UMR 8184
Thema
THéorie Économique, Modélisation et Applications

Thema Working Paper n°2013-02
Université de Cergy Pontoise, France

Commuting Time and Accessibility in a Joint Residential Location, Workplace, and Job Type Choice Model

Ignacio A. Inoa
Nathalie Picard
André de Palma

January, 2013

UNIVERSITÉ de Cergy-Pontoise
CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Commuting Time and Accessibility in a Joint Residential Location, Workplace, and Job Type Choice Model*

Ignacio A. Inoa[1], Nathalie Picard[1], and André de Palma[2]

[1] *Université de Cergy-Pontoise, Thema*†
[2] *Ecole Normale Supérieure de Cachan, CES*‡

7 January 2013

Abstract

The effect of an individual-specific measure of accessibility to jobs is analyzed using a three-level nested logit model of residential location, workplace, and job type choice. This measure takes into account the attractiveness of different job types when the workplace choice is anticipated in the residential location decision. The model allows for variation in the preferences for job types across individuals and accounts for individual heterogeneity of preferences at each choice level in the following dimensions: education, age, gender and children. Using data from the Greater Paris Area, estimation results indicate that the individual-specific accessibility measure is an important determinant of the residential location choice and its effect differ along the life cycle. Results also show that the job type attractiveness measure is a more significant predictor of workplace location than the standard measures.

Keywords: residential location, job location, accessibility, nested logit, Greater Paris.
JEL Codes: R21, C35, C51.

*The authors would like to thank Kiarash Motamedi and Mohammad Saifuzzaman for their help on travel time data computation. Useful discussion, comments, and suggestions were provided by Ismir Mulalik, Robert Pollak and participants at the 2012 Bari XIV SIET Conference, the 2012 Lausanne SustainCity Consortium Meeting, and the THEMA Lunch Seminar. The research was funded by Ademe and DRI (direction de la recherche et de l'innovation du ministère chargé du développement durable) as part of the French national research programme in land transportation PREDIT.

†133. Bd. du Port, 95011 Cergy-Pontoise Cedex, France, ignacio.inoa@u-cergy.fr, nathalie.picard@u-cergy.fr

‡61. Av. du Président Wilson, 94230 Cachan Cedex France, andre.depalma@enscachan.fr

1 Introduction

Residential location choice models have historically been estimated conditional on workplace, or vice versa. The first discrete choice models applied to residential location (Lerman, 1976; McFadden, 1978; Anas, 1981) borrowed from the Alonso-Muth-Mills literature on monocentric models the assumption of exogenous determination of workplace location (Alonso, 1964; Muth, 1969; Mills, 1972). The interdependency between residential and workplace location was studied during the late 70s, with the monocentric model extensions allowing simultaneous choice of workplace and residential location (Siegel, 1975; Simpson, 1980), and during the 80s, with the Linneman et al. (1983) joint multinomial logit model analyzing residential migration and job search.

The relevance of the exogenous workplace assumption in residential location choice models has been questioned from the early 90s following the empirical results of Waddell (1993) that showed that a *joint* logit model of workplace, tenure and residential location outperform a nested logit model of tenure and residential location choice conditional on workplace.

Subsequent applications and theoretical developments of residential location and of workplace discrete choice models were made separately. On the one hand, residential location choice models have been studied in relation, among others, with mobility or relocation (Clark et al., 1999; Lee et al., 2010), mode choice (Eliasson et al., 2000), and accessibility (Ben-Akiva et al., 1998). On the other hand, workplace location choice models have been mostly developed in the framework of aggregated travel models.

Explicit modeling of both residential location and workplace choice have been studied within the multi-worker household discrete choice literature. In this field, researchers have been mainly interested in analyzing the influence of spouses' earnings and commuting time on the choice of the household residential location and spouses' specifics job locations (Freedman et al., 1997; Abraham et al., 1997). Additionally, Waddell et al. (2007) developed a discrete choice model of joint residential location and workplace adapting methods of market segmentation for single worker households. Doing so, no a priori assumption has to be made on the exogenous choice (workplace first and residence after or vice versa) and the probability of making one

choice before the other is determined as a function of the household characteristics.

The literature on topics related to residential location and workplace range from mobility and job uncertainty to the analysis of the decision-making process in discrete choice models. Crane (1996) and Kan (1999; 2002) provide insights on modeling the effect of job changes on residential mobility, mobility expectation and commuting behavior. Introducing risk in discrete choice models provides a new research direction in residential location and workplace. Palma et al. (2008) offer a review on the implications of risk and uncertainty in the framework of discrete choice models in specific fields including residential location and transportation and provide recommendations on its implementation. A recent research strand in choice models is advocating to take into account the effects of context on the process leading to a choice. Ben-Akiva et al. (2012) provide a review and illustration of *process and context* in discrete choice models.

Commuting time is one of the main determinants of residential location. Household locations and workplaces are interdependent choices because they jointly determine commuting time. The residential location and workplace can be modeled as a textitsequential two-stage decision process. In such a decision process, the second-stage decision is made conditional on the first stage, and the second stage is anticipated in the first-stage decision. In the model considered here, workers choose a workplace conditional on their actual residential location (second stage), and they anticipate their potential workplaces (job opportunities) when choosing their residential location (first stage). In this setting, actual commuting time (between the locations chosen for living and for working) is relevant for explaining the workplace location choice (Levine, 1998; Abraham et al., 1997), whereas accessibility to jobs is the relevant variable for explaining the residential location choice (Anas, 1981; Ben-Akiva et al., 1998; Levinson, 1998).

Which decision is made first is an open research question. Is it the choice of workplace or the residential location choice? The extent to which workplace location will depend on residential location, or conversely, varies along household's life cycle and is affected by real estate and labor market rigidities, the availability of jobs, and the demographic and socioeconomic characteristics of households (Waddell, 1993; Waddell et al., 2007). The most widely used approach to model the sequential

decision-making processes in a (residential, workplace or mode) choice framework is to use discrete choice models, as in Anderson et al. (1992); Ben-Akiva et al. (1985). This will be the approach adopted here. Discrete choice models allow to study the location decision choice interdependency (nested models) and to model residence (and workplace) choice as a trade-off among local attributes that can vary across sociodemographic segments, as described in Sermons et al. (2001) and Bhat et al. (2004).

Despite the variety of contributions to the study of residential location, little has been said regarding the influence of job type attractiveness on the accessibility to jobs and accordingly on the residential location and workplace choices when individuals are forward-looking. A three-level nested logit model is developed here, allowing to study, within a behavioral framework (RUM), the interdependency of residential location and workplace, while accounting for variation across individuals on the preferences for job types. In this model, residential location is the upper level choice, and workplace location and job type are the middle and lower level choices, respectively. This nested structure allows to build an individual-specific accessibility measure, which corresponds to the expected maximum utility across all potential workplace locations (middle level). When considering accessibility to jobs, the choice of a particular workplace location is influenced by the relative distribution of jobs of the same type of the worker. Modeling the job type choice (lower level) allows for the computation of an individual-specific measure of attractiveness to job types (log-sum variable) that is used in the workplace location choice model.

In Section 2, a framework that introduce the role of risk and information in discrete choice models is developed. A three-level nested logit model is developed to analyze the residential, workplace and job type choices when information can be acquired at each choice step. In this model, residential location and workplace are related through the generalized cost of commuting. However, workplace and job type are assumed to be independent: the relative preference for one job type over another one is the same in all locations. The data are presented in Section 4. The empirical methodology and the results are provide in Section 5. Finally, Section 6 concludes with insights on implications of the proposed modeling for urban models.

2 Risk and information in discrete choice models

The role of information in discrete choice models is analyzed below, with a special focus on MultiNomial Logit model (MNL), and nested Logit model (NL). It will be applied to the context of residential location, workplace and job type in the next section.

Each individual (household) chooses a single alternative belonging to the choice set $\mathcal{L}$, with $|\mathcal{L}| = L$. In the application considered below, it will be assumed that the full set of alternatives, denoted by $\mathcal{L}_0$, is partitioned in I choice sets, or nests denoted by $\mathcal{L}_i,\ i = 1, ..., I$ ($\mathcal{L}_i$ represents the nest containing a given alternative i, $i \in \mathcal{L}_0$). The list of nests is assumed exogenous and identical for all decision makers. The utility of individual n choosing alternative i is a random variable given by $U_n(i) = V_n(i) + \varepsilon_n(i)$, where $V_n(i)$ is the deterministic (or measured) utility, and where $\varepsilon_n(i)$ is an absolutely continuous random variable. The random term, $\varepsilon_n(i)$ is an error term unknown to the modeler, corresponding to unobservable individual and/or alternative-specific characteristics (Manski et al., 1977). Such models are also referred to as Additive Random Utility Models (McFadden, 1978; Anderson et al., 1992). Alternatively, the error term $\varepsilon_n(i)$ can be interpreted as a match value between individual n and alternative i. The modeler is assumed to know the distribution of the error terms, but not their specific values. In the different cases analyzed below, different assumptions are made about the level of information available to the decision maker at the different stages of the decision process (before/after she chooses a set of alternatives).

The expectation of $\varepsilon_n(i)$ is set to zero without loss of generality, and its standard deviation is assumed the same across alternatives and denoted by μ. According to the law of comparative judgments (Thurstone, 1927), when the decision maker knows the realization of $\varepsilon_n(i),\ i \in \mathcal{L}$, she selects the alternative with the highest utility. Formally, the probability that individual n selects alternative i in choice set $\mathcal{L}$ is

$$\mathbb{P}_n(i|\mathcal{L}) = \Pr\left\{U_n(i) \geqslant U_n\left(i'\right),\ \forall\ i' \in \mathcal{L}\right\}. \tag{1}$$

We will extend this formula to the case where the decision maker does not know the realization of $\varepsilon_n(i)$. In the MNL model, the error terms are assumed inde-

pendent and identically distributed (i.i.d.). Their common law is Gumbel or double exponential distributed, with:

$$\mathbb{P}\left\{\varepsilon \leq x\right\} = \exp\left\{-\exp\left[-\left(x+\gamma\right)/\mu\right]\right\}, \tag{2}$$

where $\gamma \simeq 0.5772$ (shift parameter) is the Euler constant, which guarantees that $\mathbb{E}\left(\varepsilon\right) = 0$, and the standard deviation of ε is proportional to μ (scale parameter). The cdf of a standard Gumbel distribution is $\mathbb{P}\left\{\varepsilon \leq x\right\} = \exp\left\{-\exp\left(-x\right)\right\}$; its expectation is γ and its standard deviation is $\pi/\sqrt{6}$ (this corresponds to $\mu = 1$). The pdf of this standard random variable is denoted by f and given by $f\left(x\right) = \exp\left[-x\right]\exp\left\{-\exp\left[-x\right]\right\}$.

When the error terms are i.i.d. Gumbel with scale parameter μ, the choice probabilities have a closed form (McFadden, 1978):

$$\mathbb{P}_n\left(i|\mathcal{L}\right) = \frac{\exp\left(\frac{V_n(i)}{\mu}\right)}{\sum\limits_{i'\in\mathcal{L}}\exp\left(\frac{V_n(i')}{\mu}\right)},\ i \in \mathcal{L}. \tag{3}$$

Clearly, the same expression would hold for any shift parameter (other than γ or 0).

