Discrete choice models with capacity constraints: an empirical analysis of the housing market of the greater Paris region

André de Palma* Nathalie Picard† Paul Waddell‡

January 30, 2006

Abstract

Discrete choice models are based on the idea that each user can chose freely and independently from other users in a given set of alternatives. But this is not the case in several situations. In particular, interactions and limitations can occur when the number of available products is smaller than the total demand for some alternatives; as a consequence, some individuals can be denied the good. We develop a methodology to address this problem and apply it to residential location choice, where there is reason to suspect availability constraints may limit choices. The analysis provides some theoretical developments and elaborates an iterative procedure for estimating demand in the presence of capacity constraints. The empirical application relies on the location choice model developed and estimated in de Palma et al (2006) for Ile de France (Paris region) and generalizes it to integrate capacity constraints.

Keywords: Residential location, constrained Logit, capacity constraints, sampling, price endogeneity, Ile de France.

*University of Cergy-Pontoise, THEMA and Ecole Nationale des Ponts et Chaussée. Member of the Institut universiatire de France.

†University of Cergy-Pontoise, THEMA.

‡University of Washington at Seattle.

1 Introduction

The choices that individual households make in the housing market produce aggregate outcomes that shape urban traffic conditions, patterns of poverty concentration and ethnic segregation, the quality of public schools, access to economic opportunities, and the decline and revitalization of neighborhoods – all of which are important, long-term, and related policy concerns. Efforts to model individual household residential location choices using discrete choice frameworks date at least to the path-breaking work by [Lerman1977], [McFadden1978], [Quigley1976], [Weisbrod and Ben-Akiva1980], and [Williams1979], among others. Over the past 25 years an extensive literature has examined household location choices in the housing market, significantly advancing the behavioral underpinnings and the methods used, and including efforts to represent the residential location choice as a dynamic process: see, e.g., [Anas and Cho1988].

One important issue that has not yet received sufficient attention in the literature, and is the central focus of the empirical analysis presented in this paper, is the role of availability constraints in discrete choice models. In estimating location choice models using agents observed to make choices among a set of alternatives, an implicit assumption is made that the alternatives are all available, as they would commonly be if the choice is among standard commodities such as consumer electronics. However, in the housing market and in many other market situations, such as for air travel, a problem of limited availability is not at all uncommon. A particular neighborhood may be highly desired, and few vacancies may be available to those that are searching in the area. Seats may be sold out on the particular flight itinerary a traveler wishes to take. A standard assumption in economics is that prices adjust to clear the market, and therefore putting prices on the right hand side of the model is sufficient to address this concern. This assumption may be too strong, however, in many market conditions, including the housing market. Various forms of friction in the housing market make it less than perfectly efficient. High transactions costs, attachments to social networks, non-trivial search costs, low turnover in some locations, and sellers' willingness to withold a property from the market rather than suffer a loss, among other factors, suggest that prices may not fully clear the market.

Whether prices actually do fully clear the housing market should be an empirical question rather than a strong assumption. If the assumption that prices clear the market is not valid, then it follows that coefficients estimated for discrete choice models in markets that experience some level of availability constraints will be biased, confounding the effect of the constraints with the choice preferences of agents. An important policy implication of this methodological concern is that if these constraint effects are not corrected for in estimation of a choice model, predicted shifts in demand in response to an exogenous change, such as the change in accessibility due to major transportation investments, would also be biased, leading to potentially misleading conclusions regarding the relative costs and benefits of alternative policy choices. A related concern is that housing prices are jointly determined with location choices in constrained housing markets due to the role of prices in clearing the market, even if they do not completely clear it. It is therefore necessary to address this source of endogeneity bias in the estimation process [de Palma *et al.*2006].

We are interested here in the allocation process when supply is locally smaller than demand. This allocation could be described as a search mechanism in the line of [de Palma and Lefèvre1981], later applied in the regulated housing market in Holland by [de Palma and Rouwendal1996]. Here we adopt a completely different approach to address the problem of disequilibrium between supply and demand. We assume that the final allocation mechanism should obey some simple rules (or assumptions). From these assumptions, we are able to compute the ex post allocation. That is, we are able to describe both the ex ante choice (i.e. the choice which ignores the capacity constraints) and the ex post allocation (i.e. the individual choice once the competition for scarce housing ressources has taken place). Below we describe the organization of the paper.

In this paper we develop models of residential location and housing price for the Ile-de-France metropolitan region centered on Paris, testing for endogeneity between prices and location choices. We develop and apply an empirical estimation procedure that accounts for the effects of constraints on the availability of some alternatives. Following a brief description of the study area and the data used in the empirical work, we develop the model specifications and estimation algorithm. We then present the estimation results and an analyis of the sensitivity of the algorithm. The paper concludes with an assessment of the contribution of this research and proposed extensions of it.

2 The study context

The Ile-de-France region houses 11 million people (2 million in Paris) and 5.1 million jobs within a space of 12,000 sq. km. Careful preservation of the historic central city of Paris and its many landmarks has had the effect of making Paris a major tourist mecca for the globe, largely by restricting supply of new housing and office space in the historic core, and protecting historic buildings from alteration or replacement. As a consequence, real estate prices are high in the center, and lower-income housing is concentrated in the suburbs, principally in the eastern portion of the region. This is an important motivation for the focus of this paper, since the desirability of central Paris, coupled with strong supply constraints, provides a prime example of the kind of constrained market we wish to analyze. Most of the growth in population and employment are outside the core of the city, fueling rapid suburbanization and growth in travel. The regional express roadway network contains 4,500 lane-km, and despite the traditional rush-hour traffic jams, the average duration of a car trip is still only 19 minutes (EGT, 2001). With a ubiquitous metro system in central Paris and half of households in the city not owning a car, transit ridership is quite high. Half of commuting trips in the region in 2001 were by private cars, with 36% using public transportation and 14% using a bicycle or walking, though the transit mode share has declined 6% over the past twenty years as the region has decentralized.

Table 1 shows important differences in average housing prices by district: prices are higher in Paris, intermediate in the close suburbs and decline in the more distant suburbs, which we refer to as the inner ring and outer ring, respectively. In addition, prices are higher in the western part of the area than in the eastern part, consistent with the spatial patterns of social stratification.

Table (1): Prices by district.

Sub-region	District	Average	Standard Deviation	Minimum	Maximum
Paris	75	294,500	165,241	83,939	694,375
Close Suburbs (Inner Ring)	92 (West)	247,556	205,038	66,966	1,198,950
	93 (North)	115,709	49,055	47,876	259,163
	94 (South)	144,098	74,603	53,356	373,499
Far away suburbs (Outer Ring)	78 (West)	135,122	65,714	38,112	373,815
	91 (South)	114,826	46,740	24,719	332,338
	95 (North)	104,375	41,670	25,154	241,692
	77 (East)	91,539	37,220	18,028	253,827

Source: Author's computations from notaries' database.

The geographic units used in this analysis follow the administrative boundaries used in France. The smallest administrative unit, and the one we use as the basis for the residential location choice model is the commune, of which there are 1280 in the Ile-de-France. These roughly correspond to small cities, in terms of having local administrative control of land use decisions. Since Paris is one commune and disproportionately large compared to the remaining communes, we have used the Arrondisements in Paris to subdivide the city, resulting in 1300 zones for use in the model. The 1300 resulting geographic units used in this analysis (which we will still refer to as communes for simplicity) are grouped into 8 districts (departements). Table 2 presents the origin and destination rings and districts for the moves during 1998. Most of the moves have are within the same district, although households who move to Paris come predominantly from outside Ile-de-France, following a classical migration staging pattern. The most common destination for inmigrants from outside the region is Paris, followed by commune 92 in the inner ring, which is shares many attributes with Paris.

Table (2): The distribution of moves between different rings (origin by destination).

		Origin district				
Current District		Outside	Paris	Inner Ring	Outer Ring	Total
Paris	Frequency	77.579	67.027	18.192	18.023	180.821
	Percent	42.9%	37.1%	10.1%	10.0%	30.68%
Inner Ring	Frequency	61.135	22.633	103.205	20.168	207.141
	Percent	29.5%	10.9%	49.8%	9.7%	35.15%
Outer Ring	Frequency	49.936	9.299	23.967	118.191	201.393
	Percent	24.8%	4.6%	11.9%	58.7%	34.17%
Ile de France	Frequency	188.650	98.959	145.364	156.382	589.355
	Percent	32.01%	16.79%	24.66%	26.53%	100.00%

Source: Census, 1999.

According to the 1999 Census, 8.06% of the dwellings located in Ile de France were vacant in March 1999. While this represents a reasonable estimate of the vacancy rate over a very short period of time, for our analysis we wish to approximate the availability of housing over a longer period of a year.[1] Unfortunately, we have no information on how long each dwelling was vacant before a household moved in. However, during the period between the time a household moved out and another moved in, the dwelling was vacant and could have been chosen by another household. For this reason, some fraction, denoted by λ, of the dwellings occupied in March 1999 by a household which moved in 1998 should be included in the one year supply. If we assume that dwellings are vacant, on average, during half a year ($\lambda = 0.5$), then the one-year supply is made of $409,491$ dwellings vacant in March 1999, and half the $589,355$ dwellings in which a household moved in 1998, which represents a total supply of $704,168.5$ dwellings in the region. We consider it as an upper bound for the supply, leading to a lower bound for the constraints, and we will also explore smaller values for λ.

