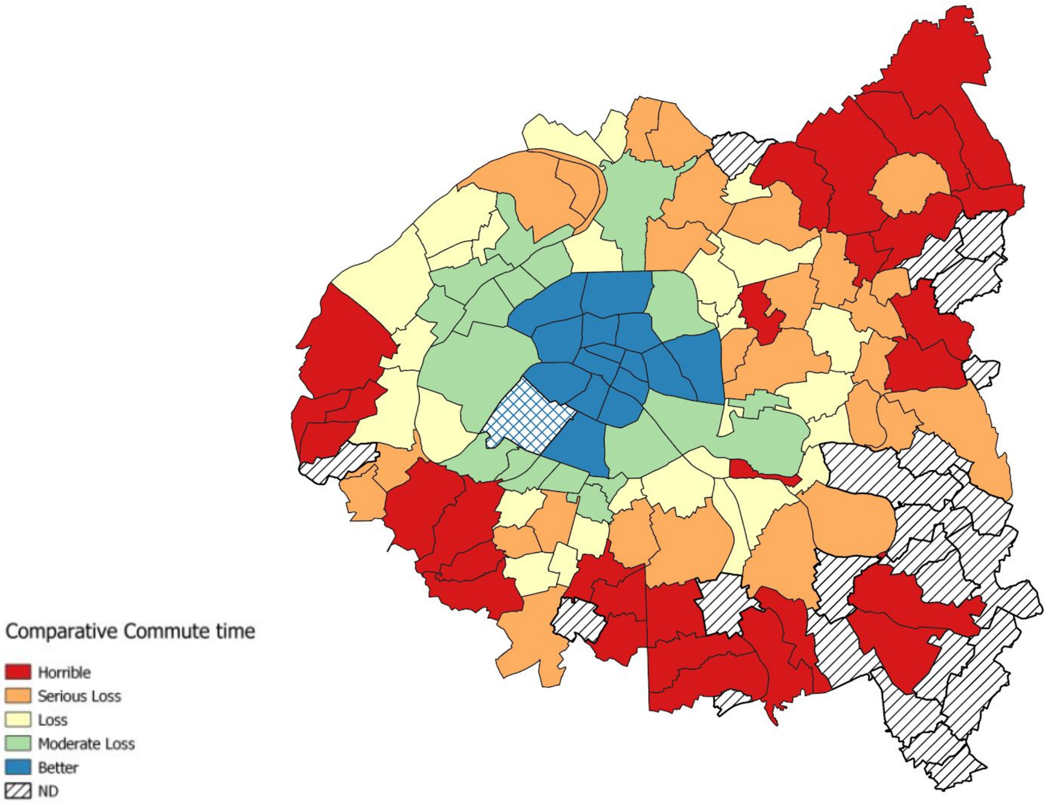

Fig. 3: Commute time similarity compared to Paris 15th arrondissement (blue grid)

III. Arbitrage mechanism

At this stage of our study, we have not fully develop the trade-off mechanism that would allow us to understand the differences in price variations across the studied area. The following as to be understood as a first draft that takes the main points of the model that is yet to be constructed.

The demand for housing increases sharply with the access to public transportation, that is especially true with a mono-centric city such as Paris and the Paris metropolitan area. A first and good rule of thumb to describe the spatial distribution of price is a downward gradient from the center of Paris to the suburb.

As they want to increase the size of their housing units, because of a new birth for example, or just to maintain the same size of habitation despite the price surge, Parisians and commuters have to move away from the center to look for more affordable price. We argue that these migrations are directed by the public transportation lines.

Let us consider a household i, characterized by its size and net annual revenue at date t, living in a arrondissement or city A. Its revenue and the market housing loan rate at date t allow him to purchase a surface of $HPP_{i,A,t}$ per person in its municipality of residence. The same household i from city A can purchase a surface per person of $HPP_{i,B,t}$ in municipality B at the same date. In our framework, the difference between municipalities A and B for the household is the sole difference of commute time implied by living in B rather than in A, measured as explain earlier. If the household i does have a housing purchase project, we thus argue that the probability that it moves from A to B is proportional to the log of its increased in purchasing power and inversely proportional to the exponential cost of increasing commute time :

$$Prob(i, A \to B, t) \sim \frac{log(1+max(HPP_{i,A,t}-HPP_{i,B,t}\ ;\ 0))}{exp(similarity(A,B))}$$

The numerator represents the gain in purchasing power the household would benefit from migrate from A to B. In our framework households would only move to a place where they can buy a bigger home. We take a log of it considering a diminishing marginal utility of the size of the units. The denominator represents the average gain / loss in commute time the household will face from moving from A to B (see part 3). We take the exponential of it considering an increasing marginal cost of commute time (or a diminishing marginal utility of a reduction of commute time). Note that as we used both the purchasing power index and the connectivity index, we need to have both, thus we reduce our study to the 112 municipalities we have been able to compute them both.