Consider two alternatives i_1 and i_2 included in two choice sets, $\mathcal{L}$ and $\mathcal{L}$'. Then, the following property holds:

$$\begin{aligned}\frac{\mathbb{P}_n\left(i_1|\mathcal{L}\right)}{\mathbb{P}_n\left(i_2|\mathcal{L}\right)} &= \frac{\exp\left(\frac{V_n(i_1)}{\mu}\right)}{\sum\limits_{i'\in\mathcal{L}}\exp\left(\frac{V_n(i')}{\mu}\right)} \Big/ \frac{\exp\left(\frac{V_n(i_2)}{\mu}\right)}{\sum\limits_{i'\in\mathcal{L}}\exp\left(\frac{V_n(i')}{\mu}\right)} = \frac{\exp\left(\frac{V_n(i_1)}{\mu}\right)}{\exp\left(\frac{V_n(i_2)}{\mu}\right)} \\ &= \frac{\exp\left(\frac{V_n(i_1)}{\mu}\right)}{\sum\limits_{i''\in\mathcal{L}'}\exp\left(\frac{V_n(i'')}{\mu}\right)} \Big/ \frac{\exp\left(\frac{V_n(i_2)}{\mu}\right)}{\sum\limits_{i''\in\mathcal{L}'}\exp\left(\frac{V_n(i'')}{\mu}\right)} \\ &= \frac{\mathbb{P}_n\left(i_1|\mathcal{L}'\right)}{\mathbb{P}_n\left(i_2|\mathcal{L}'\right)}.\end{aligned} \tag{4}$$

This corresponds to the Independence of Irrelevant Alternatives (IIA) property which characterizes the Multinomial Logit model. The relative probability that individual n selects alternative i_1 rather than alternative i_2 is the same in all the choice sets containing both i_1 and i_2.

Now consider an individual who should select between different choice sets. She needs to evaluate the benefit (or surplus) of the different choice sets beforehand, i.e.

before the choice is made based on the information available at that stage.

The terminology *ex-ante* and *ex-post* is now introduced. It characterizes the level of information available to the decision maker. *Ex-ante* utility of the choice set $\mathcal{L}$ corresponds to the value associated to $\mathcal{L}$ before it has been selected. E*x-post* utility of the choice set $\mathcal{L}$ corresponds to the value associated to $\mathcal{L}$ after it has been selected, based on the information available at that stage. Different cases are possible.

2.1 Information revealed *ex-post*

This case will be referred, for convenience, to *learning*. In this case, individual has no information *ex-ante* about the exact value of the idiosyncratic terms, but she acquires such information after selecting the choice set.

The *ex-post* surplus is derived as follows. Once the individual has selected the choice set, she has access to information about all the values of the idiosyncratic terms. That is, she can observe the realizations $\mathbf{e}_n \equiv (e_1, ..., e_L)$ of the random terms $\varepsilon_n(1), .., \varepsilon_n(L)$. Once the choice set $\mathcal{L}$ is selected, the utility of alternative i is a number denoted by $\tilde{U}_n(i) = V_n(i) + e_i,\ i = 1, ..., L$. The *ex-post* value of the utility of choice set $\mathcal{L}$ is a number equal to:

$$\tilde{U}_n(\mathcal{L}; \mathbf{e}_n) = \max_{i \in \mathcal{L}} \tilde{U}_n(i). \tag{5}$$

Ex-ante, the values of the error terms are unknown, but the probability distribution of $\varepsilon_n(1), ..., \varepsilon_n(L)$, evaluated at any point $\mathbf{e}_n$ is known, and given by $f(e_1),\ f(e_2),\ ....f(e_L)$. The surplus of choice set $\mathcal{L}$ corresponds to its expected utility, computed *ex-ante*. It is given by:

$$\mathbb{E}U_n(\mathcal{L}) \equiv \int_{\mathbb{R}^L} \tilde{U}_n(\mathcal{L}; \mathbf{e}_n) f(e_1) f(e_2) ....f(e_L)\, de_1 de_2 ... de_L. \tag{6}$$

This formula is in the vein of von Neumann-Morgenstern Expected Utility function. It involves the computation of L integrals.

Using the Gumbel specification, Ben-Akiva et al. (1985) have derived an exact formula for the above expression. In this case, the *ex-ante* utility of choice set $\mathcal{L}$ is

a surplus corresponding to the expected utility of choice set $\mathcal{L}$:

$$S_n(\mathcal{L}) \equiv \mathbb{E}U_n(\mathcal{L}) = \mu \log\left\{\sum_{i\epsilon\mathcal{L}} \exp\left(\frac{V_n(i)}{\mu}\right)\right\}. \tag{7}$$

Assume, without loss of generality, that $V_n(1) < ... < V_n(i) < ...V_n(L)$. Then,

$$\lim_{\mu\to 0} \mathbb{E}U_n(\mathcal{L}) = V_n(L). \tag{8}$$

The interpretation is as follows: when the importance of idiosyncratic terms vanishes, the model becomes deterministic. In this case, individual n selects with probability 1 the best alternative L and gets the corresponding utility $V_n(L)$.

Moreover, as the variance of the idiosyncratic terms gets very large, each MNL choice probability tend to $1/L$ and the expected utility of the chosen alternative tends to infinity:

$$\lim_{\mu\to\infty} \mathbb{E}U_n(\mathcal{L}) = \infty. \tag{9}$$

This is because the maximum of i.i.d. Gumbels has a Gumbel distribution with the same variance (here $\mu \to \infty$) and with the same expectation (which tends here to infinity).

As expected, the attractiveness of a choice set increases as alternatives are added (since positive terms are added to the sum).

Assume now that $V_n(1) = ... = V_n(i) = ... = V_n(L) \equiv V_n$. Then

$$\mathbb{E}U_n(\mathcal{L}) = V_n + \mu \log(L), \tag{10}$$

so that the marginal contribution of an extra alternative is decreasing as the number of alternative increases. To sum up, the expected utility is an increasing and concave function of the number of alternatives, L.

2.2 Information available *ex-ante* (and *ex-post*)

This case will be referred to *full information*. Here, the decision maker knows beforehand the value of the idiosyncratic terms. Therefore, in this case, the *ex-ante* value of the utility of the choice set is the same as its *ex-post* value. This common

value is equal to:

$$\tilde{U}_n\left(\mathcal{L};\mathbf{e}_n\right) = \max_{i\in\mathcal{L}} \tilde{U}_n\left(i\right). \tag{11}$$

This common value is also equal to the *ex-post* value of the choice set in the previous case, when information is revealed only *ex-post*. By contrast, the *ex-ante* value of the choice set is different depending on whether the decision maker has information *ex-ante* or not. *Ex-ante* information has a strictly positive value because it helps the decision maker selecting her choice set, which is illustrated in the following very simple example, sum up in Table (1).

[Insert Table 1]

Consider two choice sets containing two alternatives each. The expected utility $V_i,\ i = 1,...,4$ is equal to 0 for the two alternatives of the first choice set, and 1 for the two alternatives of the second choice set. The realizations $e_i,\ i = 1,...,4$ of the random terms are $(2;1)$ for the first choice set and $(0;-1)$ for the second one. The utilities $U_i,\ i = 1,...,4$ are then $(2;1)$ in the first choice set and $(1;0)$ in the second choice set. If she has no information *ex ante*, the decision maker ignores the realizations e_i of the random variables. Therefore, since these realizations are i.i.d. across alternatives, the decision maker will select the second choice set (with the highest expected utilities V_i) at the first stage of the choice process. Once this choice set is selected, the decision maker can observe the realizations in this choice set, and therefore she will select the third alternative (since it has the highest utility U_i in the chosen choice set).

On the opposite, if the decision maker has information *ex ante*, she will select the first choice set and the first alternative. In the case, the decision maker acts as if she were facing a one-step choice process, i.e. as if she were selecting in the full set of alternatives (alternative 1).

In the *learning* case, the choice set selected *ex ante* often differs (as in the above example) from the choice set which would be the best *ex post*. The decision maker selects in the second step the alternative which is the best *ex post* in the small choice set selected in the first step, but the alternative finally selected is usually not optimal *ex post* in the full choice set. This implies that, in the *learning* case, A suboptimal alternative is selected when the alternative which is optimal *ex post* is

not included in the choice subset which is optimal *ex ante*.

The individual-specific value of the choice set in the *full information* case holds for the decision maker n characterized by the realization $\mathbf{e}_n \equiv (e_1, ..., e_L)$. This value is known *ex ante* to the decision maker, but it is unknown *ex-ante* (and *ex-post*) to the modeler. The modeler does not know the realizations of idiosyncratic error terms for a given individual, but she knows the probability that the individual idiosyncratic terms lie in the infinitesimal L-dimensional hypercube $d(\mathbf{e}_n) \equiv de_1 de_2 ... de_L$ is $f(e_1) f(e_2) .... f(e_L) de_1 de_2 ... de_L$. Therefore, the expected utility of choice set $\mathcal{L}$ for individual n, computed by the modeler (who ignores the value of the idiosyncratic tastes), is:

$$\int_{\mathbb{R}^L} \tilde{U}_n(\mathcal{L}; \mathbf{e}_n) f(e_1) f(e_2) .... f(e_L) de_1 de_2 ... de_L. \tag{12}$$

The probability, computed by the modeler, that individual n chooses alternative i in the choice set $\mathcal{L}$ corresponds to the probability that the vector $\mathbf{e}_n$ lies in the region such that $V_n(i) + e_n(i) > V_n(i') + e_n(i')$ for all i' in $\mathcal{L}$. It is given by the MNL formula given in (3). This formula holds both when the decision maker chooses in the full set $\mathcal{L}_0$ and when she chooses in a subset $\mathcal{L}_i$.

When the decision maker has access to information *ex ante*, the modeler knows that she will select the same choice set and alternative than if she were selecting in the full set of alternatives, that is, if she were choosing in a single step. As a consequence, when the decision maker has access to information *ex ante*, the choice probabilities computed by the modeler are the same as if the decision maker were choosing in a single step.

2.3 No information *ex-ante* and *ex-post*

In the last case, referred to as the *no information* case, the decision maker knows the distributions of the idiosyncratic terms, but ignores the realizations of these terms both before and after selecting her choice set. Since the decision maker does not acquire any new information after selecting her choice set, the decision process can be clearly reduced to a one-step procedure.

The *no information* case can also be interpreted as a situation where the decision maker choice is the outcome of an embedded matching model. In this case, the decision maker devotes some level of effort in order to determine an optimal quality of

matching. This quality determines the matching probabilities, i.e. the probabilities that a given alternative is matched with the decision maker. The level of effort is optimized such that the risk of unsatisfactory matching is minimized. The exact procedure remains in this article a black box which has an output given by the MNL probabilities:

$$\mathbb{P}_n\left(i|\mathcal{L}\right) = \frac{\exp\left(\frac{V_n(i)}{\mu}\right)}{\sum\limits_{i'\in\mathcal{L}} \exp\left(\frac{V_n(i')}{\mu}\right)},\ i \in \mathcal{L}. \tag{13}$$

This formula holds both when the decision maker chooses in the full set ($\mathcal{L} = \mathcal{L}_0$) and when she chooses in a nest ($\mathcal{L} = \mathcal{L}_i$).

The probability that the decision maker is allocated to nest $\mathcal{L}_i$ can be computed as the sum of the probabilities that she is allocated to each alternative in the nest $\mathcal{L}_i$:

$$\mathbb{P}_n\left(\mathcal{L}_i|\mathcal{L}_0\right) = \frac{\sum\limits_{i''\in\mathcal{L}_i} \exp\left(\frac{V_n(i'')}{\mu}\right)}{\sum\limits_{i'\in\mathcal{L}_0} \exp\left(\frac{V_n(i')}{\mu}\right)}, \mathcal{L}_i \subset \mathcal{L}_0. \tag{14}$$

It is easy to check, from the above formulae, that $\mathbb{P}_n\left(i|\mathcal{L}_0\right) = \mathbb{P}_n\left(i|\mathcal{L}_i\right)\mathbb{P}_n\left(\mathcal{L}_i|\mathcal{L}_0\right)$, which means that, in the no information case, the probability that a decision maker choosing in two stages is allocated to a given alternative is the same than if she were choosing in a single step. The same probability can be computed by the modeler, who has access to the same information as the decision maker.

Obviously, the case involving information *ex ante* and no information *ex post* does not make sense, since *ex ante* information is included in *ex post* information.