Most of the movers (71%) are male headed. The "poor households" (that is, the 33% households in the region with the lowest per capita income, defined as household income divided by the square root of the number of persons in the household) are unevenly distributed in the region: only 26% of households living in district 78, to the west of Paris; are poor, whereas this fraction goes up to 41% in district 93, to the east. These same two districts contain the highest (38% in district 78) and the lowest (21% in district 93) concentrations of rich households. Single-person households are highly concentrated in Paris city (52% of households in Paris are single). Between 25 to 30% percent of households of all the counties have two members. The larger families represent a larger share in rural counties in outer ring. 25% percent of households have no working member, and of these, 28% percent live in Paris city. Nearly 50% of the families in outer ring have two or more workers. Foreign households are most concentrated in district 93 (19%), and are less represented in the outer ring (9%). 25% of households have a young head. They have a bigger share in Paris center and 92 (31% and 27%) and their share is uniform in other counties (23%).

[1] We observe the households which moved during one year (1998), without information on the exact date they moved. In addition, we know which county they come from, but not which municipality they come from, and we have no information on the number of households who left Ile de France in 1998.

3 Model specification

We develop in this section the specification for the model, beginning with the household residential location choice, and elaborating the basic model to address constraints. For clarity of exposition, we begin with a stronger assumption of homogeneous agents (no individual characteristics are considered), and then relax this to accommodate heterogeneous agents. We follow with the description of the iterative solution algorithm, first with the case in which the demand is known (parameters are given for the residential location choice model), and finally the unknown demand case in which we need to estimate the parameters of the location choice model. The housing price model is a hedonic regression model that incorporates measures of aggregate demand and aggregate supply in each commune. We refer to this variation of the hedonic model as a semi-hedonic model, since it allows estimation of the degree to which prices adjust to help clear imbalances in the market, by incorporation both aggregate demand and supply. We turn now to the specification of the demand model.

3.1 Household residential location choice: basic model and notation

We refer to households that make a location choice within a given year as movers. These represent the set of agents producing demand for housing. Supply of housing is considered to be the set of housing units that are available for locating households. The set of alternatives in the housing market are represented by 1300 communes, denoted by $\mathcal{J}$, with $Card\left(\mathcal{J}\right) = J$. The demand for alternative j is denoted by D_j and the supply (or capacity) of alternative j is denoted by S_j. We consider an alternative to be constrained if the demand for the alternative exceeds its supply. Individual decision makers are indexed by i, $i = 1...N$. The utility of individual i selecting alternative j is:

$$U_j^i = V_j^i + \varepsilon_j^i, \; i = 1...N, \; j \in \mathcal{J}, \tag{1}$$

where V_j^i represents the systematic component of the utility and where ε_j^i are i.i.d., with a double exponential distribution. The probability $\mathbb{P}_j^i$ that individual i prefers alternative j is given by the multinomial logit formula (See Anderson, de Palma and Thisse, 1993 or McFadden, 2001 for details):

$$\mathbb{P}_j^i = \frac{\exp\left(V_j^i\right)}{\sum\limits_{j' \in \mathcal{J}} \exp\left(V_{j'}^i\right)}, \; i = 1...N, \; j \in \mathcal{J}. \tag{2}$$

In the homogeneous case where we treat agents as though they were identical, systematic utilities and choice probabilities do not vary between individuals, so they are denoted, respectively, by V_j and $\mathbb{P}_j^i = \exp\left(V_j^i\right) / \sum\limits_{j' \in \mathcal{J}} \exp\left(V_{j'}^i\right)$. The expected demand, D_j, for alternative j is:

$$D_j = \sum_{i=1}^{N} \mathbb{P}_j^i, \; j \in \mathcal{J}. \tag{3}$$

3.2 Introducing capacity constraints

The analysis of the residential choices in Ile-de-France (base on preliminary demand estimates) shows that many alternatives (nearly one half) have greater demand than supply. We denote $D_j - S_j$ as the *excess demand*, which is positive for at least one alternative when the system is constrained. We consider below the situation where demand D_j strictly exceeds supply S_j for at least one alternative j. This initial estimate of demand, which we will refer to as the ex ante demand, is different from the choices we observe households making, because the constraints are binding and some households are forced to take an alternative which was not their first preference. We refer to the demand after constraints are imposed as the ex post demand. It is in this sense that we will refer to *equilibrium*, or *market clearing.* The time frame for the analysis is considered short-term, with supply assumed to be constant.

We say that the system is *constrained* if the ex ante demand is not equal to the ex post demand for at least one alternative. However, in order to guarantee that there exists at least one feasable allocation, we assume that contraints are not globally too severe. That is, we assume a feasibility condition that aggregate supply is sufficient to accommodate aggregate demand, resulting in a vacancy rate that is strictly positive:

$$N = \sum_{j\in\mathcal{J}} D_j < \sum_{j\in\mathcal{J}} S_j \tag{4}$$

When the system is constrained, the choices based only on preferences, or *ex ante* choices, differ from the actual allocation consistent with preferences and with capacity constraints, or *ex post* allocation. In this case, the probability that individual i is allocated to alternative j is denoted by $\tilde{\mathbb{P}}_j$ in the homogeneous case (or by $\tilde{\mathbb{P}}_j^i$ in the heterogeneous case developed later). The ex ante demand D_j defined in (3) corresponds to ex ante choices, whereas the *observed demand*, denoted by $\tilde{D}_j$, corresponds to the ex post allocation:

$$\tilde{D}_j = \begin{cases} \sum_{i=1}^{N} \tilde{\mathbb{P}}_j^i \text{ in the heterogeneous case} \\ N\tilde{\mathbb{P}}_j \text{ in the homogeneous case} \end{cases}. \tag{5}$$

Note that, when the system is not constrained, then $\tilde{\mathbb{P}}_j^i = \mathbb{P}_j^i$, $j \in \mathcal{J}$, and the observed demand is equal to the ex ante demand for all alternatives and all individuals). When the system is constrained, $\tilde{D}_j$ is bounded by S_j $(j = 1...J)$ and the constraint is binding for at least one alternative. Two situations may arise:

1. Alternative j is *unconstrained ex ante*: $D_j < S_j$. In this case, alternative j is *unconstrained ex post if* $\tilde{D}_j < S_j$ while alternative j is *constrained ex post if* $\tilde{D}_j = S_j$.
2. Alternative j is *constrained ex ante*: $D_j \geq S_j$. In this case, it can be shown that $\tilde{D}_j = S_j$ i.e. alternative j is also *constrained ex post.*

We will also define the following sets of constrained alternatives: $\mathcal{C} = \left\{ j \in \mathcal{J} \mid \tilde{D}_j = S_j \right\}$ denotes the alternatives constrained ex post and $\mathcal{C}(0) = \{ j \in \mathcal{J} \mid D_j \geq S_j \}$ denotes the set of alternatives constrained ex ante.

Similarly, we define the following sets of unconstrained alternatives: $\bar{\mathcal{C}} = \left\{ j \in \mathcal{J} \mid \tilde{D}_j < S_j \right\}$ denotes the alternatives unconstrained ex post and $\bar{\mathcal{C}}(0) = \{ j \in \mathcal{J} \mid D_j < S_j \}$ denotes the alternatives unconstrained ex ante.

These four sets determine two partitions of $\mathcal{J}$ (since $\mathcal{C} \cap \bar{\mathcal{C}} = \mathcal{C}(0) \cap \bar{\mathcal{C}}(0) = \varnothing$ and $\mathcal{C} \cup \bar{\mathcal{C}} = \mathcal{C}(0) \cup \bar{\mathcal{C}}(0) = \mathcal{J}$). When, for at least one alternative j, the (ex ante) demand D_j is larger than the capacity S_j the actual allocation is the result of a complex mechanism which depends on the priority rules developed below.

Constraints have two consequences: first, if alternative j is constrained ex ante ($j \in \mathcal{C}(0)$), a fraction of the individuals who would select alternative j without capacity constraints (D_j) must instead choose a less desirable alternative k. Second, due to the first consequence, the excess demand generated in alternatives constrained ex ante is reallocated to alternatives which were not constrained ex ante. Therefore, the observed demand is larger than (or equal to) the ex ante demand in all alternatives unconstrained ex ante. Some of the alternatives which are not constrained ex ante may be constrained ex post, due to the reallocation of excess demand.

We now introduce a first assumption: *free allocation,* which means that if an individual prefers an alternative j, which is unconstrained ex post, he can be sure to be allocated to it. However, he may be denied access to his preferred choice j if it is constrained ex post, and be forced to choose another alternative (that may be constrained or unconstrained ex ante).

Assumption 1 (Free allocation) *.Let $j \in \bar{\mathcal{C}}$. Then*

$$\mathbb{P}\left(\{i \text{ allocated to } j \mid i \text{ prefers } j\}\right) = 1, \forall\ i = 1...N.$$

Assumption 1 implies that the IIA property (specific to the MNL model) is valid for the unconstrained alternative, that is, for two alternatives unconstrained ex post, the ratio of (ex post) allocation probabilities is equal to the ratio of (ex ante) choice probabilities.

This assumption implies the IIA properties in the following sense:

Lemma 1 (Alternative unconstrained ex post) *If Assumption 1 holds and if alternative j and k are unconstrainted ex post, then the allocation probabilities satisfy the IIA property:*

$$\frac{\tilde{\mathbb{P}}_j^i}{\tilde{\mathbb{P}}_k^i} = \frac{\mathbb{P}_j^i}{\mathbb{P}_k^i},\ \forall\ i = 1...N,\ \forall\ j,k \in \bar{\mathcal{C}}. \tag{6}$$

Moreover:

$$\tilde{\mathbb{P}}_j^i \geq \mathbb{P}_j^i \text{ and } \tilde{D}_j \geq D_j,\ \forall\ i = 1...N,\ \forall\ j \in \bar{\mathcal{C}}.$$

The equality (6) states that the allocation probabilities satisfy the IIA property for unconstrained alternatived. The inequalities state that the probability that individual i is allocated to j is larger than the probability that he prefers j and the observed demand addressed to j is larger than the ex ante demand if j is unconstrained. As an immediate consequence of Assumption 1, we have: $\mathcal{C}(0) \subset \mathcal{C}$ and $\bar{\mathcal{C}} \subset \bar{\mathcal{C}}(0)$. That is: if alternative j is constrained ex ante, then it is also constrained ex post. If alternative j is unconstrained ex post, it is also unconstrained ex ante.