Summing over A and i, weighting by the number of households of type i in city A at time t : $N_{i,A,t}$, and dividing by the number of households in B at time t : $N_{B,t}$, we get a measure of attractiveness of city B versus the other cities of the metropolitan area at date t. We argue that the extra demand for city B, compared to the rest of the area is proportional to this score of attractiveness. Formally :

$$Score_{B,t} = \frac{\sum_{A,i} Prob(i,A \to B,t)\, N_{i,A,t}}{N_{B,t}}$$

We cannot directly observe, nor trace the origin of, the demand but, thank to the price indexes developed in 2, we are able to know the difference in price variation of municipalities over the whole area. It seems fair to assume housing prices are directly correlated to the extra demand and thus to our attractiveness measure. More precisely the performance versus the rest of the Paris area housing market should be correlated to our attractiveness measure. Thus, computing the attractiveness measure at a given time, would allow us to predict the price variation between this time and the next period. That what we will test in part 5.

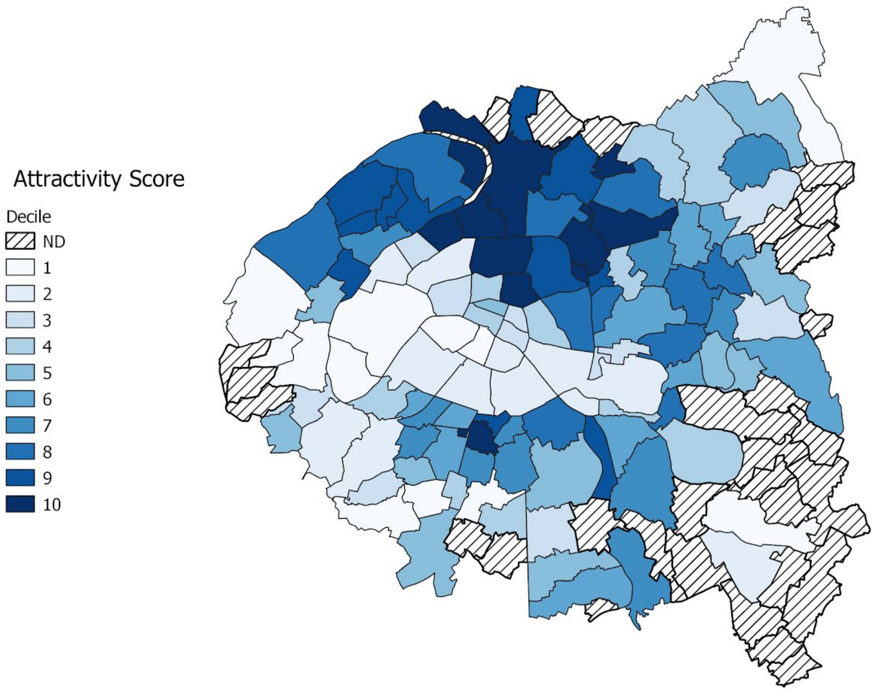

Fig. 1:Attractivity score in 2001

IV. Econometrics results

In order to test if the attractiveness score we computed from our arbitrage mechanism is a good predictor of the price variation over time, we performed panel regression of price variations from one year to the other of each municipality against our attractiveness score as computed previously. We control from fixed effects of the year and/or of the department the municipality is a part of. Controlling from an effect related to a particular year is natural, as the model is not designed to capture how prices react to macroeconomic conditions. We also controlled for fixed effect of the department as a robustness check. Indeed urban transportation connectivity is strongly related to geographical position of cities, given the strong geographical autocorrelation of housing price, one could wonder if we are not just measuring that some parts of the Paris area are just more desirable than others. For the readability of our result, we normalized the attractiveness score by dividing by the maximum attained value of our sample.