2.4 One-step versus two-step decision process: MNL versus Nested Logit

To sum up the three cases, the decision is achieved in one step when information is the same *ex ante* and *ex post*, and in two steps when information is acquired *ex post* (*learning* case). The latter case, corresponding to a two-step nested decision, is relevant when the number of alternatives is so large that the decision maker is not able to acquire information on all alternatives. In this case, she selects a choice set in the first step based on the *ex ante* value of this choice set, and then, in the second step, she acquires information about the alternatives contained in the choice set that she has selected *ex ante*, in the first step (and not on alternatives included

in other choice sets).

In the *full information* case, the choice is deterministic for the decision maker, but probabilistic and given by the MNL formula (3) for the modeler who does not know the value of the idiosyncratic terms.

In the *no information* case, the choice is random both for the decision maker and for the modeler, and the choice probabilities are given by the same MNL formula (3).

The fact that the modeler computes exactly the same probability in the *full information* and in the *no information* cases implies that it is not possible to test one of these two models against the other. In other words, the data contain no information which could be used to test whether the decision maker has full information or no information.

On the opposite, the probability computed by the modeler is different in the *learning* case because the modeler knows that the decision maker will acquire information *ex post* in the *learning* case, even though the modeler herself does not acquire information *ex post*. In the *learning* case, the decision maker selects for sure in the first step the choice set which is the best *ex ante* for her, i.e. which maximizes the surplus computed as in (7). For the subset $\mathcal{L}_{i'}$, the surplus is given accordingly by:

$$S_n\left(\mathcal{L}_{i'}\right) \equiv \mathbb{E}U_n\left(\mathcal{L}_{i'}\right) = \mu_1 \log\left\{\sum_{i''\in\mathcal{L}_{i'}} \exp\left(\frac{V_n\left(i''\right)}{\mu_1}\right)\right\}, \tag{15}$$

where μ_1 is proportional to the standard deviation of the idiosyncratic terms within $\mathcal{L}_{i'}$.

It is assumed that the decision maker preferences differ as to their perception of the quality of the subset. The corresponding idiosyncratic terms are assumed to be i.i.d. Gumbel distributed with zero mean and a standard deviation proportional to μ_2. Using the same reasoning as before, the probability, computed by the modeler, that the decision maker selects subset $\mathcal{L}_i,\ i = 1, ..., I$ is

$$\mathbb{P}_n\left(\mathcal{L}_i\right) = \frac{\exp\left(\frac{S_n(\mathcal{L}_i)}{\mu_2}\right)}{\sum_{i'=1}^{I} \exp\left(\frac{S_n(\mathcal{L}_{i'})}{\mu_2}\right)},\ i = 1, ..., I.$$

Clearly, when $\mu_1 = \mu_2$, the NL probabilities reduce to the MNL probabilities. Inter-

estingly, this means that the test of the equality $\mu_1 = \mu_2$ amounts to test that there is no information acquisition from the first to the second decision stage.

The framework developed above will now be applied to a multilevel nested model in which individuals decide upon their residential location, workplace, and job type, and may acquire information at each choice step. Each level in this nested model is described by a MNL model, so the the full model is a multi-level NL model when ths scale parameters differs, and a MNL model when the scale parameters are equal, i.e. when there is no information acquisition.

3 A Nested Logit model for job type, workplace and residential location

Individual n chooses a residential location i, a workplace j and a specific job l of type k in a set denoted by $\mathcal{E}_n$. There are I locations and K job types. Individual utility, denoted by U_n, is equal to:

$$U_n(l,k,j,i) = U_n^T(l,k) + U_n^W(j) - C_n^{WR}(j,i) + U_n^R(i) \quad \forall \quad (l,k,j,i) \in \mathcal{E}_n, \qquad (16)$$

where $U_n^T(l,k)$, $U_n^W(j)$, and $U_n^R(i)$ denote the utility specific to job l of type k, the utility specific to the workplace j, and the utility of living in residential location i, respectively. The term $C_n^{WR}(j,i)$ captures the generalized commuting cost between residential location i and workplace j.

The model concentrates on two major choices: the selection of a specific job, including its type and location, and the choice of a residential location. These choices are analyzed by a three-stage model solved by backward induction (Figure 1). At the lower level, individual n chooses a specific job l of type k, conditional on workplace j and residential location i. At the middle level, individual n chooses a workplace j, conditional on residential location i and anticipating job l of type k. Finally, at the upper level, individual n chooses a residential location i, anticipating the work related choices (j,k,l).

[Insert Figure 1]

Additive separability between the deterministic and stochastic components of the utility is imposed at each level, like in the previous section. The utility $U_n^T(l,k)$

provided by job l of type k, in equation (16), is decomposed into a deterministic term $V_n^T(k)$ depending on type k and two random terms, $\varepsilon_n^0(l)$ and $\varepsilon_n^1(k)$ depending on specific job l and on type k, respectively. The deterministic term $V_n^T(k)$ represents the intrinsic preferences of individual n for job type k. A deterministic term specific to the utility of performing a specific job l could be added if job characteristics could be observed. This would add a level into the tree. Under *full information*, the random terms $\varepsilon_n^0(l)$ and $\varepsilon_n^1(k)$ represent the idiosyncratic preference of individual n for the specific job l, and for the job type k, respectively. In the *no information* and *learning* cases, these random terms rather represent job-specific and type-specific characteristics unknown to the decision maker before she chooses her workplace and job type. In the *learning* case, after selecting job type k, the decision maker acquires information about the realization of $\varepsilon_n^0(l)$ for all the jobs of type l located in j.

The deterministic terms $V_n^W(j)$ and $V_n^R(i)$ measure the intrinsic preference for working in j and living in i, respectively. The choices of residential location and workplace are de facto related through the generalized commuting cost $C_n^{WR}(j,i)$ and cannot be assumed independent. Under *full information*, the random terms $\varepsilon_n^2(j)$ and $\varepsilon_n^3(i)$ correspond to the idiosyncratic preference (unobserved heterogeneity of preferences) of individual n for working in j and living in i. An additional random term could be considered explicitly for the generalized commuting cost but it would be impossible to disentangle it from $\varepsilon_n^2(j)$. As a consequence, $\varepsilon_n^2(j)$ also includes idiosyncratic preference for commuting between i and j. In the *no information* case, $\varepsilon_n^2(j)$ and $\varepsilon_n^3(i)$ rather represent local characteristics unknown to the decision maker throughout the decision process. In the *learning case*, $\varepsilon_n^2(j)$ and $\varepsilon_n^3(i)$ also represent local characteristics and $\varepsilon_n^3(i)$ is observed by the decision maker at the first stage of her decision process, whereas she acquires information about $\varepsilon_n^2(j)$ only after selecting residential location i (which is plausible for commuting costs).

The random terms $\varepsilon_n^\iota(\cdot)$, $\iota = i,j,k,l$, are independent from each other for a given individual n and independent across individuals.

To sum up, utility $U_n(l,k,j,i)$ can be decomposed as:

$$U_n(l,k,j,i) = V_n^T(k) + \varepsilon_n^0(l) + \varepsilon_n^1(k) + V_n^W(j) - C_n^{WR}(j,i) + \varepsilon_n^2(j) + V_n^R(i) + \varepsilon_n^3(i)$$
$$\forall \quad (l,k,j,i) \in \mathcal{E}_n. \tag{17}$$

3.1 Lower Level Choice: Specific Job and Job Type

The additive separability assumed in (17) means that the preference of individual n for a specific job l of type k is assumed independent from the job location. This preference may be related, for example, to the wage, the number of working hours and other working conditions. All these characteristics vary significantly across job types and these job characteristics, or the utility attached to these job characteristics, depend on individual characteristics such as gender, education or age. This is the reason why $V_n^T(k)$, $\varepsilon_n^0(l)$ and $\varepsilon_n^1(k)$ are indexed by n. We assume that the difference between the utilities of two job types for a given individual is the same whatever their location. This is the reason why $V_n^T(k) + \varepsilon_n^0(l) + \varepsilon_n^1(k)$ does not depend on job and residential locations j and i. This does not exclude that job characteristics, or the utility attached to these characteristics, vary geographically. For example, wages may be systematically higher in a location j than in location j', for all job types. These geographical differences, if any, can be included in $V_n^W(j) + \varepsilon_n^2(j)$, provided they are homogeneous across job types. The only restriction we impose is that we exclude the case where geographical differences would be specific to job type. To be more specific, we exclude, for example, the case where wages would be systematically higher in j then in j' for blue collars, but wages would be statistically identical in j and in j' for white collars. This interpretation holds in the full information case, but a similar one could be provided in the *learning* and *no information* cases. As a result, the choice between the various jobs located in j only depends on individual characteristics and job types, and is not affected by local observed or unobserved characteristics of workplace and/or residential location.

Let $\mathcal{T}_{kj}$ denote the set of jobs of type k available to an individual n in location j, with $\left|\mathcal{T}_{kj}\right| = N_{kj}$. Since the deterministic part of the utility $V_n^T(k)$ depends only on job type k, but not on the specific job l, all the jobs in $\mathcal{T}_{kj}$ have the same probability $1/N_{kj}$ to be selected, and equality (10) implies that the expected value of job type

k in location j is

$$\mathbb{E}U_n\left(\mathcal{T}_{kj}\right) = V_n^T\left(k\right) + \mu^0 \log\left(N_{kj}\right), \tag{18}$$

where μ^0 denotes the scaling factor of $\varepsilon_n^0\left(l\right)$.

The probability that individual n chooses job type k given workplace j is then equal to

$$\mathbb{P}_n^1\left(k\right) = \frac{\exp\left(\frac{V_n^T(k)+\mu^0 \ln\left(N_{kj}\right)}{\mu^1}\right)}{\sum\limits_{k'=1,...,K;N_{k'j}>0} \exp\left(\frac{V_n^T(k')+\mu^0 \ln\left(N_{k'j}\right)}{\mu^1}\right)}, \quad \forall \quad k = 1,...,K; N_{kj} > 0, \tag{19}$$

where μ^1 denotes the scaling factor of $\max\limits_{l\in\mathcal{T}} \varepsilon_n^0\left(l\right) + \varepsilon_n^1\left(k\right)$, which is assumed to follow a Gumbel distribution in the *learning* case. The ratio $\frac{\mu^0}{\mu^1}$ then corresponds to the ratio of the standard error of idiosyncratic preferences at the job-specific level and at the job-type level, that is the relative intensity of unobserved preferences between jobs of the same type and between job types.

In the two other cases (*full information* and *no information*), $\mu^0 = \mu^1$, and the coefficient of $\ln\left(N_{kj}\right)$ simplifies to 1. This result can be checked as follows. When the decision maker acquires no information after choosing her nest (here job type), her choice probabilities are the same as if she were choosing in one step in the full choice set, and given by the MNL formula. In this case, the probability of specific job l is given by

$$\mathbb{P}_n^0\left(l,k\right) = \frac{\exp\left(\frac{V_n^T(k)}{\mu^1}\right)}{\sum\limits_{k'=1,...,K,N_{k'j}>0} \left(\sum\limits_{l\in\mathcal{T}_{k'j}} \exp\left(\frac{V_n^T(k')}{\mu^1}\right)\right)}, \tag{20}$$

and the choice probability of job type k given workplace j is equal to

$$\begin{aligned}\mathbb{P}_n^1\left(k\right) &= \sum_{l\in\mathcal{T}_{k'j}} \frac{\exp\left(\frac{V_n^T(k)}{\mu^1}\right)}{\sum\limits_{k'=1,...,K,N_{k'j}>0} \left(\sum\limits_{l\in\mathcal{T}_{k'j}} \exp\left(\frac{V_n^T(k')}{\mu^1}\right)\right)} = \frac{N_{kj}\exp\left(\frac{V_n^T(k)}{\mu^1}\right)}{\sum\limits_{k'=1,...,K,N_{k'j}>0} \left(N_{k'j}\exp\left(\frac{V_n^T(k')}{\mu^1}\right)\right)} \\ &= \frac{\exp\left(\frac{V_n^T(k)+\mu^1 \ln\left(N_{kj}\right)}{\mu^1}\right)}{\sum\limits_{k'=1,...,K,N_{k'j}>0} \left(\exp\left(\frac{V_n^T(k')+\mu^1 \ln\left(N_{k'j}\right)}{\mu^1}\right)\right)}.\end{aligned} \tag{21}$$

Allowing μ^0/μ^1 to vary across individual types (and then be denoted by μ_n^0/μ_n^1)

amounts to considering that relative intensity of unobserved preferences between jobs of the same type and between job types varies across individuals. Probability (21) then becomes:

$$\mathbb{P}_n^1(k) = \frac{\exp\left(\delta_n^1 + \delta_n^0 \ln\left(N_{kj}\right)\right)}{\sum_{k'=1,\ldots,K, N_{k'j}>0} \exp\left(\delta_n^1 + \delta_n^0 \ln\left(N_{k'j}\right)\right)}, \tag{22}$$

with $\delta_n^0 = \dfrac{\mu_n^0}{\mu_n^1}$ and $\delta_n^1 = \dfrac{V_n^T(k)}{\mu_n^1}$.