Proposition (1) states that the individual ratio $\tilde{\mathbb{P}}_j^i/\mathbb{P}_j^i$ of the actual allocation probability to the choice probability is the *same accross all unconstrained alternatives*, since equality (6) can be rewritten as:

$$\frac{\tilde{\mathbb{P}}_j^i}{\mathbb{P}_j^i} = \frac{\tilde{\mathbb{P}}_k^i}{\mathbb{P}_k^i} \stackrel{def}{=} \Omega^i,\ \forall\ i = 1...N,\ \forall\ j,k \in \bar{\mathcal{C}},$$

where the common value of this ratio is denoted by Ω^i. We interpret Ω^i as the *individual allocation ratio* for individual i. The computation of Ω^i cannot be done before the allocation rules are defined for the ex post constrained alternatives.

We now introduce a second assumption concerning the allocation in alternatives constrained ex post. We assume that if an individual i has a stronger preference (ex post) for constrained alternative j than another individual i', in the sense that his choice probability is larger, he will also have proportionaly more opportunity to be allocated ex post to this alternative j in the following sense:

Assumption 2 (No priority rule) *If $j \in \mathcal{C}$, the individual allocation ratio of alternative j, constrained ex post is the same for all individuals:*

$$\frac{\tilde{\mathbb{P}}_j^i}{\mathbb{P}_j^i} = \frac{\tilde{\mathbb{P}}_j^{i'}}{\mathbb{P}_j^{i'}} \stackrel{def}{=} \Phi_j,\ \forall\ i,i' = 1...N,\ \forall\ j \in \mathcal{C}.$$

Assumption (2), which states that the individual allocation ratio is the same for each individual ($i = 1...N$), and denoted by Φ_j, can also be interpreteted as a fairness criterion.

This assumption suffices to determine the queue discipline for the alternatives which are constrained ex post. Indeed, it is straighforward to show that the common value (accross individuals) of the allocation ratio $\tilde{\mathbb{P}}_j^i/\mathbb{P}_j^i$ is equal to the relative supply (measured by S_j/D_j):

Lemma 2 (Alternative constrained ex post) *Consider an individual i. If Assumption 2 holds and alternative j is constrained ex post, the common value of the individual allocation ratio is equal to the relative supply, i.e.:*

$$\frac{\tilde{\mathbb{P}}_j^i}{\mathbb{P}_j^i} = \Phi_j = \frac{S_j}{D_j} < 1,\ \forall\ i = 1...N,\ \forall\ j \in \mathcal{C}.$$

These two assumptions are essentially sufficient to solve for an equilibrium. It is necessary, also to eliminate extreme preferences (which are not observed in our data). The interested reader is refered to the footnote[2]. We will provide the solution for the homogeneous case and sketch the solution for the heterogenous case. Assumptions 1 and 2 hold, throughout the rest of the paper.

3.3 Equilibrium solution with capacity constraints

The solution requires the computation of two unknowns: (1) which alternatives are constrained ex post and (2) the value of Ω^i. We assume for the moment that we know which alternatives are constrained ex post and will return to its computation subsequently.

Recall that, for the unconstrained alternatives, the allocation ratio of alternative j defined as $\tilde{\mathbb{P}}_j^i/\mathbb{P}_j^i$ is independent of the alternative and denoted by Ω^i For the unconstrained alternative ($S_j < D_j$), the value of Ω^i is larger than 1 since some constrained individual are reallocated to the unconstrained alterantives. The value of the ratio is given by:

Lemma 3 (Individual allocation ratio) *Consider an individual i. The availability ratio is the same for each alternative unconstrained ex post and is given by :*

$$\frac{\tilde{\mathbb{P}}_j^i}{\mathbb{P}_j^i} = \Omega^i = \frac{1 - \sum\limits_{j\in\mathcal{C}} \frac{S_j}{D_j}\mathbb{P}_j^i}{\sum\limits_{k\in\bar{\mathcal{C}}} \mathbb{P}_k^i} > 1, \forall j \in \bar{\mathcal{C}}.$$

Lemma 3 shows that the individual allocation ratio $\Omega^i = \tilde{\mathbb{P}}_j^i/\mathbb{P}_j^i,\ \forall\ j \in \bar{\mathcal{C}}$ is uniquely defined as a function of the set $\mathcal{C}$ of alternatives constrained ex post. The ex ante demand is given by (3), while the ex post demand is defined by: $\tilde{D}_j = \sum_{i=1}^N \tilde{\mathbb{P}}_j^i$ (see equation (5)). The aggregate allocation ratio, $\bar{\Omega}_j$, is defined as a weighted average of individual allocation ratios Ω^i:

$$\bar{\Omega}_j \overset{def}{=} \frac{\sum_{i=1}^N \Omega^i \mathbb{P}_j^i}{\sum_{i=1}^N \mathbb{P}_j^i} > 1, \forall\ j \in \mathcal{J}. \tag{7}$$

Note that $\bar{\Omega}_j > 1$ if $j \in \bar{\mathcal{C}}$ (see Lemma 3). Collecting the previous results, we have:

Lemma 4 (Aggregate allocation ratio) *Consider an alternative unconstrained ex post $j \in \bar{\mathcal{C}}$. The alternative-specific aggregate allocation ratio is :*

$$\bar{\Omega}_j = \frac{\sum_{i=1}^N \Omega^i \mathbb{P}_j^i}{\sum_{i=1}^N \mathbb{P}_j^i} = \frac{\tilde{D}_j}{D_j},\ j \in \bar{\mathcal{C}},$$

where Ω^i are given by Lemma 3.

[2] We need also to eliminate extreme preference. One supplementary assumption is the "no preference reversal" : no individual in the population has a strong preference for one alternative, which is constrained ex post but not ex ante. Moreover, we need to assume that $\sum_{j\in\mathcal{C}\cap\mathcal{C}(0)} \mathbb{P}_j^i$ is not too large.

Note that the value of the individual allocation ratio Ω^i can be computed once the set $\mathcal{C}$ is known (see Lemma 3). Conversely, assume that the ratios Ω^i are known. An alternative j is constrained ex post if and only if the demand allocated to alternative j is contrained, i.e. if $\sum_{i=1}^{N} \Omega^i \mathbb{P}_j^i$ is larger than (or equal to) S_j. Therefore, an alternative j is constrained ex post if and only is:

$$\bar{\Omega}_j = \frac{\sum_{i=1}^{N} \Omega^i \mathbb{P}_j^i}{\sum_{i=1}^{N} \mathbb{P}_j^i} \geq \frac{S_j}{D_j}. \tag{8}$$

Note that if the alternative j is constrained ex ante, then $S_j/D_j > 1$, and the condition (8) is always satisfied (ex ante constraint implies ex post constraint). We refer the reader to de Palma, Picard and Waddell (2006) for the proofs of existence and uniqueness of a global solution $(\Omega^i, \mathcal{C})$.

Below, we provide a method for computing the global solution, either when the demand is known and the parameters are given, or when it is unknown and the parameters need to be estimated in tandem with the iterative procedure.

3.4 Computational method

3.4.1 Allocation Probabilities

The allocation probabilities are given by different expressions according to whether the alternatives are constrained or not ex post. The allocation probabilities can still be written as a multinomial logit model, but with a additional term, or "correction factor", $\ln(\pi_j^i)$, which expresses the allocation ratio.

Theorem 1 (Allocation probabilities) *If Assumptions 1 and 2 hold, the allocation probabilities are given by the adjusted MNL formula:*

$$\tilde{\mathbb{P}}_j^i = \frac{\exp\left(\tilde{V}_j^i\right)}{\sum_{k\in\mathcal{J}} \exp\left(\tilde{V}_k^i\right)}, \text{ with} \tag{9a}$$

$$\tilde{V}_j^i = V_j^i + \ln(\pi_j^i), \text{ with} \tag{9b}$$

$$\pi_j^i = \begin{cases} \frac{S_j}{D_j} & \text{if } j \in \mathcal{C} \\ \Omega^i & \text{if } j \in \bar{\mathcal{C}} \end{cases}. \tag{9c}$$

Proof. *The denominator in (9a) is*

$$\sum_{k\in\mathcal{J}} \exp\left(\tilde{V}_k^i\right) = \sum_{k\in\mathcal{C}} \frac{S_k}{D_k} \exp\left(V_k^i\right) + \Omega^i \times \sum_{k\in\bar{\mathcal{C}}} \exp\left(V_k^i\right)$$

$$= \sum_{j\in\mathcal{J}} \exp\left(V_j^i\right) \times \left[\sum_{k\in\mathcal{C}} \frac{S_k}{D_k} \frac{\exp\left(V_k^i\right)}{\sum_{j\in\mathcal{J}} \exp\left(V_j^i\right)} + \Omega^i \sum_{k\in\bar{\mathcal{C}}} \frac{\exp\left(V_k^i\right)}{\sum_{j\in\mathcal{J}} \exp\left(V_j^i\right)}\right]$$

$$= \sum_{j\in\mathcal{J}} \exp\left(V_j^i\right) \times \left[\sum_{k\in\mathcal{C}} \frac{S_k}{D_k} \mathbb{P}_k^i + \Omega^i \sum_{k\in\bar{\mathcal{C}}} \mathbb{P}_k^i\right]$$

$$= \sum_{j\in\mathcal{J}} \exp\left(V_j^i\right) \times \sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j^i = \sum_{j\in\mathcal{J}} \exp\left(V_j^i\right).$$

Therefore:

$$\frac{\exp\left(\tilde{V}_j^i\right)}{\sum_{k\in\mathcal{J}}\exp\left(\tilde{V}_k^i\right)} = \pi_j^i \times \frac{\exp\left(V_j^i\right)}{\sum_{j\in\mathcal{J}}\exp\left(V_j^i\right)} = \tilde{\mathbb{P}}_j^i.$$

■

The reader should still keep in mind that the difficult part of this approach (and of any approach dealing with constraints) is the determination of the alternatives which are constrained ex post.