Our results are regrouped in the following table :

	model 1	model 2	model 3	model 4
coef attractiveness	8.065e-02	4.159e-02	9.862e-02	5.473e-02
p-value	3.9e-04	1.48e-03	4.54e-05	8.21e-05
Control “year’	N	Y	N	Y
Control “department”	N	N	Y	Y
R^2 projected model	1.1%	0.9%	%	1.4%
R^2 full model	1.1%	68.3%	1.7%	68.7%
#observations	1120	1120	1120	1120

Table 3 : Panel regression results of year to year price variation against attractivity score

The coefficient related to our attractiveness score is positive, as expected, significative and robust to control of the year and geography. The magnitude of the coefficient is coherent with the data as it indicates that the most attractive municipalities would face a year to year variation 7 percentage points more important than the least attractive one. That is roughly twice the standard deviation of year to year variation (see table 1). However, the explanatory power of our score is modest as R^2 of the projected model remains under 2%. That is not entirely a surprise as real estate prices is strongly correlated to macroeconomic conditions, as the strong R^2 of the full model with control over the year shows.

A unitary inspection of the score assigned to each couple municipality/year indicate that our model seems to overweight the most attractive municipalities. Indeed a regression of the yearly price variation against the log of our score (normalized by the max of this logged score) achieves better performance despite being hard to link to a tangible quantity (see annex 4 for result).

It is worth noticing that the explanatory power of our attractiveness score seems to increase as we control with the department. This could indicate that our model tends to underestimate the stigma that some poor municipalities suffer from, despite their low real estate price (as annex 5, where we created interaction variable between departments and our score, seems to confirm).

For comparison, we report in annex 6 the performance of more naïve models that try to predict the yearly variations of prices based on previous variation or previous surperformance compared to the rest of the Paris area. Those naïve models seem unable to achieve good predictions.

As real estate prices tend to peg to their previous values and as the individuals may not always be able to seize good opportunities even though they detect it, as they buy property rarely (average 3% turnover of real estate property in France, source INSEE), the attractiveness score we compute should be able to forecast price variation on a longer horizon than one year. Thus we regress price variation over different timeframe (2 years, 5 years and 10 years) against the attractiveness score of the municipalities as computed on the previous year. Results are reported below:

dependant variable	2 years variation	5 years variation	10 years variation
coef attractiveness	8.292e-02	25.92e-02	66.13e-02
p-value	1.86e-04	1.02e-07	9.88e-05
Control “year’	Y	Y	N
Control “department”	Y	Y	Y
R^2 projected model	1.4%	4.2%	13.3%
R^2 full model	78.8%	82.1%	20.7%
#observations	1008	672	112

Table 3 : Panel regression results of prices variations at 2, 5 and 10 years against attractivity score

Once again, the coefficient related to our score of attractiveness is positive, significant and of the expected magnitude (roughly two standard deviation recorded in table 1 for the 10 years variation). Note that the R^2 of the projected model increased as we try to forecast over a longer period of time. This could seem surprising but indicates that the arbitrage between purchasing power and commute time appears as structuring force of the dynamic of real estate price in the Paris area.

One can be surprised that a factor that take prices into account, through the purchasing power, can help forecast price variation ten years ahead, especially as prices experience such a surge. Two factors could help to understand this. First, as we have seen, most of this surge is due to macroeconomic factors that have affected every municipality symmetrically. Second, the very fact that our attractiveness score seems to play a small role from one year to another assure that the good opportunities it detects in terms of purchasing power survive as the increase of price due to “arbitrageurs” over short periods of time is marginal.

On a technical note, contrary to the projected model, the R^2 of the full model of the 10-year forecast is weaker to the forecast over a shorter period of time. The reason is we cannot control of the year here, as only the scores calculated for the year 2001 can be the object of a forecast over 10 years, our price indexes track evolutions over the 2001-2011 period. That is why we only have 112 observations here. As a matter of fact, every time we try to forecast one year further in the future, we lose 112 observations.

V. Conclusion

Aggregating real estate transaction data across the Paris area, along an eleven-year period from 2001 and 2011, we created 125 housing prices index, one for every Parisian arrondissement and most of the cities of the three departments surrounding the capital. Over the course of the period prices grew on average of +143% with a wide range from +73% to +230%. Combining these indexes with revenue census and housing loan rate, we computed housing purchasing power index for every typology of households (size, percentile of revenue and municipality of residence). The housing purchasing power of every typology of the household was negatively impacted by the price boom.