Interestingly, the choices of job type k and job location j are related only through the number N_{kj} of jobs of type k in location j, denoted by N_{kj}. This result is a direct implication of the assumption of additive separability of utility in Equation (17).

In the case of job type choice conditional on workplace, the surplus (15), or expected utility resulting from the choice of the best job type conditional on workplace j is denoted by $S_n(j)$, and is equal to:

$$\begin{aligned} S_n(j) &= \mu_n^1 \ln\left(\sum_{k=1,\ldots,K; N_{kj}>0}^{K} \exp\left(\frac{V_n^T(k) + \mu^0 \ln\left(N_{kj}\right)}{\mu^1}\right)\right) \\ &= \mu_n^1 \ln\left(\sum_{k=1,\ldots,K; N_{kj}>0}^{K} \exp\left(\delta_n^1 + \delta_n^0 \ln\left(N_{kj}\right)\right)\right). \end{aligned} \tag{23}$$

It measures the attractiveness of workplace j.

3.2 Middle Level Choice: Workplace Location

Let $\mathcal{I}$ denote the set of all potential (residential or workplace) locations, with $|\mathcal{I}| = I$. These locations are assumed available for each individual both for working and for living, so $(j,i) \in \mathcal{I}^2$. Considering the decision tree assumed here, an individual n will choose a workplace j conditional on her current residential location i, and so actual travel time is relevant for explaining workplace location and the generalized travel cost, $C_n^{WR}(j,i)$, is considered here, in the middle level choice.

Using the assumptions above, from equation (16), the utility of workplace location j, including generalized commuting cost, $C_n^{WR}(j,i)$, between residential location i and workplace j, can be expressed as:

$$U_n^W(j) - C_n^{WR}(j,i) = V_n^W(j) - C_n^{WR}(j,i) + \varepsilon_n^2(j) \qquad \forall \quad j \in \mathcal{I}. \tag{24}$$

Similarly to the lower level, in the *full information* case, the error term $\varepsilon_n^2(j)$ representing the idiosyncratic preference of individual n attributable to workplace j, is distributed so that the random part of $\max_{(k,l)} U_n^T(l,k) + \varepsilon_n^2(j)$ has a Gumbel distribution with scale parameter μ_n^2 specific to individual n (See the discussion about μ_n^1 below equation (21)). The probability of choosing workplace location j is then equal to:

$$\mathbb{P}_n^2(j) = \frac{\exp\left(\frac{V_n^W(j;) - C_n^{WR}(j,i) + S_n(j)}{\mu_n^2}\right)}{\sum_{j' \in \mathcal{I}} \exp\left(\frac{V_n^W(j') - C_n^{WR}(j',i) + S_n(j')}{\mu_n^2}\right)} \quad \forall \quad j \in \mathcal{I}. \tag{25}$$

In the case of workplace choice conditional on residential location, the surplus (15), or expected utility resulting from the choice of the best workplace conditional on residential location i is denoted by $LS_n(i)$, and is equal to:

$$LS_n(i) = \mu_n^2 \ln\left(\sum_{j \in \mathcal{I}} \exp\left(\frac{V_n^W(j) - C_n^{WR}(j,i) + S_n(j)}{\mu_n^2}\right)\right). \tag{26}$$

It measures the accessibility to jobs of residential location i.

3.3 Upper Level: Residential Location

In the upper level of the decision tree, the individual anticipates the workplace, job type and specific job choices when she chooses her residential location. The utility of living in residential location i (equation (16)) is:

$$U_n^R(i) = V_n^R(i; X_n, Z_i) + \varepsilon_n^3(i) \quad \forall \quad i \in \mathcal{I}. \tag{27}$$

The residual term $\varepsilon_n^3(i)$ accounts for the idiosyncratic preference of individual n for residential location i. It expresses unobserved location attributes, variation in individual tastes, and model misspecification. Similarly to other levels, this residual term is distributed so that the random part of $\max_{(j,k,l)} U_n^T(l,k) + U_n^W(j) - C_n^{WR}(j,i) + \varepsilon_n^3(i)$ is type I extreme value distributed with scale parameter μ_n^3. The probability

of choosing residential location i is then:

$$\mathbb{P}_n^3(j) = \frac{\exp\left(\frac{V_n^R(i; X_n, Z_i) + LS_n(i)}{\mu^3}\right)}{\sum_{i' \in \mathcal{I}} \exp\left(\frac{V_n^R(i'; X_n, Z_{i'}) + LS_n(i')}{\mu^3}\right)} \quad \forall \quad i \in \mathcal{I}. \tag{28}$$

4 Data, methodology and results

4.1 Greater Paris Data

The model is estimated using exhaustive household data from the last French General Census conducted in 1999 in Ile-de-France Region (IDF). In this census, job type and individual characteristics are observed for 100% of the population, corresponding to about 11 million inhabitants and 5 million households. Residential location and workplace are observed at the commune (municipality) level for a 5% sample of the working population; that is around 240,000 workers in 1999. The commune is the smallest administrative unit used in France. The IDF region is composed by 1,300 communes, of which 20 form the central city of Paris. The 1,300 communes are grouped into eight departments or districts, the central one corresponding to Paris. The central city of Paris accounts for about 2 million inhabitants. The inner ring (close suburbs) is made of three districts, while the outer ring is composed by four districts (Figure 2).

[Insert Figure 2]

The study area exhibits spatial disparities in the supply of jobs. In particular, many outer ring communes have little or no job supply. Almost 25% of the 1,300 communes (almost entirely located in the outer ring) are very small communes in terms of number of jobs (Figure 3). In order to circumvent this small number problem, small adjacent communes were grouped into "pseudo-communes" following a simple pairwise aggregation strategy until each pseudo-commune contained at least 100 jobs. The resulting 950 pseudo-communes with 100 jobs or more were used as unit of location for both jobs and residence.

[Insert Figure 3]

The census data was aggregated at the pseudo-commune level, and variables measuring prices and local amenities were computed at the same level. Price data originally come from Cote Callon, which reports average prices per m^2 for offices and dwellings by type and tenure status for communes with more than 5,000 inhabitants (287 communes, each of them containing at least 100 jobs). Hedonic price regressions were estimated jointly for each tenure and dwelling type as well as for offices, and the results were used to predict prices in smaller pseudo-communes. Palma et al., 2007 provide a detailed description of local amenities, of price equations and of the method used for correcting the endogeneity of prices.

Origin-Destination (O-D) matrices of travel time using public transportation were obtained from the Regional Department of Infrastructure and Transportation Planning (DRIEA) transport model MODUS, whereas O-D matrices of travel time for private car were computed using the dynamic transport network model METROPOLIS described in (Palma et al., 1997). The transport model METROPOLIS has a dynamic traffic simulator which uses a mesoscopic approach: vehicles are simulated individually while the traffic dynamics is modeled at the aggregate level. The disaggregate representation of demand allows to consider the heterogeneity of the population and trips. Saifuzzaman et al. (2012) provide a more detailed information on METROPOLIS calibration for the Paris Region.

The sample analyzed contains 239,499 people living and working in Ile-de-France. The lower and middle level models (job type and workplace) are estimated separately in 24 subsamples in order to reflect how individual preferences and job opportunities depend on age, education, gender, and fertility. More precisely, the sample is split in two age groups of approximately equal size, the "young" being less than 35 years. Education groups (elementary; secondary; undergraduate; graduate) were defined according to preliminary results measuring the influence of education on the job type choice. Similarly, the influence of fertility on female job type choice was measured by a dummy variable indicating whether the woman considered has at least one child less than 12 years old. The combination of education, age, gender and fertility categories results in 24 categories (Table 2).

[Insert Table 2]

The results of the three models (residential location, workplace, and job type) outlined in Section 2 are presented after some methodological considerations.

4.2 Methodological considerations

Since the nested logit is estimated sequentially by backward induction, the first model estimated corresponds to the job type choice, the second one corresponds to workplace choice, and the last model estimated corresponds to residential location choice. A Multinomial Logit (MNL) model is estimated at each level, including a log-sum variable at the middle and upper levels.

Given the large number of alternatives (950 pseudo-communes), it would have been too cumbersome to consider all the 950 alternatives at the middle and upper levels (job and residential location choices). This problem can be circumvented by using random sampling, which consists in randomly selecting a small number of alternatives for each individual, with equal probabilities of selection in the choice set . McFadden (1978) showed that random sampling leads to consistent estimates of the coefficients a MNL model under the IIA assumption. Ben-Akiva et al. (1985) further showed that importance sampling improves the efficiency of the estimates. Importance sampling consists in increasing the probability that a given alternative is included in the choice set. Ben-Akiva et al. (1985) also show that importance sampling usually induces a bias in the coefficients, which should be corrected. In our case, the probability that a pseudo-commune is included in the choice set is proportional to the number of dwellings (upper level) or jobs (middle level) in that pseudo-commune. Since no information is available concerning the dwellings (intrinsic) characteristics, all dwellings of the same type and tenure status located in the same pseudo-commune are statistically identical and provide the same expected utility and the same odds of being selected by a specific household. Similarly, since information is available concerning the jobs characteristics, all jobs of the same type located in the same pseudo-commune are statistically identical. As a consequence, the importance sampling of pseudo-communes considered here is equivalent to uniform random sampling of dwellings (or jobs), so the coefficients are not biased and no correction is needed.

The strong segmentation of the dwelling market in France has two implications

here. First, prices vary across dwelling types and tenure status, so it is more relevant to consider dwelling prices specific each of the resulting four sub-markets. Second, at a given point in the life cycle, a given household is usually not flexible concerning dwelling type and tenure status, which adds a level at the top of the tree (Figure 4). Modeling endogenous choice of dwelling type and tenure status is out of the scope of this paper. Instead, residential location is simply estimated separately in the four sub-samples defined by dwelling type and tenure status, which amounts to assuming that dwelling type and tenure status is exogenous to residential location choice. Getting rid of this implicit assumption is left for future research.

[Insert Figure 4]

Finally, the residential location choice is restricted to one-worker households, in order to avoid the bargaining considerations that would arise in a multiple-worker household. In such households, the jobs of the different workers are usually located in different places, and each household member has to bargain so that the household locates closer to his/her job.

4.3 Job Type Choice

A multinomial logit (MNL) model is estimated for each of the 24 categories. This is, 24 different MNL choice models between the following job types: blue collar, employee, intermediate, manager and independent. Given the small number of alternatives (5 job types), no random sampling is needed at this level.

The results of the MNL model of job type are presented in Table 3. The reference alternative is blue collar. Almost all the estimated coefficients by job type are highly significant. The measure of goodness of fit presented in the last column of Table 3 suggests that the explanatory power increases with education for men. This suggests that the less educated men accept any job and are randomly assigned to job types such as blue collar, employee, or independent. By contrast, the most educated men typically only accept the job types best suited to them (manager, intermediate and independent). The role of education is more ambiguous for women. Conditional on age and fertility, what influences most the decision to work or not for a woman is the education rather than the choice of job type.