We first describe the iterative procedure to find the allocation probabilities and the ex post constrained alternatives, when the demand is known, as would be the case once the model is estimated and is being used to make predictions. This simplifies the initial exposition.

3.4.2 Iterative procedure with known demand

The formulas are given by Lemma 3 for the individual allocation ratio Ω^i, by Lemma 4 for the aggregate availability ratio $\bar{\Omega}_j$ and by condition (8) for the set $\mathcal{C}$ of alternatives j constrained ex post. The following algorithm allows computation of Ω^i, $\bar{\Omega}_j$ and $\mathcal{C}$ with less than J iterations.

Algorithm 1: known demand

1. Check that aggregate demand is smaller than aggregate supply (see condition 4)).
2. Iteration $l = 0$ (initialization): set $\Omega^i = 1$; compute the set $\mathcal{C}(0) = \{j \in \mathcal{J} \mid S_j \leq D_j\}$ of alternatives constrained initially (that is at iteration zero). $\bar{\mathcal{C}}(0) = \mathcal{J} \setminus \mathcal{C}(0)$.
3. Compute the individual allocation ratios $\Omega^i(0) = \frac{1-\sum_{j\in\mathcal{C}(0)}\frac{S_j}{D_j}\mathbb{P}_j^i}{1-\sum_{j\in\mathcal{C}(0)}\mathbb{P}_k^i}$, using Lemma 3).
4. Compute the alternative-specific allocation ratios using Equation (7): $\bar{\Omega}_j(0) = \frac{\sum_{i=1}^{N}\Omega^i(0)\mathbb{P}_j^i}{\sum_i \mathbb{P}_j^i}$
5. Update iteration: $l \to l+1$. Update the constrained choice set using condition (8):
$$\mathcal{C}(l+1) = \left\{j \in \mathcal{J} \mid S_j \leq \bar{\Omega}_j(l)\, D_j = \sum_{i=1}^{N}\Omega^i(l)\,\mathbb{P}_j^i\right\}$$
6. Update $\Omega^i(l+1) = \frac{1-\sum_{j\in\mathcal{C}(l+1)}\frac{S_j}{D_j}\mathbb{P}_j^i}{\sum_{k\in\bar{\mathcal{C}}(l+1)}\mathbb{P}_k^i}$
7. Update $\bar{\Omega}_j(l+1) = \frac{\sum_{i=1}^{N}\Omega^i(l+1)\mathbb{P}_j^i}{\sum_{i=1}^{N}\mathbb{P}_j^i}$
8. Stop at iteration $l+1$ if $\mathcal{C}(l+1) = \mathcal{C}(l)$ (and $\Omega^i(l+1) = \Omega^i(l)$ for all $i = 1...N$), else go to 5.
9. Compute the correction factor $\ln\left(\pi_j^i\right)$ using Theorem 1.

We have shown, in the homogenous case, that this algorithm converges. Simulation experiments suggest that it also converge in the heterogenous case.

3.4.3 Iterative procedure with unknown demand

When the demand is unknown, we need to jointly estimate the parameters of the model and to determine the allocated demand, given a set of parameters. The first algorithm determines the alternatives contrained ex post, the allocated demand and the correction factors given the demand parameters. The second algorithm has the same output as the first algorithm, but relies on a demand system which needs to be estimated.

The approach in this second algorithm is to iterate between an estimation module, which incorporates the correcting factor $\ln\left(\pi_j^i\right)$, computed during the previous iteration and algorithm 1, which in turn relies on the previous iteration estimated demand. The estimation module and algoritm 1 are embedded in a double loop. The outer loop (estimation) is indexed by l' and the inner loop (algorithm 1) is indexed by l.

We require additional notations: $\hat{D}_j\left(l'\right)$, $\hat{\mathbb{P}}_j^i\left(l'\right)$ and $\ln\left[\pi_j^i\left(l'\right)\right]$ represent the estimation of the demand, the choice probabilities and the correction factor at iteration l'. Moreover, $\mathcal{C}\left(l', l\right)$ denotes the set of alternatives constrained ex post at iteration $\left(l', l\right)$, where l' corresponds to the demand estimation and l corresponds to the constraints. We define $\Omega^i\left(l', l\right)$ and $\bar{\Omega}_j\left(l', l\right)$ in a similar way. Below, we sketch the structure of this second algorithm.

Algorithm 2: unknown demand

1. Check that the aggregate demand is smaller than the aggregate supply (Assumption 4).
2. Iteration $l' = 0$ (initialization of demand of the estimation module): estimate a MNL model assuming initially no constraints (i.e. $\Omega^i = \bar{\Omega}_j = \pi_j^i = 1$). This gives the estimated choice probabilities $\hat{\mathbb{P}}_j^i(0)$ and the estimated demand $\hat{D}_j(0) = \sum_{i=1}^N \hat{\mathbb{P}}_j^i(0)$. The correcting factors $\pi_j^i(0)$ are initialized to 1.

   (a) Iteration $l = 0$ (initialization of the constraints): compute the set

   $$\mathcal{C}(0,0) = \left\{ j \in \mathcal{J} \mid S_j \leq \hat{D}_j(0) \right\}$$

   of alternatives constrained ex ante for the initial estimated demand, using condition (8).

   (b) Compute the individual allocation ratio, using Lemma 3):

   $$\Omega^i(0,0) = \left(1 - \sum_{j \in \mathcal{C}(0,0)} \frac{S_j}{\hat{D}_j(0)} \hat{\mathbb{P}}_j^i(0)\right) \Bigg/ \left(1 - \sum_{j \in \mathcal{C}(0,0)} \hat{\mathbb{P}}_j^i(0)\right).$$

   (c) Compute the alternative-specific allocation ratio using Equation (7):

   $$\bar{\Omega}_j(0,0) = \left(\sum_{i=1}^N \Omega^i(0,0)\, \hat{\mathbb{P}}_j^i(0)\right) \Bigg/ \left(\sum_{i=1}^N \hat{\mathbb{P}}_j^i(0)\right).$$

   (d) Update iteration for algorithm 1: $l \to l+1$: update the set of alternatives constrained using condition (8):

   $$\mathcal{C}(0,l+1) = \left\{ j \in \mathcal{J} \mid S_j \leq \bar{\Omega}_j(0,l)\, \hat{D}_j(0) = \sum_{i=1}^N \Omega^i(0,l)\, \hat{\mathbb{P}}_j^i(0) \right\}$$

   (e) Update $\Omega^i(0,l+1) = \left(1 - \sum_{j \in \mathcal{C}(0,l+1)} \frac{S_j}{\hat{D}_j(0)} \hat{\mathbb{P}}_j^i(0)\right) \Bigg/ \left(1 - \sum_{j \in \mathcal{C}(0,l+1)} \hat{\mathbb{P}}_j^i(0)\right)$

(f) Update $\bar{\Omega}_j(0,l+1) = \left(\sum_{i=1}^{N} \Omega^i(0,l+1)\,\hat{\mathbb{P}}_j^i(0)\right) \Big/ \left(\sum_{i=1}^{N} \hat{\mathbb{P}}_j^i(0)\right)$.

(g) Stop at iteration $l+1$ if $\mathcal{C}(0,l+1) = \mathcal{C}(0,l)$ (and therefore $\Omega(0,l+1) = \Omega(0,l)$), else go to step d.

(h) When convergence is attained ($l=\infty$), we have $\mathcal{C}(0) \overset{def}{=} \left\{ j \in \mathcal{J} \mid S_j \leq \bar{\Omega}_j(0,\infty)\,\hat{D}_j(0) \right\}$, $\Omega^i(0) \overset{def}{=} \Omega^i(0,\infty)$, and $\bar{\Omega}_j(0) \overset{def}{=} \bar{\Omega}_j(0,\infty)$

3. Update iteration of the estimation module: $l' \to l'+1$: update $\pi_j^i(l'+1) = \begin{cases} \frac{S_j}{\hat{D}_j(l')} & \text{if } j \in \mathcal{C}(l') \\ \Omega^i(l') & \text{if } j \in \bar{\mathcal{C}}(l') \end{cases}$.

4. Stop if $\pi_j^i(l'+1) = \pi_j^i(l')$, the solution is: $D_j = \hat{D}_j(l') = \hat{D}_j(l'+1)$, $\mathcal{C} = \mathcal{C}(l') = \mathcal{C}(l'+1)$, $\Omega^i = \Omega^i(l') = \Omega^i(l'+1)$ and $\bar{\Omega}_j = \bar{\Omega}_j(l') = \bar{\Omega}_j(l'+1)$; else go to step 5.

