UMR 1041 INRA – AGROSUP

CESAER

Centre d'Economie et Sociologie
appliquées à l'Agriculture et aux Espaces Ruraux

Is sprawling residential behavior influenced by climate?

Cyrus Grout

Jean Cavailhès

Cécile Détang-Dessendre

Alban Thomas

Working Paper

2013/2

UMR 1041 CESAER INRA – AGROSUP
26, Bd Dr Petitjean - BP 87999 - 21079 Dijon cedex

Is sprawling residential behavior influenced by climate?

Cyrus Grout[1]
Jean Cavailhès[2]
Cécile Détang-Dessendre[2]
Alban Thomas[3]

Abstract

This paper addresses the question of a causal link between global warming and urban sprawl by focusing on the role local climate plays in determining household behavior regarding housing decisions. We introduce a theoretical model with a climatic amenity along urban economics lines, and consider the hypothesis that under a warmer climate, households will locate in larger plots, farther away from city centers. This hypothesis is tested empirically on household data, and by controlling for selection in simultaneous equations for housing size and distance to community center. We find evidence that housing decisions on plot size and distance to the city are related to climate differences. Global warming and urban sprawl strengthen each other in a vicious circle.

JEL *Classification*: C3, C4, Q2, R1
Keywords: urban economics, urban sprawl, global warming, sample selection, simultaneous equations

---

[1] INRA, UMR 1401, CESAER, 26, Bd. Docteur Petitjean, BP87999, 21079 Dijon Cedex, France. Corresponding author: cyrusgrout@gmail.com; 3011 NW 94th Street, Seattle, WA, 98115, USA; (1) 206.819.0878

[2] INRA, UMR 1401, CESAER, 26, Bd. Docteur Petitjean, BP87999, 21079 Dijon Cedex, France.

[3] Toulouse School of Economics, INRA and IDEI, Université des Sciences Sociales, 21 Allée de Brienne, F-31000 Toulouse

1. Introduction

Over the past decades, global warming and urban sprawl have emerged as two major environmental concerns. It is well-known that the latter contributes to the former. Here, we explore the opposite causal link: does warming influence urban sprawl?

The existence of anthropogenic climate change and its primacy as an environmental concern is well established. The Intergovernmental Panel on Climate Change (IPPC) finds in its 2007 assessment (IPCC, 2007) that "*Observational evidence from all continents and most oceans shows that many natural systems are being affected by regional climate changes, particularly temperature increases.*" Unmitigated climate change is projected to have widespread impacts affecting among other things, accessibility to freshwater resources, biodiversity and ecosystem integrity, agricultural and forest productivity, ocean flooding, and societal stability.

The primary environmental concerns associated with urban sprawl are excessive consumption of land resources and increased greenhouse gas emissions due to long commuting that accompany diffuse and uncoordinated patterns of urban expansion. Economic analyses of the causes of urban sprawl have identified rising incomes and auto-driven changes in transportation costs as primary drivers of sprawling development (Nechyba and Walsh, 2004; Glaeser and Kahn, 2004), as well as the attraction of rural amenities (Wu, 2006) and the movement of high income groups to the suburbs in response to fiscal and social problems in city centers.

Environmental concerns associated with urban sprawl include its indirect impact on climate change. A number of studies have analyzed the links between sprawl-induced increases in energy use and consumption of green space, and increases in the concentration of green-house-gases in the atmosphere (e.g., Bart, 2010). The famous figure by Newman and Kenworthy shows an intricate relationship between urban density and fuel consumption: slack North American cities are more energy intensive than dense Asian cities (Newman and Kenworthy, 1999). But what about the reverse link, *i.e.* the effect of climate on the extent of urban sprawl? To our knowledge, only a few papers account for the potential effect of climate on sprawl, and only incidentally. For example, Burchfield et al. (2006) and Pattachini and Zenou (2009) find that climate is a significant determinant of the extent of urban sprawl. However, the contribution of climate to urban sprawl is not a focus of these studies; rather, they use regional climate as a control variable in cross-sectional analyses of urban area-level measures of sprawl. How urban growth patterns may adapt to climate change impacts is generally left unstudied (Irwin et al. 2009).