[Insert Table 3]

Based on this lower-level estimates, a log-sum can be computed. It corresponds to the sum of the log-number of jobs by type, weighted by the individual-specific probability to choose a particular job type. This individual-specific measure of attractiveness, defined in equation (23), varies between job locations and between individual characteristics. The job type attractiveness of each workplace, computed as the log-sum across local job types, is represented in Figures 5 and 6 by gender and education.

[Insert Figures 5 and 6]

4.4 Workplace Location Choice

The only criteria to choose a workplace location considered here are the generalized commuting cost $C_n^{WR}(j,i)$ and the availability of jobs of different types that can be found in the job type choice inclusive value or surplus term $S_n(j)$. Past empirical works have included average wage by job type as an explanatory variable of workplace location choice (Waddell (1993)). Any significant difference in wage between locations in the geographical units of study is already in the job type surplus term $S_n(j)$, because this term allows us to account for differences in the employment structure between different workplace locations. With the purpose of constructing a parsimonious workplace choice model the $V_n^W(j; X_n, Z_j)$ term is considered nil.

The workplace location choice of a pseudo-commune is considered to depends on its job type attractiveness (the individual-specific measure computed in the job type choice model) and the commuting travel time of individuals. For this second choice level, MNL models are estimated separately for each of the 24 categories described. The results of the 24 workplace location choice models are presented in Table 4.

[Insert Table 4]

In the Attractiveness column of Table 4 the association between the measure of attractiveness and the workplace location is explored. The estimated coefficients indicate that the most educated and older men are more sensitive to the job type attractiveness of the workplace than the younger and less educated. Women (especially the more educated) are less sensitive than men to the job type attractiveness.

Columns "Travel time" and "(Travel time)^2" allow for a quadratic specification of travel time. The results suggest that the workplace location utility is decreasing and concave in travel time for each of the 24 groups. The value of time depends then on age, education, gender and children.

In order to explore further the gain of using a job type attractiveness measure, 24 workplace location choice models were estimated using a size measure (log number of jobs) instead of the measure of job type attractiveness chosen here, while keeping the quadratic specification of travel time. The last column of Table 4 presents the difference between the Likelihood Ratio (LR) of the workplace location choice model estimated with the attractiveness measure and the LR of the models estimated with the size measure instead. Results indicate that the measure of attractiveness (specific to each individual) is a better predictor of the workplace location choice than the size measure commonly used (log of the number of jobs).

This choice level allows us to develop an accessibility measure specific to each individual: the log-sum of workplace locations. That is to say, the expected maximum utility of all job opportunities. This measure varies between residential location of households and individual characteristics. Accessibility differs across groups because local employment prospects and the value of time differ across them. The computed measure of accessibility to jobs has been mapped in Figures (7) and (8) by gender and education in the Appendix. Difference of accessibility are particularly strong as the education level of individual increases (Figure 8).

[Insert Figures 7 and 8]

4.5 Residential Location Choice

The results presented in Tables 6 and 7 are limited to households with only one worker. In households with more workers, the choice of residential location and workplace is modified by the negotiation process within the household. Bargaining considerations are left for future work. The location model is estimated separately by tenure type (owner and tenant) and dwelling type (apartment and single dwelling). Sample sizes by tenure and dwelling type are presented in Table 5.

[Insert Table 5]

In the last row of Tables (6) and (7) the measure of goodness of fit is presented. The explanatory power is higher for owners than for renters. This is consistent with the fact that purchasing decisions are much more developed or matured (and so less random) than renting decisions. Similarly, the explanatory power is higher for the choice of single dwellings than for the choice of apartments. This result is consistent with the rotation rates, which are higher for renters than for owners, and for apartments than for single dwellings. Location decisions are more thoughtful when it regards the longer term.

[Insert Table 6]

From the accessibility and transport estimated coefficients and presented in Table 6, when comparing between ownership status and dwelling types, owners are more sensitive to accessibility than tenants; and sensitivity to accessibility is more pronounced for households living in apartments than for those living in a single dwelling. These results are consistent with considerations of life cycle and geographical distribution of single dwellings and apartments. In the early stages of the life cycle, when jobs are less stable and when households do not have children yet, households usually rent an apartment strategically located in relation to potential jobs. At later stages of the life cycle, when employment stabilizes and couples have children, households buy single dwellings that are usually far away (and less accessible) in the suburbs. In the decision-making process of choice of residence, the more the households move through their life cycle, the more they are willing to sacrifice accessibility to jobs to access to ownership and gain in residence space.

The numbers of subways and suburban train stations (RER and SNCF suburban trains) only attract households who rent apartments. For other households, the effect is ambiguous or insignificant, which is logical in the single-worker household sample used here.

The results of the influence of price in the residential location choice can be found in the lower rows of Table 6. For households with an average income, the price has a negative impact on the probability of location, with the exception of households that rent a single dwelling (very small sample). The negative effect of price decreases with income, and may be positive for the richest households.

To test for the influence of unobserved local amenities, regional dummies are considered. All other things being equal, an apartment in the outer ring has a lower probability of being selected, and conversely a single dwelling in the outer ring has a higher probability to be chosen. Similarly, apartments located in a Planned City have greater probability of being selected, while there are no significant differences between Planned Cities and the other locations for single dwellings. All other things being equal, an apartment in Paris has a lower probability of being selected, which may seem surprising at first sight. However, this can be explained by the fact that the reasons why Paris attract households are already taken into account by other explanatory variables in the model (number of subway stations and accessibility to jobs are particularly favorable for Paris).

The fourth group of explanatory variables taken into consideration and the last group presented on Table 6 are the local taxes variables. The effect of the residence tax (for ownership and tenancy) and property taxes (for ownership) is ambiguous. Higher taxes have a direct negative effect, but they are usually associated with local services (such as child care center or streets amenities, not measured here), which exert an attractive effect.

[Insert Table 7]

The second table of results dedicated to the residential location choice model (Table 7), displays the coefficients measuring the influence of local amenities, and population composition variables. As expected, the probability of choosing a commune increases with the density for households living in an apartment, and decreases for those living in a single dwelling, as well as for ownership with respect to tenancy. The other usual local amenity variables present the expected signs.

Variables related to social mix are among the most significant for explaining residential location choice: households are attracted by households with similar characteristics regarding age, size, and income. Single dwelling owners are more attracted by communes with a high percentage of foreigners, which can be explained by the fact that the (rich) foreigners who settled in the Paris Region tend to buy a dwelling close to their compatriots. Moreover, beyond a threshold, the percentage of foreigners (rather poor) can be seen as a negative characteristic, but the communes

concerned generally have little owners. For renters, the percentage of foreigners has a positive effect, which decreases with education.

5 Conclusion

The choices between residential location, workplace, and job type are modeled here. An econometric framework is developed to study the interdependency between the residential location and workplace, including the attractiveness of jobs by type. This framework provides a way to compute individual-specific measures that are very relevant for public policy analysis: accessibility to jobs, travel time, value of time, and job type attractiveness.

The three-level nested logit model proposed allows for a new concept of accessibility to jobs that takes into account the individual-specific job type attractiveness and the heterogeneity in the preferences of individuals, in the education, age, gender, and children dimensions.

Estimation results show that the job type attractiveness measure is a more significant predictor of workplace than the usual (total number of jobs) measure. It means that workers are not attracted equally by any job, but they are more attracted by the jobs which are better suited to them. This selective attraction of jobs is more pronounced for highly educated men than for less educated men or for women. Results also show that the individual-specific accessibility measure is an important determinant of the residential location choice, and its effect strongly differ along the life cycle.

The model developed here bridges the gap between micro-simulation and general equilibrium urban models. On the one hand, micro-simulation urban models ignore the joint nature of the residential location, workplace, and job type decision. On the other hand, general equilibrium urban models consider only limited heterogeneity. Empirical results draw the attention to the pertinence of considering residential location, workplace, and job type all together and allowing for greater heterogeneity, especially when individual-specific accessibility, attractiveness, and travel time measures are computed for policy study.

References

Abraham, J. E. and J. D. Hunt (1997). "Specification and estimation of nested logit model of home, workplaces, and commuter mode choices by multiple-worker households". *Transportation Research Record* 1606, pp. 17–24.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Harvard University Press.

Anas, A. (1981). "The estimation of multinomial logit models of joint location and travel mode choice from aggregated data". *Journal of Regional Science* 21.2, pp. 223–242.

Anderson, S., A. de Palma, and J.-F. Thisse (1992). *Discrete Choice Theory of Product Differentiation.* Cambridge, MA: The MIT Press.

Ben-Akiva, M. and S. R. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press, Cambridge, Massachusetts.

Ben-Akiva, M. and J. L. Bowman (1998). "Integration of an activity-based model system and a residential location model". *Urban Studies* 35.7, pp. 1131–1153.

Ben-Akiva, M., A. de Palma, D. McFadden, M. Abou-Zeid, P.-A. Chiappori, M. de Lapparent, S. Durlauf, M. Fosgerau, D. Fukuda, S. Hess, C. Manski, A. Pakes, N. Picard, and J. Walker (2012). "Process and context in choice models". *Marketing Letters.*

Bhat, C. R. and J. Guo (2004). "A mixed spatially correlated logit model: formulation and application to residential choice modeling". *Transportation Research Part B: Methodological* 38.2, pp. 147 –168.

Clark, W. A. V. and S. Davies Withers (1999). "Changing jobs and changing houses: Mobility outcomes of employment transitions". *Journal of Regional Science* 39.4, pp. 653–673.

Crane, R. (1996). "The influence of uncertain job location on urban form and the journey to work". *Journal of Urban Economics* 39.3, pp. 342 –356.

Eliasson, J. and L.-G. Mattsson (2000). "A model for integrated analysis of household location and travel choices". *Transportation Research Part A: Policy and Practice* 34.5, pp. 375 –394.

Freedman, O. and C. R. Kern (1997). "A model of workplace and residence choice in two-worker households". *Regional Science and Urban Economics* 27.3, pp. 241 –260.

Kan, K. (1999). "Expected and unexpected residential mobility". *Journal of Urban Economics* 45.1, pp. 72 –96.

— (2002). "Residential mobility with job location uncertainty". *Journal of Urban Economics* 52.3, pp. 501 –523.

Lee, B. and P. Waddell (2010). "Residential mobility and location choice: a nested logit model with sampling of alternatives". *Transportation* 37 (4), pp. 587–601.

Lerman, S. R. (1976). "Location, housing, automobile ownership, and mode to work: A joint choice model". *Transportation Research Record* 610, pp. 6–11.

Levine, J. (1998). "Rethinking Accessibility and Jobs-Housing Balance". *Journal of the American Planning Association* 64.2, pp. 133–149.

Levinson, D. M. (1998). *Accessibility and the Journey to Work.* Working Papers 199802. University of Minnesota: Nexus Research Group.

Linneman, P. and P. E. Graves (1983). "Migration and job change: A multinomial logit approach". *Journal of Urban Economics* 14.3, pp. 263 –279.

Manski, C. F. and S. R. Lerman (1977). "The Estimation of Choice Probabilities from Choice Based Samples". *Econometrica* 45.8, pp. 1977–88.

McFadden, D (1978). "Modeling the residential location". *Transportation Research Record* 673, pp. 72–77.

Mills, E. S. (1972). *Studies in the Structure of the Urban Economy.* The Johns Hopkins Press, Baltimore, Maryland.

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use.* University of Chicago Press.

Palma, A., F. Marchal, and Y. Nesterov (1997). "METROPOLIS : A modular architecture for dynamic traffic simulation". *Transportation Research Record* 1607, pp. 178–184.

Palma, A., K. Motamedi, N. Picard, and P. Waddell (2007). "Accessibility and environmental quality: Inequality in the paris housing market". *European Transport* 36, pp. 47–74.

Palma, A., M. Ben-Akiva, D. Brownstone, C. Holt, T. Magnac, D. McFadden, P. Moffatt, N. Picard, K. Train, P. Wakker, and J. Walker (2008). "Risk, uncertainty and discrete choice models". *Marketing Letters* 19 (3), pp. 269–285.