5. Update the MNL estimates with updated $\pi_j^i(l'+1)$ among the explanatory variables in $\tilde{V}_j^i$. This gives $\tilde{\mathbb{P}}_j^i(l'+1) = \frac{\exp\left(\tilde{V}_j^i(l'+1)\right)}{\sum\limits_{k\in\mathcal{J}} \exp\left(\tilde{V}_k^i(l'+1)\right)}$ and $\hat{\mathbb{P}}_j^i(l'+1) = \frac{\tilde{\mathbb{P}}_j^i(l'+1)}{\pi_j^i(l'+1)}$ and $\hat{D}_j(l'+1) = \sum_{i=1}^{N} \hat{\mathbb{P}}_j^i(l'+1)$

(a) Iteration $l=0$ (initialization): Compute the set $\mathcal{C}(l'+1,0) = \left\{ j \in \mathcal{J} \mid S_j \leq \hat{D}_j(l'+1) \right\}$ of alternatives constrained ex ante for demand at step $(l'+1)$

(b) Compute the individual allocation ratio
$\Omega^i(l'+1,0) = \left(1 - \sum\limits_{j\in\mathcal{C}(l'+1,0)} \frac{S_j}{\hat{D}_j(l'+1)} \hat{\mathbb{P}}_j^i(l'+1)\right) \Big/ \left(1 - \sum\limits_{j\in\mathcal{C}(l'+1,0)} \hat{\mathbb{P}}_j^i(l'+1)\right)$

(c) Compute the alternative-specific allocation ratio $\bar{\Omega}_j(l'+1,0) = \frac{\sum_{i=1}^{N} \Omega^i(l'+1,0)\hat{\mathbb{P}}_j^i(l'+1)}{\sum_{i=1}^{N} \hat{\mathbb{P}}_j^i(l'+1)}$

(d) Update iteration of Algorithm 1: $l \to l+1$: Update
$\mathcal{C}(l'+1,l+1) = \left\{ j \in \mathcal{J} \mid S_j \leq \bar{\Omega}_j(l'+1,l)\,\hat{D}_j(l'+1) = \sum_{i=1}^{N} \Omega^i(l'+1,l)\,\hat{\mathbb{P}}_j^i(l'+1) \right\}$

(e) Update $\Omega^i(l'+1,l+1) = \left(1 - \sum\limits_{j\in\mathcal{C}(l'+1,l+1)} \frac{S_j}{\hat{D}_j(l'+1)} \hat{\mathbb{P}}_j^i(l'+1)\right) \Big/ \left(1 - \sum\limits_{j\in\mathcal{C}(l'+1,l+1)} \hat{\mathbb{P}}_j^i(l'+1)\right)$

(f) Update $\bar{\Omega}_j(l'+1,l+1) = \left(\sum_{i=1}^{N} \Omega^i(l'+1,l+1)\,\hat{\mathbb{P}}_j^i(l'+1)\right) \Big/ \left(\sum_{i=1}^{N} \hat{\mathbb{P}}_j^i(l'+1)\right)$

(g) Stop Algorithm 1 at iteration $l+1$ if $\mathcal{C}(l'+1,l+1) = \mathcal{C}(l'+1,l)$ (and therefore $\Omega^i(l'+1,l+1) = \Omega^i(l'+1,l)$ and $\bar{\Omega}_j(l'+1,l+1) = \bar{\Omega}_j(l'+1,l)$), else go to step d.

(h) When convergence of Algorithm 1 is attained ($l=\infty$), we have
$\mathcal{C}(l'+1) = \left\{ j \in \mathcal{J} \mid S_j \leq \bar{\Omega}_j(l'+1,\infty)\,\hat{D}_j(l'+1) \right\}$, $\Omega^i(l'+1) = \Omega^i(l'+1,\infty)$, and $\bar{\Omega}_j(l'+1) = \bar{\Omega}_j(l'+1,\infty)$.

6. Go to step 3.

In the next section we provide some numerical results to illustrate this method. Note that we have not shown (even with known demand) that algorithm 2 converges, even if numerical experiments with real data strongly susggest that this is the case.

4 Empirical results

The data for household location model come from the 1999 census, for which we were able to access household-level data for the entire population, allowing us to test the sensitivity of the estimation to using a range of sampling rates from the population. The analysis focuses on "recent movers": households who settled or moved to the region recently, that is during year 1998. Among the 4,510,369 households living in the study area in March 1999, 589,355 moved into or within the region during year 1998. The housing price data used in the model come from the "base de données des Notaires" and contains the average price of single-family dwellings sold in the commune in 1999. Note that the single-family dwellings are usually larger and are on larger lots in the suburbs, so the differences between Paris and the suburbs are smaller than differences in price per square meter. The attributes of the alternatives have been computed from different sources, mainly drawing on data from the IAURIF metropolitan planning agency. See [de Palma *et al.*2006] for details.

4.1 Housing price

We estimate a semi-hedonic regression model to predict housing prices, and use these predicted prices in the residential location choice model. To reflect the economic endogeneity between prices, demand and supply, we put measures of housing supply (considered static in the short-run period we are modeling), and demand, which is computed from the residential location choice model, on the right hand side of the hedonic regression. This specification, linking the hedonic regression and the residential location choice model via aggregate demand, allows us to empirically test the degree to which price adjustments based on the varying relationship of demand and supply serve to clear the market. As is the norm in the hedonic literature, we specify the model as semi-log, with the natural log of housing prices as the dependent variable.

The estimated coefficients for housing price model are presented in table 3. The R^2 for the model is 0.53, which is rather high considering that we are using only the average sales price in each commune and therefore have no attributes of individual houses entered in the model. The only effects are commune characteristics and the aggregate market conditions of demand and supply. We obtain the expected signs for demand and supply but they are not exactly opposed. A purely structural equation (results not reported here, available on request) with only supply and demand gives coefficients exactly opposed, which means that the price only depends on the supply/demand ratio, and not separately on supply and demand. This result does not hold ceteris paribus.

A decrease in average travel time significantly increases the price: 10 minutes less imply a 2.8% increase in housing price. The price is very sensitive to socio-economic structure of the commune: a 10% increase in the proportion of one-member households is associated with a 50% increase of the price. Such an increase for the proportion of two-members households corresponds to a 19% increase of the price. Similarly, the fraction of households with no or only one working member has a positive effect on the price. Surprisingly, the fraction of foreign households has a positive effect on price. We should notice however, that the data do not distinguish the nationality of the foreigners, and make no difference between OECD countries and third world ones, and we are controlling for the income of the commune, which shows negative and highly significant effects of the proportion of low and intermediate income families on the price.

Table (3): Housing Price Estimation Results

Variable	Coefficient	Standard error	t-statistic	p-value
Intercept	11.02668	0.12800	86.14	<.0001
Log(Supply)	-0.04791	0.02466	-1.94	0.0522
Log(Demand)	0.09918	0.02244	4.42	<.0001
Average travel time from j to work (minutes)	-0.00280	0.00085119	-3.28	0.0011
% households with 1 member	5.09136	0.37884	13.44	<.0001
% households with 2 members	1.87960	0.34135	5.51	<.0001
% households with no working member	1.25241	0.30954	4.05	<.0001
% households with 1 working member	0.82300	0.33762	2.44	0.0149
% poor households	-6.63187	0.50316	-13.18	<.0001
% households with medium income	-4.54311	0.33102	-13.72	<.0001
% households with a foreign head	1.58406	0.36279	4.37	<.0001

4.2 Location choice without capacity constraints

Table 4 contains the results of the residential location choice model estimation, estimated on a 100% sample of households, and assuming no availability constraints. With a pseudo-R^2 of 22% this model has a moderate explanatory power at the individual level. Later on, we will explore the validity of this model at the aggregate level (see Section 4.4.4). We find a very significant effect of the "same district as before" variable, confirming (as expected) a strong preference of households to move in the same district or neighbourhood in which they lived befor the move. Testing the effect of the distance from last residence may be interesting but it was not possible with our available data, which did not contain information on residence location more detailed than the commune. The Paris dummy variable has a negative coefficient, implying that, ceteris paribus, the households who live in Paris and decide to move have a slightly higher probability of relocating to a different district than do residents living in other districts. Note that this is consistent with the intra-metropolitan migration patterns shown in Table 2, and with general expectations that households moving into the region, and new households formed within the region locate initially within Paris, and may relocate to suburban neighbourhoods later. Note, however, that some of the other variables in the model, such as better accessibility in Paris, tend to have effects that at least partially offset this suburbanization preference, while others, such as housing prices, tend to reinforce it.

As expected, housing price has a negative effect on location preference for a commune. This effect increases with the age of the household head and decreases with as the household income increases. The older heads of households are more sensitive to price and the richer households are less sensitive to it. Since price is entered using three variables to capture average effects as well as interactions with age and income, the combined effects are complex. We note that the average price effect as well as the age and income interactions, all have expected signs. However, for a small subset of the population, namely very young and very rich households, the net price effect from the interaction of these three coefficients would be predicted by this model to show a slight positive preference for higher prices in communes where they the neighbouring households are in the same socio-economic category and which have more amenities. The relative sensitivity to price is as we would expect, though the potential for a small positive preference for higher prices for this specific subpopulation and sample of locations is likely to be due to some amenities that are not accounted for in the model, rather than an actual preference to may more for housing, ceteris paribus.

Increase of the average travel time by public transit decreases the preference of households headed by a woman, though this effect is insignificant for male-headed households.