In this paper, we focus on the role local climate characteristics play in determining some aspects of urban sprawl. First, we introduce a theoretical urban micro-economic model in the tradition of Alonso (1964) and Muth (1969), and analytically synthesized by Fujita (1989), that introduces a climatic amenity. We hypothesize that under a warmer climate households will consume more of a residential good and choose locations with less proximity to the urban center. Then, we empirically test if, in a temperate country (France), households' preferences for larger plots more distant to city centers, are stronger in milder climates. To do so, we test for the role of climate by considering both temperature and the number of days of precipitation, using detailed household-level survey data to account for differences in individual household characteristics. The model predicts residential location choices represented by distance to city center and residential lot size, as a function of household and dwelling characteristics, and local climate. The latter is not only expected to influence distance to city center and size, but also preferences for outdoor space.

As pointed out by Smith (1988), every household is expected by belong to a submarket that is differentiated by location, dwelling type, tenure form, age, quality, and financing. Because of such differentiation in the housing market, it is important to select a homogenous market when modeling housing decisions. For this reason, the population we consider when testing the above hypothesis is composed of households living in a single detached house outside central cities of urbanized areas. However, restricting the analysis of the impact of climate on housing to a particular type of dwellings may introduce a sample selection bias if the distribution of households within the submarket is not random. We propose an estimation strategy to simultaneously predict the distance to the city center and the residential plot's size by a simultaneous-equation procedure suggested by Gouriéroux (2000). We also control for selection resulting from plot size being observed only when the household chooses a single-detached house and distance to the urban center being censored at 0 for households living anywhere within the central municipality.

2. What is Urban Sprawl?

Although "urban sprawl" is a well-recognized term, it lacks a definition that is precise and generally accepted. Brueckner (2001) defines urban sprawl as the "spatial growth of cities that is excessive to what is socially desirable." This definition reflects the negative connotations that typically accompany urban sprawl, which is often associated with environmental and economic costs resulting from the irreversible consumption of valuable resource lands, the inefficient use of infrastructure, increased energy consumption, impeding efforts to mitigate climate change, and social segregation. In its report *Urban Sprawl in Europe: the Ignored Challenge* (European Commission, 2006), the European Environment Agency remarks, "Sprawl threatens the very culture of Europe, as it creates environmental, social and economic impacts for both the city and countryside of Europe."[4]

Other authors adopt definitions of sprawl that are more circumspect about its desirability. Glaeser and Kahn (2004) discuss urban sprawl in terms of decentralization and density: that is, the spreading of employment and population throughout an urban area and the degree to which it is concentrated. They, and others such as Nechyba and Walsh (2004), find that consequences of urban sprawl have been largely positive due to the higher household consumption of land and housing it has enabled. They identify the primary problem with sprawl as the inequality resulting

---

[4] "The European Environment Agency (EEA) has described sprawl as the physical pattern of low-density expansion of large urban areas, under market conditions, mainly into the surrounding agricultural areas. Sprawl is the leading edge of urban growth and implies little planning control of land subdivision. Development is patchy, scattered and strung out, with a tendency for discontinuity. It leap-frogs over areas, leaving agricultural enclaves." (European Commission, 2006)

from sprawl-related fiscal reorganization that has occurred in many urban areas, particularly the migration of wealthier households away from urban centers.

2.1 Measuring urban sprawl

Increasing concern about the causes and consequences of urban sprawl, and the contentious debate around what to do (or not do) about it, have led to calls for precise definitions and robust empirical measures of the phenomenon by geographers (e.g., Ewing, et al., 2002; Galster, et al., 2001; Torrens and Alberti, 2001), economists (e.g., Irwin, et al., 2006; Song and Knapp, 2004), and researchers of both disciplines (Burchfield et al., 2006). Of course, the multi-faceted nature of urban sprawl does not lend itself to any particular quantitative measure. Indeed, empirical analyses of urban sprawl tend to (necessarily) define it in relation to available data, the theoretical perspectives of the researcher, and particular social or environmental concerns (of which there are many to choose from).

The past decade has produced a number of studies in the geography literature focused on the establishment of quantitative measurements of urban sprawl. Galster et al. (2001) argue that policy attempts to deal with urban sprawl have been hampered by the lack of a consistent metric that enables cross-sectional and dynamic comparisons. These authors note that sprawl has been described in terms of land-use *patterns*, the *process* of urban expansion, the *causes* of particular patterns, and the *consequences* of particular patterns.