We argue that to compensate this loss of purchasing power, household chose to move away from the most expensive municipalities keeping their public transportation connectivity as good as possible. To verify it, we computed average commute time of 123 municipalities to the main transportation hubs. With a simple trade-off model that forecast the probability of every typology of households to move to a given municipality, comparing their housing purchasing power there and the connectivity to public transport, we computed attractiveness score for every municipality, every year. Even if it certainly needs more development, this attractiveness score is a good predictor of the difference in the municipalities housing prices variations, both year to year and on the long term and confirms our hypothesis.

Our work still have clear limitations. First of all as noticed before, our trade-off model needs to be more precisely defined. Secondly, we took the current time table to explain price variations as far back as 18 years ago. Since then some stations have been opened and the timetable has changed. Finally we considered municipalities' prices as homogene when they are not and that the impact of public transport stations on prices is probably concentrated around them.

As the "Grand Paris Express" project, a new 200 km long automatic metro planned for 2024, is expected to totally change the face of the public transportation of the region in the coming years, a natural continuation of this work would be to apply our model given the forecast of the commute time modifications it will cause. Another would be to understand, as our model seems to overestimate the attractiveness of some poorer municipalities, how poverty stigma impact the housing price and how it can be overcome at it was the case for some of the top performer municipalities such as Montreuil.

Bibliography

Anas A, Shyong Duann L (1985), *Dynamic forecasting of travel demand, residential location and land development : policy simulation with the chicago area transportation/land use system.* Papers, Regional Science Association, 56: 38-58.

Alonso W (1964). *Location and Land Use, toward a general theory of land rent*

Boniver V (1979), *Un aperçu de la nouvelle microéconomie urbaine*. Revue d'économie régionale et urbaine, n°3/4 : 326-361.

Derycke PH, Gannon F (1990), *Distance et coûts de transports. Quelques réflexions sur les politiques de réduction de la congestion urbaine.* Revenue d'économie Régionale et Urbaine n°2.

Gannon F (1992), *Modèles de la ville et politiques urbaines optimales.* Thèse pour le Doctorat es Sciences économiques, Paris, Université paris X Nanterre

Gleyze J-F (2007), *La desserte de Paris par le réseau de Metro et de RER*. HAL Sciences de l'Homme et de la Société.

Johann Heinrich von Thünen (1986), *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie oder Untersuchungen über den Einfluss, den die Getreidepreise, der Reichthum des Bodens und die Abgaben auf den Ackerbau ausüben*, Düsseldorf, Wissenschaft und Finanzen GmhH, p. 290

Masson S (1998), *Interactions entre système de transport et système de localisation. De l'héritage des modèles traditionnels à l'apport des modèles interactifs de transport et d'occupation des sol*s. Les Cahiers scientifiques du transport , AFITL, 79-108

Wingo L (1961), *An economic model of the utilization of urban land*. Regional Science, 7, 1, 191-205

Annex

Annex 1 : Filters description

Statics filters :

- price per meter square > 100€
- price > 0
- surface > 0
- surface is not null
- room count is not null
- room count ∈ [1;15]
- bathroom count is not null
- bathroom count ∈ [0;3]
- secondary room count ∈ [0;2]
- parking count ∈ [0;2]

Dynamic filters :

Dynamic filters are intended to get rid of the most extreme transactions, that could perturb our mesures. Thus we filter the top and bottom 0.5% of the transaction according to three dimensions : area, price and price per square meter. Overall those “dynamic filters” discard 3% of the data

Annex 2 : List of variables name for hedonic price index

Features	variable name	modality
Floor the apartment is situated at	f_ND	floor is not documented
	f_0	ground floor
	f_123	floors 1 to 3
	f_456	floors 4 to 6
	f_78	floors 7 to 8
	f_9plus	floor 9 and upper
Elevator	e_ND	elevator is not documented
	e_1	there is an elevator
	e_0	there is no elevator
Number of rooms	r_1	one room
	r_2	two rooms
	r_3	three rooms
	r_4	four rooms
	r_5plus	five rooms and more
Number of bathrooms	b_0	no bathroom
	b_1	one bathroom
	b_2plus	two bathrooms or more
Building period	per_nd	non documented
	per_a	built prior to 1850
	per_b	built between 1850 and 1913
	per_c	built between 1914 and 1947
	per_d	built between 1948 and 1969
	per_e	built between 1970 and 1980
	per_f	built between 1981 and 1991
	per_g	built between 1992 and 2000
	per_h	built between 2001 and 2010
Number of parking spot	park_0	no parking spot or ND
	park_1	one parking spot
	park_2	two parking spot
Secondary rooms (or maidrooms)	secr_0	no secondary room or ND
	secr_1	one secondary room
	secr_2	two secondary room
Cellar	cellar_0	no cellar or ND
	cellar_1	one cellar or more