Table (4): Residential Location Choice, 100% sample, no constraints

Variable	Coefficient	Std error	t-stat.	p-value
Same district as before move	2.5515	0.004194	608.31	<.0001
Paris	-0.3386	0.0123	-27.62	<.0001
Log(Price)	-1.7243	0.0471	-36.62	<.0001
Log(Price)* (Age-20)/10	-0.0639	0.002073	-30.81	<.0001
Log(Price)* Log(Income)	0.1774	0.004661	38.06	<.0001
Number Railway stations	-0.0137	0.001186	-11.53	<.0001
Number Subway stations	0.007164	0.000523	13.69	<.0001
Average travel time from j, commuting (TC) [100']	-0.0026	0.0212	-0.12	0.9023
TC*(Dummy female) [100']	-0.6129	0.0315	-19.48	<.0001
Average travel time from j, by private car (VP) [100']	0.5651	0.0349	16.18	<.0001
Distance to highway [km]	-0.002822	2.749E-4	-10.27	<.0001
% households with 1 member * 1 member in h	2.5965	0.0377	68.91	<.0001
% households with 2 members* 2 members in h	0.9065	0.1360	6.66	0.0022
% households with 3+ members* 3+ member in h	3.2398	0.0349	92.80	<.0001
% hh with no working member * no working member in h	6.1624	0.1005	61.34	<.0001
% hh with 1 working member * 1 working member in h	0.1497	0.0631	2.37	0.0177
% hh with 2+ working member * 2+ working member in h	0.7512	0.0440	17.06	<.0001
% hh with a young head * young head in h	4.8530	0.0478	101.45	<.0001
% poor households * h poor	0.7796	0.0499	15.62	0.0240
% households with a foreign head * foreign head in h	5.9707	0.0719	83.00	<.0001
% households with a foreign head * French head in h	-2.8506	0.0429	-66.39	<.0001
Density (Population/Surface) [1000 persons/km]	-0.00519	0.000456	-11.37	<.0001
Log(Population)	0.0909	0.002264	40.14	<.0001
% change in population, 1990 to 1999	0.0793	0.007151	11.09	<.0001

The number of subway (metro) stations in a commune increases the probability of location but the number of railway stations decreases it, after accounting for transit accessibility and other effects. These results may reflect the relative effects of positive and negative externalities associated with subway stations and railway stations. Metro stations are more likely than railway stations to be located within clusters of shopping and service employment or adjacent to major cultural attractions, and railway stations are larger and may be more likely to have negative localized externalities on the immediate neighbourhood, such as traffic, noise, and possibly petty crime. The average travel time by private car and the distance to the highway have a negative effect on the preference for a commune, as expected.

The estimated coefficients corresponding to the socio-economic structure of the commune show a general preference of the households to live with neighbors of the same social category. This preference is very strong for households without workers, or with a foreign or young head. The households with one worker are less sensitive to the concentration of similar households. Households of French origin tend to avoid locations in which there are higher concentrations of foreign households. The coefficients for the percentage of young head households and the total number of jobs are insignificant. Households prefer more populated but less dense communes, and the communes that have absorbed more population during the 1990-99 period attract still more households.

Adding the residuals of the price equation as an explanatory variable, the estimated coefficients changes trivially and the coefficient of this new variable is not at all significant. This result confirms that housing price is *not endogenous* (in the econometric sense) with regard to the location choice model. In other words, the variables used in these two models fully explain the correlation between prices and location choice.

4.3 Sensitivity analysis

In order to determine how robust these results are with respect to the sampling procedures we are using, we present below the results of sensitivity analyses to test for the effects of different sampling rates for households, and different sizes of sampled alternatives.

4.3.1 Sampling households

Since we have access to the full population of the region at a household level, we can estimate the model on the full population rather than a sample of households. As this is a very unusual circumstance, we wish to learn whether the results deteriorate significantly as we reduce the sample size. Table 5 presents the estimation results of the location choice model with a randomly sampled choice set of 8 alternatives (including the chosen alternative, and using importance sampling), on different household samples, randomly selected.

Table (5): Impact of sampling ratio on Residential Location Choice Estimation Results

Variable \ Household Sampling Rate	100%	20%	2%	1%
Same district as before move	2.5515‡	2.5461‡	2.5873‡	2.6575‡
Paris	-0.3386‡	-0.299‡	-0.3402‡	-0.3077†
Log(Price)	-1.7243‡	-1.7285‡	-1.0927‡	-1.312‡
Log(Price)* (Age-20)/10	-0.0639‡	-0.0653‡	-0.0585‡	-0.0501†
Log(Price)* Log(Income)	0.1774‡	0.1783‡	0.1155‡	0.1334‡
Number Railway stations	-0.0137‡	-0.0129‡	-0.0155*	-0.0241*
Number Subway stations [10]	0.007164‡	0.07070‡	0.06413*	0.09369*
Average travel time from j, commuting (TC) [100']	-0.0026	0.0561	0.129	0.1014
TC*(Dummy female) [100']	-0.6129‡	-0.684‡	-0.652‡	-1.12‡
Average travel time from j, by private car (VP) [100']	0.5651‡	-1.391†	7.364‡	8.389†
Distance to highway [km]	-2.82E-3‡	-3.39E-3‡	-3.81E-3*	-6.34E-3†
% households with 1 member * 1 member in h	2.5965‡	2.6327‡	2.3846‡	2.4915‡
% households with 2 members* 2 members in h	0.9065‡	0.9366‡	0.9857	0.3499
% households with 3+ members* 3+ member in h	3.2398‡	3.2437‡	3.4122‡	3.477‡
% hh with no working memb. * no working memb. in h	6.1624‡	6.1790‡	5.2497‡	5.084‡
% hh with 1 working memb. * 1 working memb. in h	0.1497†	0.3384†	0.3152	-0.0793
% hh with 2+ working memb. * 2+ working memb. in h	0.7512‡	0.7132‡	0.7320†	1.0542†
% hh with a young head * young head in h	4.8530‡	4.7947‡	4.7072‡	4.665‡
% poor households * h poor	0.7796‡	0.3853†	0.9607‡	0.0847
% households with a foreign head * foreign head in h	5.9707‡	6.2094‡	5.5039‡	4.9780‡
% households with a foreign head * French head in h	-2.8506‡	-2.7905‡	-3.2432‡	-3.098‡
Density (Population/Surface) [1000 persons/km]	-5.191E-3‡	-4.62E-3‡	-1.95E-3	-6.41E-3
Log(Population)	0.0909‡	0.0931‡	0.1001‡	0.1209‡
% change in population, 1990 to 1999	0.0793‡	0.0931‡	0.115†	0.00097
Pseudo R^2	22.2%	22.17%	22.8%	23.78%

*: significant at the 10% level; †: significant at the 5% level; ‡: significant at the 1% level

Computing time is significantly affected by household sampling ratio, especially when dataset size is above RAM capacity (2 Gb). The aggregate results appear not to be very sensitive to sampling size concerning the fraction of households used to estimate demand. Indeed, the correlations between aggregate demand (D_j) estimated in the 4 samples is very high. However, the precision of aggregate demand deteriorates in small municipalities when sample size becomes too small (see Section 4.4.4.

The price coefficients (by itself and crossed with age and income) are sensitive to sample size, although they remain very significant in all samples. The coefficients of the numbers of railway and metro stations become not significant at the 5% level in the 1% and 2% samples. Women appear nearly two times more sensitive than men to accessibility (commuting) in the 1% sample. The coefficient of distance to highway also nearly doubles, and it becomes less significant in the 1% sample. The coefficient of accessibility by car is very unstable and even changes sign from one sample to the other. This is not very surprising since the accessibility variables are highly correlated, and this suggests that it is not possible to identify the coefficients of more than one accessibility variable in the smallest samples. Most of the population composition coefficients remain stable across sample sizes, except for the least significant ones. Once again, it shows the need to be parsimonious with respect to the number of variables measuring population composition in the smallest samples. The coefficients of the three size variables are also unstable across sample sizes, probably because they are too highly correlated.

4.3.2 A more parsimonious location choice model

Based on the above results, it seems that the location choice model should be more parsimonious in reduced samples, and we decided to select, in addition to the "same district" dummy, only one variable for each of the three following items: price, accessibility, population composition, and size. We now test the sensitivity to the number of alternatives used in random sampling, in such a parsimonious model. These results are shown in Table 6.

Table (6): A more parsimonious model

Variable \ sampling rate	100%	20%	2%	1%
Same district as before move	2.5071‡	2.5036‡	2.5370‡	2.6081‡
Log(Price)	0.0448‡	0.0419‡	0.0521‡	0.0378
Average travel time from j [100']	-0.3923‡	-0.402‡	-0.236†	-0.2742
% households with a foreign head	-0.7257‡	-0.748‡	-0.848‡	-1.116‡
Log(Population)	0.0328‡	0.0936	0.1071‡	0.1239
Pseudo R^2	18.98%	18.97%	19.52%	20.66%

The results of the sampling Sections will be useful for specifying the location choice model with constraints, which will be estimated using a universal choice set (all the 1300 municipalities) on reduced samples of 1% and 2% of moving households.

As a conclusion of this sensitivity analysis without capacity constraints, the parsimonious model estimated on a 1% or 2% sample of households with universal choice set seems a good benchmark for the estimates with constraints.

4.4 Location choice with capacity constraints

Since random sampling introduces heterogeneity in the individual allocation ratios Ω^i, the homogeneous model can only be estimated using the universal set of alternatives ($1,300$ municipalities). However, the dataset with universal choice set and 100% of the movers would contain 766 million lines and its size would be over 1 terabyte with all the variables used in the previous section! In order to illustrate the method on a more tractable data set and for obtaining preliminary results, we selected a small number of explanatory variables (Log(Price); Average travel time from j, commuting (TC); % households with a foreign head; Log(Population); and Same district as before move in the heterogeneous case), and we used a representative random sample of 2% households. This represents $15,321,800$ lines, $11,786$ movers and slightly more than 1 Gb. Although convergence is attained after a small number of iterations (less than 5 iterations for constraints at each estimation loop, less than 10 demand loops), about 2 hours are necessary for estimating the homogeneous version of the model, and computing time goes over 15 hours for the heterogeneous version model.[3] In the near future, working on a more powerful computer and improving the efficiency of the algorithm will allow increasing significantly both the sampling rate and the number of explanatory variables. A trade-off between the number of households, the number of alternatives by individual and the number of explanatory variables will remain, however.