While the recent geography literature has, for the most part, adopted the conceptualization of sprawl as a *pattern* of land use, the choices of what to measure and how to measure it have been varied. Metrics used in geographic analyses of sprawl include:

- Density (residential and/or employment) and the distribution of density
- Continuity or conversely, fragmentation (i.e., the spatial organization of density)
- Centrality (of population to central business districts and sub-centers)
- Proximity/accessibility (to shopping areas, transportation, jobs)

An important aspect of sprawl that emerges from geographers' efforts at quantification is the complexity and multidimensionality of sprawl. By defining sprawl along a number of dimensions, geographers have in essence created a typology of sprawl. Empirical investigations of the causes and consequences of sprawl are unlikely to address all dimensions of sprawl simultaneously. While the measurement of particular patterns of land use is not generally the primary research concern, it does provide a useful hub between the causes and consequences urban sprawl. A difficulty is that cross-sectional analyses of sprawl that develop city-level indices of sprawl have proven to be sensitive to the way sprawl is quantified (e.g., density vs. spatial composition) and how the chosen unit of analysis is defined and/or measured (see, for example, Wolman et al., 2005).

Economic analyses of urban sprawl have tended to focus on one or two of the dimensions described above. Standard models of urban structure and the process of suburbanization, which focus on transportation costs, income distribution, and agglomeration effects as drivers of urban structure, tend to define sprawl along the dimensions of density and centrality. In general, these studies measure the degree of sprawl in terms of the distribution of jobs and/or residential population relative to a metropolitan city center as defined by density gradients (e.g., Anas, et al., 1998, Glaeser, et al., 2001, Glaeser and Kahn, 2004, Mieszkowski and Mills, 1993). Aside from measuring proximity to the central business district, and in some cases to sub-centers of employment, these studies do not tend to account for the spatial organization of development.

A few more recent studies have defined urban sprawl as the degree of scatteredness (fragmentation) or urban development. Burchfield et al. (2006) measure the scatteredness of new urban development using aerial photographs and high resolution land-cover data for the entire United States. On a parcel level, the degree of sprawl is measured as the proportion of undeveloped land in a square kilometer surrounding a newly developed parcel of land.[5] Recent studies of urban sprawl in Europe have focused on the variation over time of cities' total urban areas and population levels. Kasanko et al. (2006) develop a series of indicators of urban sprawl applied to 15 urban areas in Europe: built-up areas, the continuity of residential areas, land taken by urban expansion, and population density. In an analysis of European cities using the Urban Audit data-set, Patacchini and Zenou (2009) adopt what they acknowledge to be a very narrow and limited view of urban sprawl, and measure its degree as the variation over time of urban land use and population density.

As discussed above, urban sprawl is typically defined as an objective fact that can be observed and objectively measured. Sprawl is clearly an outcome of residential behavior, but the 'sprawling behavior' of households choosing their residential locations remains a black box. At best, statistical correlations are established between the characteristics of urban area populations and variables measuring geographical settlement patterns.

In the present paper, we analyze urban sprawl from a microeconomic perspective and investigate household behavior at theoretical and empirical levels. Our variables of interest are residential lot size and commuting distance, which are characteristics linked to urban sprawl. We obtain qualitative predictions from the market-clearing conditions of an urban economic model, in terms of lot size and commuting distance as the primary household decisions. We then model these variables in a random utility framework to describe household behavior (as a reduced form of the theoretical model). To our knowledge, this is the first paper to use household level variables to represent urban sprawl: economists have tended to use urban area-level statistics as dependent variables in previous empirical analyses.

---

[5] This measure was criticized by several authors. Among the economists, Irwin and Bockstael (2007) argue that Burchfield et al.'s measure of sprawl is systematically biased against recording low-density residential development, primarily due to the resolution of their data. As a primary measure of fragmentation, they use the proportion of contrasting edges (the length of the edge boundary between two contrasting land uses), which accounts for both non-contiguous development (increasing sprawl) and in-fill development (decreasing sprawl).

2.2 Determinants of urban sprawl

2.2.1 Urban sprawl in the literature.

A number of papers advance a synthetic analysis of the causes of sprawl, see Glaeser and Kahn (2004), Brueckner (2000) and Irwin et al (2009) for literature reviews. According to (among others) Glaeser and Kahn (2004), the automobile is the primary driver of urban sprawl. They describe urban sprawl as an “inexorable product of car-based living”. Brueckner (2000) underlines the role of improved transportation infrastructure in lowering commuting costs, as well as increasing incomes, in driving urban growth. These authors also discuss the role inner-city social problems play in leading many households to move to more “socially controlled suburbs”. Wu (2006) widens the milieu of causes to households seeking amenities (in particular, green amenities).