Balcony	balcony_0	no balcony or ND
	balcony_1	balcony_available
Interaction variable rooms and bathrooms	r1b0	one room, no bathroom
	r1b1	one room, one bathroom
	r2b0	two rooms, no bathroom
	r2b1	two rooms, one bathroom
Year of sale (from 2001 to 2011)	y_20XX	Sold on year 20XX

Annex 3 : Result of the OLS regression producing the price index of Neuilly-sur-Seine

variable	effect	p-value
Intercept	3605.633	< 2e-16
f_ND	-59.426	0.55169
f_0	79.394	0.05037
f_456	228.415	6.86e-14
f_78	345.503	2.06e-05
f_9plus	437.730	0.12289
e_ND	-	-
e_0	32.478	0.26251
r1b0	-390.956	0.17031
r1b1	37.344	0.40643
r2b0	-135.053	0.62208
r_3	126.612	0.00112
r_4	356.942	2.95e-16
r_5plus	377.758	1.32e-14
b_0	-184.229	0.33020
b_2plus	214.991	2.91e-07
per_nd	-98.092	0.53422
per_a	10.406	0.95495
per_c	-107.367	0.01034
per_d	99.387	0.00372
per_e	65.369	0.12836
per_f	-77.032	0.41870
per_g	578.374	9.03e-09
per_h	1186.889	< 2e-16
park_0	565.832	< 2e-16
park_2	996.750	< 2e-16
secr_0	427.057	2.22e-16
secr_2	308.299	0.01343
cellar_1	5.402	0.88814
balcony_1	477.855	5.86e-08
y_2002	211.395	0.00676
y_2003	690.014	< 2e-16
y_2004	1192.696	< 2e-16
y_2005	1654.142	< 2e-16

y_2006	2146.809	< 2e-16
y_2007	2665.677	< 2e-16
y_2008	3177.934	< 2e-16
y_2009	2819.751	< 2e-16
y_2010	3443.617	< 2e-16
y_2011	4703.557	< 2e-16

Annex 4

A - Panel regression results of year to year price variations against logged attractiveness score

	model 1	model 2	model 3	model 4
coef attractiveness	2.708e-2	1.487e-2	4.268 e-2	5.040 e-2
p-value	1.28e-06	3.83e-06	6.14e-11	3.83e-12
Control “year’	N	Y	N	Y
Control “department”	N	N	Y	Y
R^2 projected model	2.1%	1.9%	3.8%	4.3%
R^2 full model	2.1%	68.7%	4.0%	69.2%
#observations	1120	1120	1120	1120

B - Panel regression results of prices variations at 2, 5 and 10 years against logged attractiveness score

dependant variable	2 years variation	5 years variation	10 years variation
coef attractiveness	7.626e-02	21.70e-02	63.18e-02
p-value	<2e-16	<2e-16	1.04e-12
Control “year’	Y	Y	N
Control “department”	Y	Y	Y
R^2 projected model	7.6%	16.7%	37.9%
R^2 full model	80.2%	84.5%	43.2%
#observations	1008	672	112

Annex 5 : Panel regression results of year to year prices variations against attractiveness score with interaction with departement

	Model 1	Model 2
Paris*attractiveness (p-value)	0.2229 (6.948e-3)	0.2949 (3.975e-2)
92*attractiveness (p-value)	0.28608 (4.82e-4)	0.5089 (3.34e-4)
93*attractiveness (p-value)	0.03133 (1.835e-2)	0.0600 (9.745e-3)
94*attractiveness (p-value)	0.14992 (2.920e-2)	0.3257 (6.552e-3)
Control “year’	Y	N
R^2 projected model	2.3%	2.5%
R^2 full model	68.8%	2.5%
#observations	1120	1120

Annex 6 : Panel regression results of of year to year prices variations against past performance

	model 1	model 2
previous variation (p-value)	1.888 e-02 (0.563)	-
previous surperformance (p-value)	-	1.888 e-02 (0.563)
Control “department”	Y	Y
Control “year’	Y	Y
R^2 projected model	0.0%	0.0%
R^2 full model	70.0%	70.0%
#observations	1008	1008