In order to estimate the effect of the availability constraints based on the yearly demand, it is necessary to infer an annual supply of dwellings from the current supply observed in March 1998. This annual supply depends on the λ ratio defined in Section 2. As a preliminary sensitivity analysis, we consider two cases: $\lambda = 0.5$ and $\lambda = 0.45$. We are also interested in the sensivity of the results to the sampling ratio and compare the results obtained on a 1% and on a 2% representative samples of households. Because of constraints on the data sets size, we only explore the parsimonious models studied in Section 4.2 (see Table 6). The "Same district as before move" dummy is used in the heterogeneous case, but not in the homogeneous one.

4.4.1 Extent of the constraints and allocation ratio in the homogeneous case

We turn now to an assessment in the homogeneous case of the effect on the number of constrained alternatives and on the allocation ratio of the sampling rate of households and of λ. Table 7 presents the results for 1% and 2% samples of households, using $\lambda = 0.5$ and $\lambda = 0.45$.

Table (7) Fraction alternatives constrained and allocation ratio, homogeneous case

	$\lambda = 0.5,\ 1\%$	$\lambda = 0.5,\ 2\%$	$\lambda = 0.45,\ 2\%$
% alt. constrained:			
Ex ante, initial demand	36.2%	40.2%	47.31%
Ex post, initial demand	39.3%	43.5%	52.69%
Ex ante, final demand	46.5%	51.5%	61.77%
Ex post, final demand	51.2%	56.2%	70.85%
Average [min;max] S_j/D_j ratio:			
Initial demand	1.18 $[0.06; 5.48]$	1.14 $[0.06; 4.97]$	1.09 $[0.06; 4.94]$
Final demand	1.08 $[0.01; 5.26]$	1.03 $[0.03; 4.64]$	0.91 $[0.00; 4.53]$
Allocation ratio Ω:			
Ex ante, initial demand	1.019	1.0208	1.0365
Ex post, final demand	1.050	1.0391	1.1274

For all values of λ, we have the following result: the fraction of constrained alternatives is always larger once algorithm 1 has converged. This is true when the initial demand is considered and when the final demand is considered. For the final demand, the fraction of constrained alternatives is (almost always) larger than for the initial demand at any stage of algorithm 1. In other words, convergence of the inner or outer loop increases the fraction of constrained alternatives. The reason behind this intuition is as follows: at the initial demand loop, the observed demand under-estimates the ex ante demand in constrained alternatives and over-estimates ex ante demand in unconstrained alternatives. These biases are corrected at the final demand loop. As expected, ceteris paribus, when λ decreases, the level of constraints increases since a lower value of λ means a smaller number of dwellings available on average during one year. Note that a very small variation of λ induces an important variation of the fraction of alternatives constrained.

The average ratio of supply to demand, which is inversely proportional to the degree of constraint, decreases from the initial demand to the final demand (since this increases the amount of constraints). The same reasoning applies to Ω, which is an increasing function of the level of constraints. According to the final estimates, when $\lambda = 0.5$, 4% to 5% of the households are not located in their preferred municipality. This fraction increases to 13% when $\lambda = 0.45$.

[3] Using the SAS software on an IBM laptop with 2 Gb RAM.

4.4.2 Extent of the constraints and allocation ratio in the heterogeneous case

Here we examine in the heterogeneous case the effect on the number of constrained alternatives and on the allocation ratio of the sampling rate of households and of λ. Table 8 presents the results for 1% and 2% samples of households, using $\lambda = 0.5$ and $\lambda = 0.45$.

Table (8) Fraction alternatives constrained and allocation ratio, heterogeneous case

	$\lambda = 0.5,\ 1\%$	$\lambda = 0.5,\ 2\%$	$\lambda = 0.45,\ 2\%$
% alt. constrained:			
ex ante, initial demand	36.62%	41.7%	48.69%
ex post, initial demand	41.31%	46.4%	57.08%
ex ante, final demand	43.62%	52.8%	60.62%
ex post, final demand	50.46%	58.6%	73.23%
Average [min;max] S_j/D_j ratio:			
Initial demand	1.17 $[0.06; 5.49]$	1.12 $[0.06; 5.09]$	1.08 $[0.06; 4.93]$
Final demand	1.08 $[0.01; 5.49]$	1.02 $[0.03; 4.84]$	0.92 $[0.00; 5.16]$
Average [min;max] Ω^i:			
ex ante, initial demand	1.02 $[1.00; 1.04]$	1.02 $[1.00; 1.05]$	1.04 $[1.01; 1.09]$
ex post, final demand	1.05 $[1.01; 1.11]$	1.04 $[1.01; 1.11]$	1.20 $[1.00; 1.68]$
Average [min;max] $\bar{\Omega}_j$:			
ex ante, initial demand	1.026 $[1.01; 1.03]$	1.026 $[1.01; 1.03]$	1.05 $[1.02; 1.06]$
ex post, final demand	1.07 $[1.03; 1.08]$	1.06 $[1.03; 1.08]$	1.28 $[1.10; 1.40]$

The results for the heterogeneous are qualitatively similar to the heterogeneous case in terms of the fraction of alternatives constrained and the supply over demand ratio. Note that the allocation ratio in the heterogeneous case is specific either to the household or to the alternative. When $\lambda = 0.5$, the average individual-specific allocation ratio is approximately equal to the homogeneous allocation ratio, but this is no more true when $\lambda = 0.45$.

For the final estimates in the 2% sample with $\lambda = 0.5$, the minimum individual allocation ratio is $\Omega^{i,\min} = 1.01$, which means that the least constrained households (which prefer the least constrained municipalities) have 1% probability to be denied access to their preferred choice. By contrast, the maximum individual allocation ratio is $\Omega^{i,\max} = 1.11$, which means that the most constrained households (which prefer the most constrained municipalities) have 11% probability to be denied access to their preferred choice. At the same time, the minimum alternative-specific allocation ratio is $\bar{\Omega}_{j,\min} = 1.03$, which means that 3% of the households allocated in commune with the lowest relative excess demand had their first best in another municipality. On the other hand, the maximum alternative-specific allocation ratio is $\bar{\Omega}_{j,\max}(5,3) = 1.08$, which means that 8% of the households allocated in the commune with the highest relative excess demand had their first best in another commune. Those figures are quite large, since the only source of heterogeneity in the estimated model is the dummy variable "Same district as before move". The variability of Ω^i and $\bar{\Omega}_j$ will be far more significant when more variables will be added to the demand estimation with constraints, as in Table 4. The results concerning the degree of constraints are similar using a 1% sample, but computing time is significantly reduced (from about 15 hours to about 6 hours in the heterogeneous case).

4.4.3 Estimation of the demand

The overall fit of the estimated model is quite good given the low number of explanatory variables: the Pseudo-R^2 is about 30% in the heterogeneous case and 23% in the homogeneous case. The unexpected positive sign of the price coefficient reflects an omission bias: some alternative attributes (e.g. % poor households) are correlated both with price and with demand. Because of this omission bias (which will be corrected by the introduction of additional explanatory variables), we will not interpret the value of the price coefficient, but simply note that its value and significance is very different in the 1%

and 2% samples. The most significant variable remains the dummy "same district as before".and the corresponding estimated coefficient is stable across sample sizes. Strangely enough, the price coefficient is unchanged when the corrected factor is introduced.

We could anticipate identification problems because both price and correcting factor $\log\left(\pi_j^i\right)$ reflect the supply/demand ratio in some sense: $\pi_j^i = S_j/D_j$ in the alternatives constrained ex post and equilibrium on the dwelling market suggests that price depends on the supply/demand ratio. However, the correlation between log(Price) and supply/demand ratio is only -4.85% (significant at the 10% level but not at the 5% level) for initial demand estimates and -1.12% (not significant at all), suggesting that prices are only marginally influenced by the frictions in the housing market. In other words, prices do not actually clear the market, a conclusion we can only generalize at this point to the Il-de-France housing market. Note that the price coefficient is not significant in the 1% sample and only marginally significant in the 2% sample. However, we can anticipate that the price coefficient will become more significant in larger samples and will become negative when enough covariates will be added, as shown in some of the more fully specified results presented in the paper. The coefficients of the accessibility variable (average travel time from the commune) and of the size variable (log(Population)) are stable across sample sizes and between initial and final estimates. By contrast, the fraction of foreign households is more sensitive both to the sample size and to the introduction of constraints.

Table (9): Comparison of results in the homogeneous and heterogeneous cases (2% sample)

Variable		Homogeneous case Coefficient (t-stat)	Heterogeneous case Coefficient (t-stat)
Same district as before move	initial, $\lambda = 0.45$	does not apply	2.5765 (105.71)
	final, $\lambda = 0.45$	does not apply	2.5890 (106.05)
	initial, $\lambda = 0.5$	does not apply	
	final, $\lambda = 0.5$	does not apply	
Log(Price)	initial, $\lambda = 0.45$	0.0280 (1.67)	0.0219 (1.30)
	final, $\lambda = 0.45$	0.0321 (1.92)	0.0268 (1.59)
	initial, $\lambda = 0.5$	0.002501 (0.11)	
	final, $\lambda = 0.5$	0.000250 (0.01)	
Average travel time from j	initial, $\lambda = 0.45$	-0.2443 (-2.27)	-0.2876 (-2.64)
	final, $\lambda = 0.45$	-0.1392 (-1.28)	-0.1999 (-1.82)
	initial, $\lambda = 0.5$	-0.3170 (-2.07)	
	final, $\lambda = 0.5$	-0.3302 (-2.14)	
% households with a foreign head	initial, $\lambda = 0.45$	-0.4483 (-2.28)	-0.5612 (-2.77)
	final, $\lambda = 0.45$	-0.8548 (-4.21)	-0.9530 (-4.57)
	initial, $\lambda = 0.5$	-0.9955 (-3.46)	
	final, $\lambda = 0.5$	-1.0966 (-3.71)	
Log(Population)	initial, $\lambda = 0.45$	1.0846 (105.27)	1.0722 (100.23)
	final, $\lambda = 0.45$	1.0590 (98.82)	1.0495 (94.76)
	initial, $\lambda = 0.5$	1.0807 (105.27)	
	final, $\lambda = 0.5$	1.0733 (68.33)	
$\mathrm{Log}\left(\pi_j^i\right)$	final, $\lambda = 0.45$	0.5157 (8.27)	0.3888 (7.88)
	final, $\lambda = 0.5$	0.7848 (7.22)	
Pseudo-R^2	initial, $\lambda = 0.45$	22.90%	29.68%
	final, $\lambda = 0.45$	22.94%	29.72%
	initial, $\lambda = 0.5$	23.15%	
	final, $\lambda = 0.5$	23.19%	