In an analysis of urban sprawl in Europe, the European Environment Agency synthesizes the determinants of sprawl, connecting it with new lifestyles outside the inner city (European Commission, 2006). This report also mentions the role of improved transportation links and personal mobility, in allowing individuals to live in more and more distant locations. Several determinants of urban sprawl are identified at various scales: for example, micro and macro socioeconomic trends, land prices, individual preferences, demographic trends, and land use planning policies. Moreover, according to the report, there is strong evidence that urban sprawl is bound to occur in a “mechanistic way” where unplanned and decentralized development dominates.

2.2.2 Human behavior: is climate a determinant of urban sprawl?

Climate is a well-known driver of human behavior. Its effects on population migrations have a long history (Graves, 1976; 1980; Graves and Linneman, 1979). Cheshire and Magrini (2006) have shown its impact on the growth of urban populations in Europe. However, the relationship between climate and urban sprawl is not well studied. As previously stated, Burchfield et al (2006) and Pattachini and Zenou (2009) show that urban sprawl is greater where the weather is warmer, but do not investigate climate as a driver of sprawl.

We go further by providing behavioral explanations for this phenomenon and will show here, from a theoretical model and an empirical model, that climate is a determinant of urban sprawl. More precisely, our assumptions are that where the climate is milder:

(i) People locate farther into the countryside for two reasons; first, commuting is easier than in a cold region where winter driving is difficult due to, for example, fog, glare, and snow. Second, in regions where the summer is hot, location outside of city centers is a means of avoiding the ‘heat island effects’ of cities.

(ii) People live on larger plots, because they spend more time using a larger residential lot (“pursuing an outdoor lifestyle”): for example, it may provide space for a swimming pool, a barbecue, a garden, or an outdoor game area.

An outcome of ease of transportation, i.e. point (i), is accessibility, and an outcome of lot size, i.e. point (ii), is density of population. As we will show, accessibility and density are linked to residential land-use patterns (dense or scattered), such that our two points of interest cover a wide range of measures of urban sprawl identified by geographers (as described in this section). To the best of our knowledge, these points are new in environmental and urban economics literatures, without any related theoretical model or empirical estimation from individual data.

3. Theoretical Model

3.1 The framework

We now present an urban economic model in the tradition of Alonso (1964) and Muth (1969), analytically synthesized by Fujita (1989). As in Fujita (1989, pp. 31-38), this model is 'time-extended.' Its equilibrium depends on the maximization of a utility function that includes a leisure attribute under two constraints: a budget constraint and a time constraint.

We introduce into the utility function a climatic amenity that may produce three effects:

1. The climatic amenity may be directly consumed by households, in addition to the consumption of a residential good, and an aspatial composite good consisting of all other goods and leisure.
2. The consumption of the residential good may depend on climate. As mentioned in Section 2, where climate is milder, the household may substitute more land for less inside living area and/or composite good. In other words, the parameter of the residential good in the utility function depends on climate.
3. The residential location chosen by the household regarding the city center may also depend on climate. In a cold region, all things being equal, commuting is slower and more hazardous (i.e., more costly) than in a warmer region (see Section 2). In other words, the time constraint depends on climate.

The properties of the urban equilibrium can be deduced from the usual maximization of the utility with respect to the three above effects under the constraints of time and budget.

Urban economics consists of two types of models: the open city model and the closed city model. As we are interested in sprawling behavior, and people who choose to sprawl tend to originate from within the same urban area, the closed city model is more relevant to the empirical part of our paper. For example, 80% of the migration within France in 2004 concerns mobility within the same 'department' (Baccaïni, 2005; a department is a French administrative region covering between 4,000 and 7,000 square kilometers). Therefore, we use the closed city model as the basis of our theoretical analysis.

3.2 The basic model with a climatic amenity

Space is represented by the half real line $X = [0, +\infty]$ where the CBD is the point at the origin of the line. It is assumed that all employment is concentrated in the CBD where identical households work. Identical households reside at the spatial location $x \in X, x > 0$ and consume the following goods: a residential good consisting both of a building, B, and a plot of land $S(x)$ for which land rent $R(x)$ is paid, an aspatial composite good $Z(x)$ taken as the *numéraire* ($P_z = 1$), a quantity of leisure time θ_l that is the total available time $\bar{\theta}$ minus the commuting time

θ_c (where θ_c is the unit commuting time) and the working time θ_w, and a climatic amenity A that is constant for all x. We adopt a Cobb-Douglas-type utility function:

$$U = Z(x)^{\alpha(A)} S(x)^{\beta(A)} B^{\xi} \theta_l^{\gamma}(x) A^{\delta}. \quad \textbf{(1)}$$

The exponents α and β depend on the amenity A, but $\gamma$and ξ do not for two reasons. First, a unit of leisure time provides the same utility regardless of climate because it can be spent on indoor or outdoor leisure. Second, in contrast to the plot size, which includes outdoor attributes (yard, etc.), the utility brought by a house or an apartment is independent of climate because it depends on indoor attributes (e.g., the number of rooms and bathrooms). Without loss of generality, we assume that$\alpha(A) + \beta(A) + \gamma + \xi = 1$.