Saifuzzaman, M., A. de Palma, and K. Motamedi (2012). *Calibration of METROPOLIS for Ile de-France.* SustainCity Working Paper, 7.2. CES, ENS-Cachan, France.

Sermons, M. and F. S. Koppelman (2001). "Representing the differences between female and male commute behavior in residential location choice models". *Journal of Transport Geography* 9.2, pp. 101 –110.

Siegel, J. (1975). "Intrametropolitan migration: A simultaneous model of employment and residential location of white and black households". *Journal of Urban Economics* 2.1, pp. 29 –47.

Simpson, W. (1980). "A simultaneous model of workplace and residential location incorporating job search". *Journal of Urban Economics* 8.3, pp. 330 –349.

Thurstone, L. (1927). "A law of comparative judgement". *Psychological Review* 34, pp. 273–86.

Waddell, P. (1993). "Exogenous workplace choice in residential location models: Is the assumption valid?" *Geographical Analysis* 25.1, pp. 65–82.

Waddell, P., C. Bhat, N. Eluru, L. Wang, and R. M. Pendyala (2007). "Modeling interdependence in household residence and workplace choices". *Transportation Research Record* 2007, pp. 84–92.

Table 1: Simplified Examples

Choice set	Alternative	V_i	e_i	U_i	Chosen Alternative Under *Full Information*	Chosen alternative in the *Learning Case*
1	1	0	2	2	X	
1	2	0	1	1		
2	3	1	0	1		X
2	4	1	-1	0		

Table 2: Sample Size per Category

	Men		Women			
	Young	Old	Young		Old	
Education			With Children	Without Children	With Children	Without Children
Elementary	18,270	36,813	5,002	7,700	5,974	25,577
Secondary	8,551	12,750	2,950	5,883	3,402	11,251
Undergraduate	10,441	10,234	3,354	9,569	3,145	7,791
Graduate	11,091	17,279	2,478	8,549	3,165	8,280

Note: Total sample size of 239,499 working persons. Categorization by sex, age, children, and education resulted in 24 subsamples. A person is considered young if she has less than 35 years old. Also, a woman is categorized as having children if she has at least one child of 11 years old or less.

Source: General Population Census for the Paris Region. INSEE, 1999.

Table 3: Job Type Choice Model

	Job Type Preferences (Reference: Blue Collar)					
Groups	Size Measure	Independent	Managerial	Intermediate	Employee	$\rho^{2[1]}$
Men						
Young						
Elementary	0.8222‡	-1.5332‡	-3.2529‡	-1.7087‡	-1.3610‡	0.30
	(0.0251)	(0.0452)	(0.0490)	(0.0244)	(0.0231)	
Secondary	0.8643‡	-0.9413‡	-1.704‡	-0.3940‡	-0.3618‡	0.16
	(0.0365)	(0.0709)	(0.049)	(0.0323)	(0.0342)	
Undergraduate	0.9574‡	0.0309‡	0.2486‡	1.2597‡	0.4337‡	0.21
	(0.0389)	(0.0792)	(0.0436)	(0.0387)	(0.0432)	
Graduate	0.8261‡	0.9085‡	2.9593‡	1.5861‡	0.5015‡	0.44
	(0.0398)	(0.0990)	(0.0646)	(0.0676)	(0.0748)	
Old						
Elementary	0.8063‡	-0.3283‡	-1.9485‡	-1.0148‡	-1.3334‡	0.13
	(0.0158)	(0.0249)	(0.0221)	(0.0152)	(0.0171)	
Secondary	0.8032‡	0.7045‡	0.3025‡	0.4639‡	-0.3486‡	0.07
	(0.0275)	(0.0479)	(0.0311)	(0.0302)	(0.0354)	
Undergraduate	0.7389‡	1.2021‡	1.5406‡	1.3041‡	-0.3058‡	0.19
	(0.0324)	(0.0627)	(0.0425)	(0.0435)	(0.0548)	
Graduate	0.5935‡	1.8790‡	3.1703‡	1.0478‡	-0.3125‡	0.50
	(0.0296)	(0.0653)	(0.0521)	(0.0574)	(0.0708)	
Women						
Young						
With Children						
Elementary	0.8421‡	-0.6579‡	-2.5086‡	-0.5056‡	1.3346‡	0.47
	(0.0652)	(0.1252)	(0.1372)	(0.0638)	(0.0515)	
Secondary	0.7735‡	0.0085‡	-0.6206‡	1.2581‡	2.1184‡	0.40
	(0.0983)	(0.1903)	(0.1344)	(0.0990)	(0.1000)	
Undergraduate	0.8730‡	0.7312‡	1.5153‡	3.1667‡	2.6266‡	0.37
	(0.0933)	(0.2398)	(0.1554)	(0.1472)	(0.1515)	
Graduate	0.9575‡	1.3300‡	4.1252‡	3.4052‡	1.9893‡	0.42
	(0.0884)	(0.3437)	(0.2526)	(0.2540)	(0.2623)	
Without Children						
Elementary	0.8540‡	-0.6294‡	-2.0835‡	-0.4047‡	1.3296‡	0.45
	(0.0511)	(0.0981)	(0.0921)	(0.0501)	(0.0423)	
Secondary	0.7288‡	-0.5234‡	-0.4125‡	1.1887‡	2.1063‡	0.41
	(0.0679)	(0.1480)	(0.0889)	(0.0699)	(0.0708)	
Undergraduate	0.7167‡	0.0124‡	1.1669‡	2.9215‡	2.7596‡	0.37
	(0.0585)	(0.1543)	(0.0914)	(0.0854)	(0.0885)	
Graduate	0.8210‡	0.8320‡	3.6721‡	3.2444‡	2.3833‡	0.34
	(0.0479)	(0.1790)	(0.1230)	(0.1236)	(0.1265)	
Old						
With Children						
Elementary	0.7458‡	-0.5445‡	-1.8243‡	-0.3663‡	1.2783‡	0.40
	(0.0549)	(0.1019)	(0.0919)	(0.0556)	(0.0474)	
Secondary	0.7786‡	0.8059‡	0.6396‡	1.9021‡	2.2070‡	0.30
	(0.0823)	(0.1628)	(0.1123)	(0.1018)	(0.1044)	
Undergraduate	1.1452‡	1.6729‡	1.9197‡	3.0693‡	1.8303‡	0.34
	(0.0860)	(0.2038)	(0.1495)	(0.1447)	(0.1513)	
Graduate	0.6183‡	1.6659‡	4.1522‡	3.0145‡	1.6371‡	0.45
	(0.0775)	(0.2467)	(0.2030)	(0.2064)	(0.2173)	
Without Children						
Elementary	0.8646‡	0.1262‡	-1.1662‡	-0.1209‡	1.2035‡	0.32
	(0.0250)	(0.0432)	(0.0355)	(0.0259)	(0.0231)	
Secondary	0.8230‡	1.3394‡	1.1889‡	2.2136‡	2.1471‡	0.24
	(0.0415)	(0.0855)	(0.0644)	(0.0610)	(0.0628)	
Undergraduate	0.8785‡	1.6054‡	2.2571‡	3.0564‡	1.8847‡	0.29
	(0.0515)	(0.1233)	(0.0950)	(0.0936)	(0.0980)	
Graduate	0.3816‡	1.6216‡	4.1368‡	2.9735‡	1.9107‡	0.41
	(0.0484)	(0.1428)	(0.1217)	(0.1241)	(0.1300)	

[1] ρ^2 is a measure of goodness of fit defined as the percentage increased in the log-likelihood function above the value taken at zero parameters.

‡ Significant at the 1% level, † Significant at the 5% level, * Significant at the 10% level

Table 4: Workplace Location Choice Model

Groups	Explanatory Variables: Attractiveness	Travel Time	(Travel Time)^2	$\rho^{2\,1}$	ΔLR^2
Men					
Young					
Elementary	-0.0468‡	1.2532‡	-8.4221‡	0.48	-7.0
	(0.0108)	(0.1247)	(0.1907)		
Secondary	0.0634‡	1.7142‡	-8.3713‡	0.38	-1.0
	(0.0141)	(0.1581)	(0.2418)		
Undergraduate	0.0511‡	1.4170‡	-7.0682‡	0.28	6.0
	(0.0102)	(0.1360)	(0.2034)		
Graduate	0.1277‡	1.2599‡	-5.8953‡	0.21	114.6
	(0.0104)	(0.1231)	(0.1741)		
Old					
Elementary	0.0381‡	1.7761‡	-8.7578‡	0.43	1.0
	(0.0076)	(0.0794)	(0.1227)		
Secondary	0.2090‡	1.9071‡	-8.3928‡	0.33	-6.0
	(0.0119)	(0.1288)	(0.1966)		
Undergraduate	0.1510‡	1.8766‡	-7.7952‡	0.29	25.0
	(0.0132)	(0.1365)	(0.2049)		
Graduate	0.2904‡	1.6935‡	-7.1091‡	0.25	272.0
	(0.0119)	(0.1101)	(0.1549)		
Women					
Young					
With Children					
Elementary	0.0425*	0.5459†	-7.9755‡	0.53	-2.0
	(0.0227)	(0.2412)	(0.3928)		
Secondary	0.1970‡	-0.2243	-6.2741‡	0.42	0.1
	(0.0287)	(0.2921)	(0.4595)		
Undergraduate	0.1619‡	-0.4822*	-5.8000‡	0.39	6.8
	(0.0227)	(0.2767)	(0.4285)		
Graduate	0.1078‡	-0.3486	-4.9114‡	0.30	13.9
	(0.0204)	(0.3014)	(0.4424)		
Without Children					
Elementary	0.0648‡	0.5082†	-7.8432‡	0.51	-4.0
	(0.0174)	(0.1986)	(0.3055)		
Secondary	0.2283‡	0.2109	-6.7998‡	0.42	-1.0
	(0.0217)	(0.2116)	(0.3177)		
Undergraduate	0.2453‡	0.3424†	-6.0165‡	0.32	25.0
	(0.0157)	(0.1512)	(0.2248)		
Graduate	0.1120‡	0.5039‡	-5.3448‡	0.25	45.1
	(0.0129)	(0.1484)	(0.2122)		
Old					
With Children					
Elementary	0.1761‡	0.5775‡	-8.1457‡	0.54	-7.0
	(0.0244)	(0.2244)	(0.3590)		
Secondary	0.3019‡	-0.0683	-6.9683‡	0.46	2.5
	(0.0280)	(0.2860)	(0.4470)		
Undergraduate	0.1893‡	0.2121	-6.9253‡	0.42	13.3
	(0.0190)	(0.2815)	(0.4258)		
Graduate	0.2033‡	-0.4168	-5.3695‡	0.35	20.2
	(0.0293)	(0.2838)	(0.3988)		
Without Children					
Elementary	0.1462‡	0.6924‡	-8.2738‡	0.55	-28.0
	(0.0102)	(0.1083)	(0.1699)		
Secondary	0.2961‡	0.3529†	-7.2969‡	0.45	23.0
	(0.0145)	(0.1560)	(0.2371)		
Undergraduate	0.1576‡	0.6057‡	-7.1502‡	0.41	20.0
	(0.0149)	(0.1754)	(0.2578)		
Graduate	0.1705‡	0.6792‡	-6.6626‡	0.36	22.0
	(0.0300)	(0.1684)	(0.2355)		

[1] ρ^2 is a measure of goodness of fit defined as the percentage increased in the log-likelihood function above the value taken at zero parameters.

[2] ΔLR is the difference between the Likelihood Ratio (LR) of the workplace location choice model estimated with the attractiveness measure and the LR of a model estimated with the size measure (total number of jobs): $\Delta LR = LR_{attractiveness} - LR_{size\ measure}$

‡ Significant at the 1% level, † Significant at the 5% level, * Significant at the 10% level. Standard errors in parenthesis.

Table 5: Sample Size by Tenure and Dwelling Type

	Dwelling Type[1]		
Tenure	Apartment	Single Dwelling	Total
Owner	17,047	16,121	33,168
			(37.96%)
Tenant	51,104	3,095	54,199
			(62.04%)
Total	68,151	19,216	87,367
	(78.01%)	(21.99%)	(100%)

[1]All the detached-single unit and semi-detached dwellings are defined as "houses", otherwise the dwellings are defined as "flats".