4.4.4 Comparison of constrained and unconstrained aggregate demand

Table 10 reports the correlation coefficient between the actual demand (number of households which moved in to each municipality in 1998) and the demand predicted by the different models. The "Full model" was estimated using a random sampling of 8 alternatives for each household, with 24 explanatory variables, as reported in Table 5. The "Parsimonious model" was estimated using the same random sampling of 8 alternatives, but with only 5 explanatory variables, as reported in Table 6. The estimates with constraints are those of the homogeneous case with $\lambda = 0.5$. Table 10 shows that the precision of the aggregate estimated remains satisfactory when the sampling ratio decreases. It is interesting to notice that precision is not significantly altered in the parsimonious model. It seems that the precision becomes poorer and poorer (correlation less than 90% in the 1% sample, less than 93% in the 2% sample) when the "size" (measured as the number of observed moves into the commune in 1998) of the alternatives decreases. The correlation with actual demand remains quite high in the 1148 municipalities with less than 1000 movers in 1998, but it diminishes in the 1054 municipalities with less than 500 movers in 1998 and becomes even lower in smaller alternatives. However, precision significantly improves in the 1% and 2% samples of households when a universal set of alternatives is used for the demand predictions. Note that with a universal set of altenatives the precision is the same in the 1% and 2% samples of households, whereas, with a random sampling of 8 alternatives, precision was far poorer in the 1% sample compared to the 2% sample of households (in the small municipalities). Integrating the constraints significantly improves the precision of the demand estimates, especially for small alternatives.

Table (10): Correlation between actual number of movers and demand predicted by different models

Variable \ sampling rate	All alt.	< 1000 1148	< 500 1054	< 200 948	< 100 841	< 50 725
Full model, 100%	0.99786	0.99410	0.98926	0.97642	0.95993	0.92051
Parsimonious model, 100%	0.99677	0.98774	0.97649	0.96259	0.94436	0.89790
Full model, 20%	0.99791	0.99195	0.98553	0.96724	0.93288	0.87265
Parsimonious model, 20%	0.99666	0.98643	0.97262	0.95355	0.92062	0.85946
Full model, 2%, 8 alt.	0.99679	0.97355	0.93976	0.87008	0.74415	0.62582
Parsimonious model, 2%, 8 alt.	0.99583	0.97017	0.93613	0.88778	0.77750	0.67020
Full model, 1%, 8 alt.	0.99577	0.95541	0.89516	0.76587	0.61803	0.51244
Parsimonious model, 1%, 8 alt.	0.99497	0.95310	0.89856	0.79354	0.65629	0.55291
Parsimonious model, 2%, all alt.	0.99531	0.98033	0.96350	0.94274	0.92180	0.85824
idem, with constraints	*0.99554*	*0.98345*	*0.97220*	*0.96246*	*0.94261*	*0.90053*
Parsimonious model, 1%, all alt.	0.99527	0.97994	0.96236	0.93971	0.91761	0.85824
idem, with constraints	*0.99968*	*0.98279*	*0.97119*	*0.96099*	*0.94131*	*0.90053*

5 Conclusion and extensions

Constraints on the availability of alternatives are clearly present in some markets, such as the housing market in the Ile-de-France. We have demonstrated that housing prices do tend to capitalize the imbalances between demand and supply in the short run period of one year, but do not fully accomplish the task of resolving short-term market disequilibrium. The specification linking residential location choices and housing prices through he aggregation of individual demands provides a tractable means to address this economic endogeneity, and the results demonstrate that the market clearing role for prices is significant, but even after accounting for the price effect, a large fraction of alternatives still had excess demand, imposing availability constraints on consumers that ultimately must make suboptimal choices from unconstrained alternatives. This is a significant empirical finding, since virtually all previous literature in discrete choice modeling in market contexts implicitly or explicitly assume that prices clear the market.

We have developed and demonstrated a viable mechanism to address the influence of constrained alternatives on choice outcomes, by using relatively weak assumptions consistent with the property of independence of irrelevant alternatives inherent in the multinomial logit model: 1) if an individual prefers an alternative j, which is unconstrained ex post, he can be sure to be allocated to it, and 2) the individual allocation ratio is the same for each individual. We then develop a discrete choice model specification and computational algorithm that corrects for capacity constraints in a way that is consistent with the IIA assumption, and shown that this converges to a unique equilibrium in the case of homogeneous agents. Though we have not proven the existence and uniqueness of an equilibrium solution in the heterogeneous case, simulation results show that convergence occurs. We leave the theoretical proof for future research.

Application of this constrained discrete choice algorithm confirms, first, that the share of alternatives initially estimated to be constrained significantly increases once the parameter estimates are adjusted to correct for the effect of the constraints. In our heterogeneous case, the initial estimate of ex ante constraints was that 40.7% of communes were constrained. At convergence of the algorithm, after accounting for the complex spillovers of consumers forced to make second-best choices among initially unconstrained alternatives that could become constrained due to these spillovers, this estimate increased to 58.6%.

Our sensitivity testing confirms that most parameter estimates in the unconstrained model are fairly robust to sampling of households, down to quite small proportions such as 1 to 2%, where precision decreases. Due to the size of the data and the use of a full enumeration of 1300 alternatives in the heterogeneous case, we were limited to using small sampling rates for households for testing a very parsimonious model that clearly reflects omitted variables bias, but allowed the initial testing of the algorithm to proceed. The results of this testing confirm that the algorithm converges in a small number of iterations, and that parameters can be adjusted for the bias resulting from capacity constraints. Finally, our analysis of demand predictions confirms that the constrained discrete choice algorithm significantly improves predictions of demand, particularly for those alternatives with relatively small market share. This may have important implications for research in which the alternatives of interest reflect small market shares, and the system contains some capacity constraints – even if the alternatives of interest are not themselves constrained ex ante.

The method developed so far can be improved along several directions. First, it has been assumed that the λ coefficient, corresponding to the fraction of movers who contribute to the yearly supply (and also to the average fraction of the year the dwellings were vacant before a household moved in), is the same across all municipalities. This is clearly not true, and the λ coefficient should be computed on a location-specific basis (municipality or, at least county level). Second, the sampling of households and of alternatives has been discussed, but not fully explored yet. This should be performed in a more elaborated version of the proposed model. Third, other allocation mechanisms could be envisaged. Our allocation mechanism relies to a certain extent on an assumption of equity in the adjustment of choice probabilities. Other mechanisms should also be explored. Finally, further empirical testing and refinement of the computational performance of the constrained choice algorithm, and implementation in generally-available software, remain to be completed, and these steps will make it possible to explore applications of this approach to policy questions relevant to constrained choice contexts.

6 Acknowledgements

This research has been funded in part by PREDIT, the French Department of Transportation (contract with IAURIF, University of Cergy-Pontoise and adpC). We would like to thank Dany Nguyen (IAURIF), for helpful comments and for providing us with the data. We benefited from the comments of the seminar participants of the second Khumo conference, of the Dijon conference, and from the Fiesta conference at the Ecole Nationales des Ponts et Chaussées (in particular F. Leurent, L. Swartz and J. Maurice). The third author also acknowledge the National Science Foundation Grants EIA-0121326 and IIS-0534094.

References

[Anas and Cho1988] Alex Anas and J.R. Cho. A dynamic, policy oriented model of the regulated housing market; the swedish prototype. *Regional Science Urban Economics*, 18:210–231, 1988.

[de Palma and Lefèvre1981] André de Palma and Claude Lefèvre. A probabilistic search model. *Journal of Mahematical Sociology*, 8:43–60, 1981.

[de Palma and Rouwendal1996] André de Palma and Jan Rouwendal. Availability constraints in the housing market. *Journal of housing economics*, 5:105–132, 1996.

[de Palma *et al.*2006] André de Palma, Nathalie Picard, and Paul Waddell. A model of residential location choice with endogenous housing prices and traffic for the paris region. *European Transport (in press)*, 2006.

[Lerman1977] Steve Lerman. Location, housing, automobile ownership, and the mode to work: A joint choice model. *Transportation Research Board Record*, 610:6–11, 1977.

[McFadden1978] Daniel McFadden. Modelling the choice of the residential location. In A. Karlquist et Al., editor, *Spatial Interaction Theory and Residential Location*, pages 75–96. North Holland, Amsterdam, 1978.

[Quigley1976] John Quigley. Housing demand in the short-run: An analysis of polytomous choice. *Explorations in Economic Research*, 3:76–102, 1976.

[Weisbrod and Ben-Akiva1980] Lerman S.R. Weisbrod, G.E. and M. Ben-Akiva. Tradeoffs in residential location decisions: Transportation versus other factors. *Transport Policy and Decision Making*, 1:13–26, 1980.

[Williams1979] Roberton C Jr. Williams. A logit model of demand for neighborhood. In David Segal, editor, *The Economics of Neighborhood*, pages 17–42. Academic Press, New York, 1979.