Note: Sample size of 87,367 one-worker households living and working in the Greater Paris Area.

Source: General Population Census for the Paris Region. INSEE, 1999.

Table 6: Residential Location Choice Mode, I

	Owner		Tenant	
	Apartment	Single Dwelling	Apartment	Single Dwelling
Accessibility and Transport				
Accessibility to Jobs (IV)	0.3024‡	0.0727†	0.4029‡	0.236‡
	(0.0428)	(0.0369)	(0.0236)	(0.0789)
Suburban Train × High-Income	0.0199‡	0.0161*	0.0368‡	0.0291
	(0.0064)	(0.0088)	(0.0054)	(0.0217)
Suburban Train× Middle-Income	0.000662	-0.0649‡	0.0176‡	-0.0449†
	(0.0066)	(0.0109)	(0.0038)	(0.0214)
Suburban Train× Low-Income	-0.006209	-0.0303	0.0161 ‡	-0.0724‡
	(0.0099)	(0.0195)	(0.0040)	(0.0257)
Subway × High-Income	0.004628	-0.0620‡	0.0355‡	-0.0633‡
	(0.0031)	(0.0059)	(0.0022)	(0.0129)
Subway× Middle-Income	0.004678	-0.0953‡	0.0197‡	-0.0596‡
	(0.0033)	(0.0083)	(0.0018)	(0.0120)
Subway × Low-Income	-0.009174†	-0.0818‡	-0.000154	-0.0417‡
	(0.0043)	(0.0141)	(0.0019)	(0.0125)
Prices				
AvgPrice × High-Income	1.2159‡	-0.0929	-1.3917‡	2.1822‡
	(0.1686)	(0.1387)	(0.1552)	(0.5123)
AvgPrice × Middle-Income	-0.4729‡	-0.2054	-2.4401‡	0.8922*
	(0.1954)	(0.1556)	(0.1354)	(0.4823)
AvgPrice × Low-Income	-0.8661‡	-0.6162**	-3.4165‡	0.959*
	(0.2082)	(0.2578)	(0.1293)	(0.5184)
Regional dummies				
Paris Dummy	-0.4969‡		-1.0269‡	
	(0.0585)		(0.0377)	
Outer Ring Dummy	-0.0972‡	-0.0391	-0.4347‡	0.3993‡
	(0.0370)	(0.0305)	(0.0214)	(0.0739)
Planned City Dummy	0.3938‡	-0.0374	0.0619‡	0.0351
	(0.0436)	(0.0356)	(0.0234)	(0.0780)
Local Tax Rates				
Residence Tax × High-Income	0.0388‡	0.0221‡	-0.0295‡	-0.002156
	(0.0060)	(0.0040)	(0.0038)	(0.0113)
Residence Tax × Middle-Income	0.0538‡	-0.001725	-0.003713	-0.003324
	(0.0056)	(0.0045)	(0.0026)	(0.0101)
Residence Tax × Low-Income	0.0587‡	-0.006367	0.0148‡	0.000731
	(0.0072)	(0.0081)	(0.0026)	(0.0122)
Ownership Tax × High-Income	-0.0382‡	-0.003597†		
	(0.0029)	(0.0016)		
Ownership Tax× Middle-Income	-0.0180‡	0.0119‡		
	(0.0024)	(0.0017)		
Ownership Tax × Low-Income	-0.0181‡	0.0151‡		
	(0.0032)	(0.0032)		
Observations	17,047	16,121	51,104	3,095
ρ^2	0.0598	0.2166	0.0553	0.1639

‡ Significant at the 1% level, † Significant at the 5% level, * Significant at the 10% level. Standard errors in parenthesis.

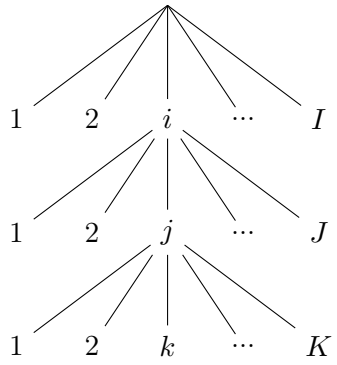
Table 7: Residential Location Choice Model, II

	Owner		Tenant	
	Apartment	Single Dwelling	Apartment	Single Dwelling
Land Use and Local Amenities				
Density	0.0140‡	-0.0750‡	0.0171‡	-0.0688‡
	(0.0018)	(0.0048)	(0.0012)	(0.0085)
%Noise (Surface)	0.1883	-0.3433‡	-0.0697‡	-0.567†
	(0.1281)	(0.1080)	(0.0719)	(0.2577)
%Water (Surface under)	0.1042	-2.4388‡	0.9928‡	-0.6122
	(0.3065)	(0.3524)	(0.1760)	(0.7749)
%Water × Children Dummy	0.3851	1.1099	0.0808	2.5462†
	(0.8023)	(0.7161)	(0.3673)	(1.2991)
% Priority Schools (Surface)	-0.0604	-0.1178‡	-0.0780‡	0.1463
	(0.0458)	(0.0435)	(0.0239)	(0.0995)
% Priority Schools × Children Dummy	0.2551‡	-0.1348*	0.4531‡	-0.0265
	(0.0856)	(0.0793)	(0.0375)	(0.1559)
%Educational Buildings (Surface)	0.5175	-11.1408‡	0.6231†	-5.9683‡
	(0.5074)	(1.1641)	(0.3163)	(2.1469)
%Education × Children Dummy	-0.0755	-0.3048	3.3991‡	-8.0326†
	(1.5100)	(1.6933)	(0.6295)	(3.2130)
Neighborhood Composition				
%Foreign HHs	8.2979‡	8.5187‡	5.2350‡	6.499‡
	(0.5791)	(0.5381)	(0.2352)	(0.9758)
%Foreign HHs × Below Secondary	3.8826‡	0.5979	3.0045‡	1.7406
	(0.6095)	(0.5024)	(0.2824)	(1.1303)
%Foreign HHs × Undergraduate	0.2241	-0.4024	0.1871	-1.4373
	(0.4228)	(0.4236)	(0.2254)	(0.9273)
%Foreign HHs × Graduate	-1.2090‡	-1.3881‡	-1.7355‡	-2.066†
	(0.4242)	(0.4690)	(0.2397)	(1.0649)
% High-Income HHs × High-Income	1.3133‡	2.6800‡	-0.0670	1.3933‡
	(0.2497)	(0.1951)	(0.1718)	(0.4723)
% Low-Income HHs × Low-Income	-0.9021†	-0.5987	0.3802†	0.8118
	(0.4337)	(0.5595)	(0.1920)	(0.8369)
% Middle-Income HHs × Middle-Income	-1.5340†	2.8920‡	1.6014‡	0.7312
	(0.6313)	(0.4131)	(0.3068)	(0.8124)
% of 1person HHs × 1 person HH	4.1973‡	-1.1671‡	4.2715‡	0.9202
	(0.1804)	(0.3930)	(0.1063)	(0.5751)
% of 2 persons HHs × 2 persons HH	-1.3139	2.1214‡	-0.1747	-1.2891
	(0.8264)	(0.6843)	(0.4722)	(1.5166)
% of 3+ persons HHs × 3+ persons HH	0.1882	3.3951‡	1.2337‡	2.9105‡
	(0.2002)	(0.2091)	(0.1093)	(0.4280)
% Young HHs × Young HH	2.4149‡	-3.7586‡	4.2568‡	-1.0913
	(0.5182)	(0.8446)	(0.2343)	(0.9658)
% Middle-age HHs × Middle-age HH	-0.6212‡	1.7560‡	-0.2820*	-0.2946
	(0.2408)	(0.2383)	(0.1492)	(0.4910)
% Old HHs × Old HH	3.6486‡	2.3031‡	1.5554‡	1.5386*
	(0.3758)	(0.3067)	(0.2951)	(0.9104)
Observations	17,047	16,121	51,104	3,095
ρ^2	0.0598	0.2166	0.0553	0.1639

Note: HH= Household Head

‡ Significant at the 1% level, † Significant at the 5% level, * Significant at the 10% level. Standard errors in parenthesis.

Figure 1: Three-level Nested Structure of Residential Location, Workplace, and Job Type Choice



3. Upper Level:
Residential Location

2. Middle Level:
Workplace

1. Lower Level:
Job Type

Figure 2: Greater Paris Area (1,300 Communes)

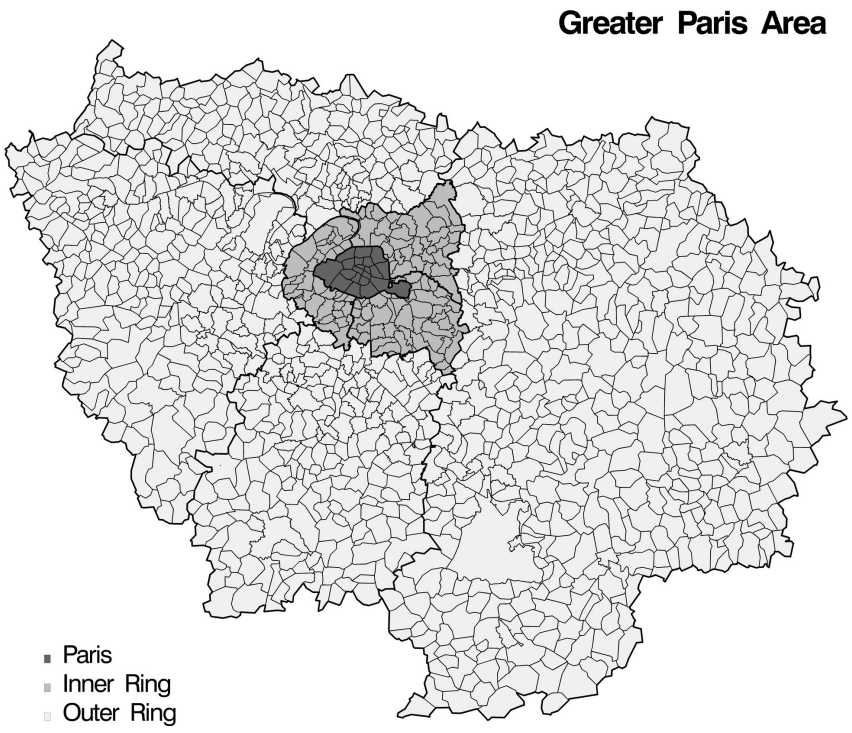


Figure 3: Aggregation of Small Adjacent Communes by Number of Jobs (950 Pseudo-Communes with More than 100 Jobs)

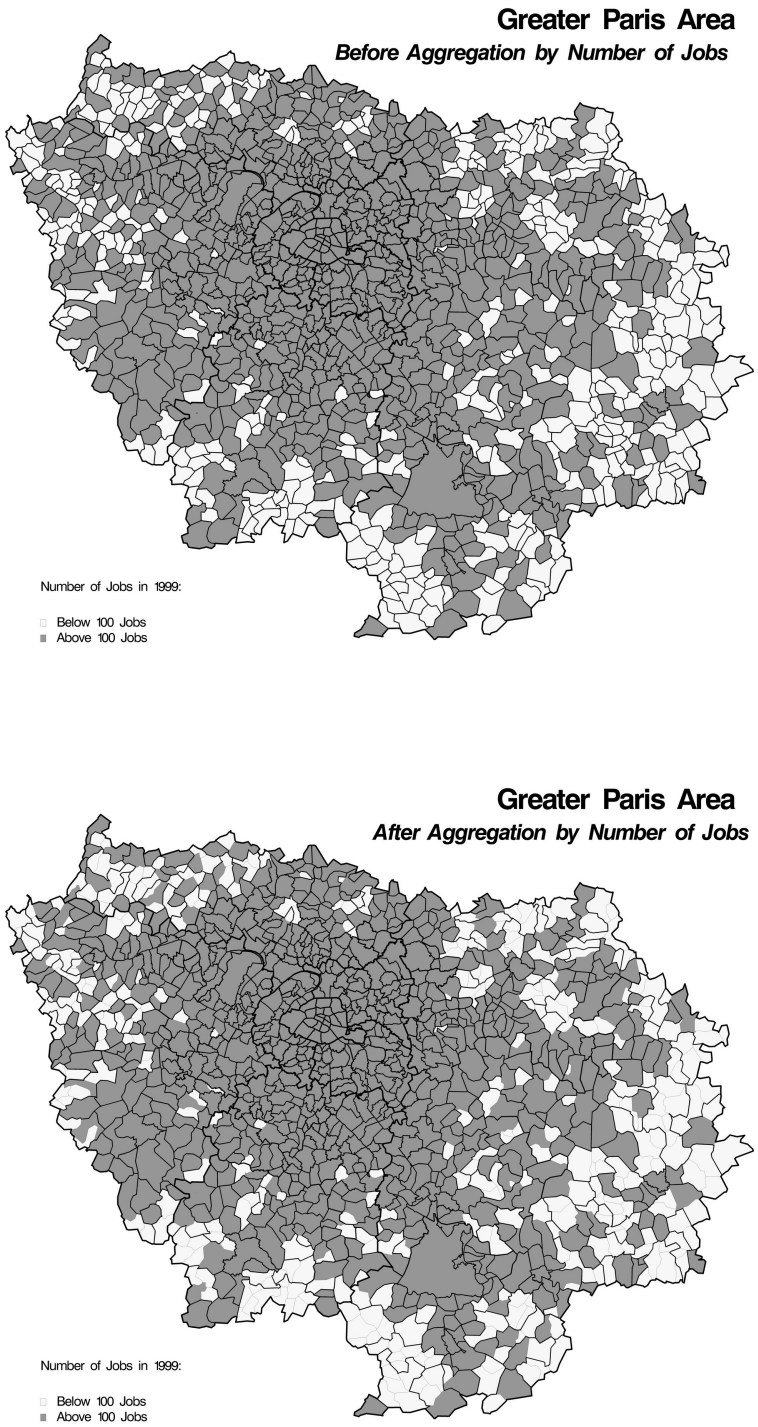


Figure 4: Three-level Nested Structure of Residential Location, Workplace, and Job Type Choice; Segmentation by Tenure and Dwelling Type

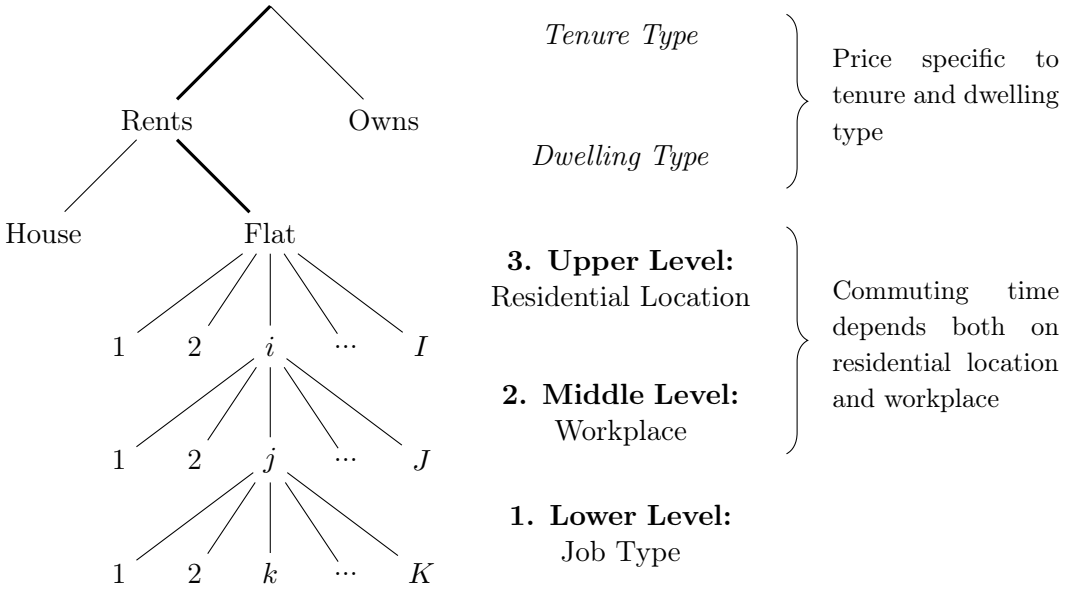


Figure 5: Attractiveness of Communes for Workers by Gender

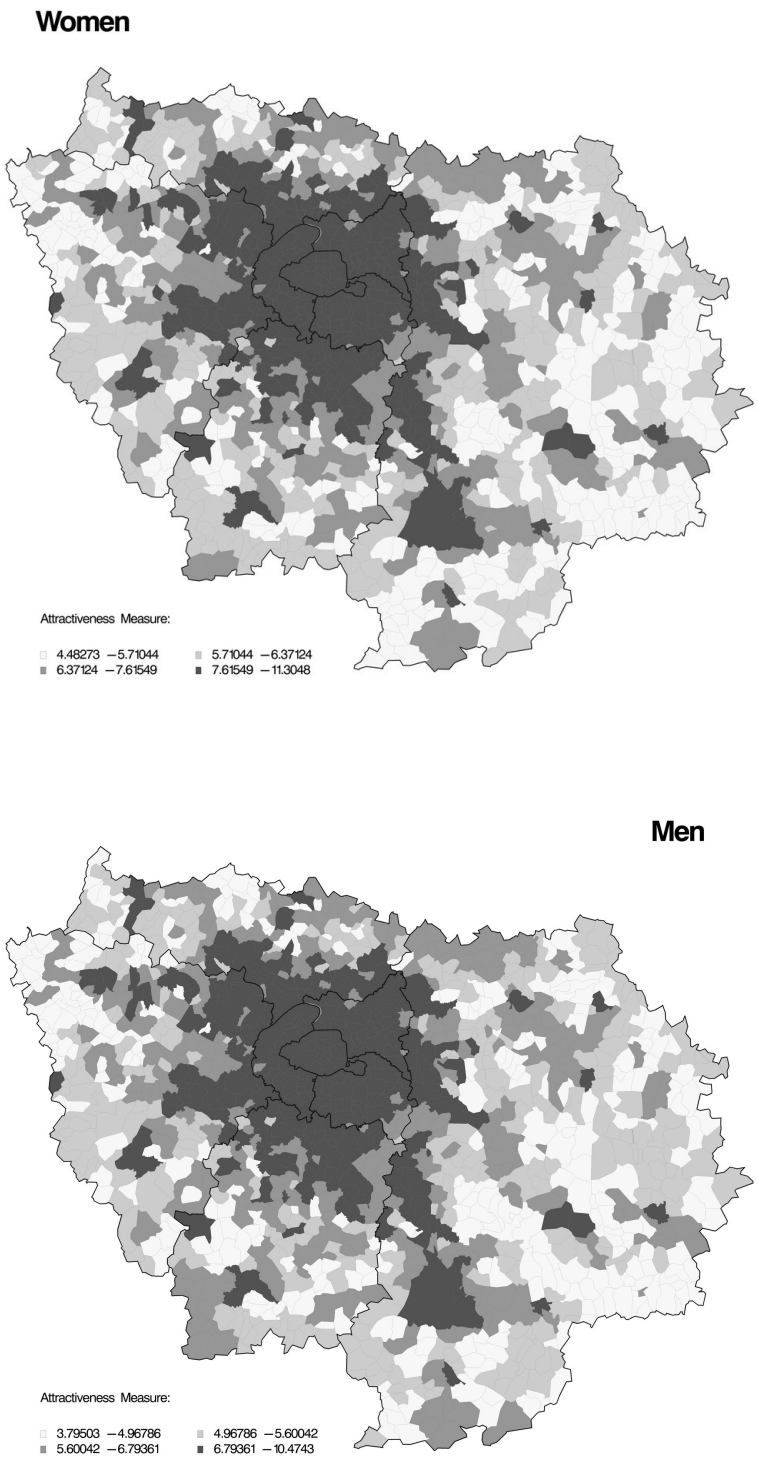


Figure 6: Attractiveness of Communes for Workers by Education Level

Figure 7: Accessibility to Jobs by Gender

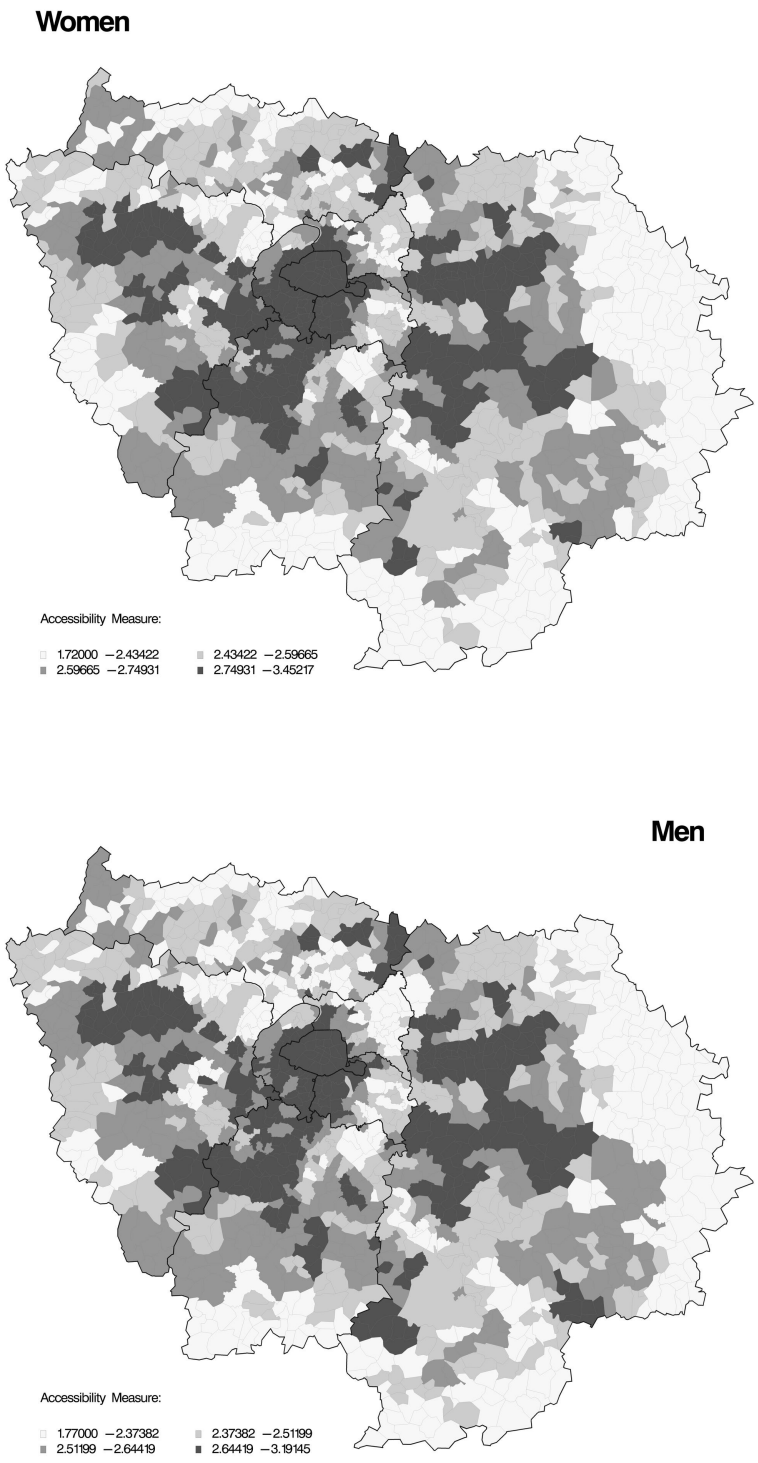


Figure 8: Accessibility to Jobs by Education Level