The household faces a budget and a time constraint:

$$Z + p_B B + RS + tx = w\theta_w + Y \qquad \theta_l + \theta_w + \theta_c x = \bar{\theta}, \quad \textbf{(2)}$$

where P_B is the price of the structure[6], t is the unit commuting cost, θ_w is the working time per a chosen unit of time (day, month, year), w is the unitary wage per unit of time, and Y is an amount of money independent of working time. We assume, in the general case, that the worker can freely choose her leisure time and working time.

From the time constraint we obtain $\theta_w = \bar{\theta} - \theta_l - \theta_c x$; substituting θ_w into the budget constraint, the model becomes

$$\max U(Z, S, B, \theta_l, A),$$

subject to :

$$Z + p_B B + RS + w\theta_l = w(\bar{\theta} - \theta_c x) + Y - tx. \quad \textbf{(3)}$$

As noted by Fujita (1989), this formulation suggests that the household makes a fictitious transaction: it sells all its available time net of commuting $(\bar{\theta} - \theta_c x)$ to the employer at the wage rate w (r.h.s. of the equation) and then purchases back its leisure time θ_l at the same unit price of time w (l.h.s. of the equation). The wage rate w also serves as the unit price of leisure time. The total commuting cost at distance x is $\Theta(x) = (t + w\theta_c)x$. Following Fujita (1989), the bid rent function is

$$\Psi = \max_{S,\theta_l}\left[\frac{w\bar{\theta} + Y - (w\theta_c + t)x - Z - p_B B - w\theta_l}{S}\right],$$

with the first-order conditions obtained as the usual marginal conditions:

$$-\frac{\partial Z}{\partial S} = \Psi \qquad -\frac{\partial Z}{\partial \theta_l} = w.$$

---

[6] The price of the structure is footloose, because the price of materials (cinderblocks, wood, etc.) and labor (mason, carpenter, etc.) do not depend on location.

To simplify notation, we let $w\bar{\theta} + Y - P_B B = I$ and assume that Y is large enough to satisfy $I > 1$. From the first-order conditions, we obtain:

$$R(x) = \alpha^{\frac{\alpha}{\beta}} \beta \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{\beta}} [I - (w\theta_c + t)x]^{\frac{1}{\beta}} A^{\frac{\delta}{\beta}} U^{-\frac{1}{\beta}}, \tag{4}$$

$$Z(x) = \alpha[I - (w\theta_c + t)x],$$

$$\theta_l(x) = \frac{\gamma}{w}[I - (w\theta_c + t)x], \tag{5}$$

$$S(x) = \alpha^{-\frac{\alpha}{\beta}} \left(\frac{\gamma}{w}\right)^{-\frac{\gamma}{\beta}} [I - (w\theta_c + t)x]^{\frac{\beta-1}{\beta}} A^{-\frac{\delta}{\beta}} U^{\frac{1}{\beta}}. \tag{6}$$

As mentioned above, we suppose that $\beta(A)$ and $\theta_c(A)$ depend on A. We focus on the effect of climate on lot size and leisure time by studying $\partial S/\partial A$ and $\partial \theta_l/\partial A$, which correspond to the comparative statics in the empirical analysis. It is easy to verify that $\partial S/\partial x$ is positive.[7]

3.3 The equilibrium

In the closed city model, we determine both the limit of the city, x_F, and the utility of the citizens by writing that the city population, N, must be housed. At the border of the city, the residential land rent equals the agricultural opportunity rent. Therefore

$$x_F = \frac{I - \alpha^{-\alpha} \beta^{-\beta} \left(\frac{\gamma}{w}\right)^{-\gamma} R_A^{\beta} A^{-\delta} U}{w\theta_c + t} \tag{7}$$

The population N is housed if

$$\frac{N}{2} = \int_0^{x_F} \frac{1}{S} dx.$$

We deduce from (6) that

$$U = \alpha^{\alpha} \beta^{\beta} \left(\frac{\gamma}{w}\right)^{\gamma} A^{\delta} \left[\frac{N}{2}(w\theta_c + t) + R_A\right]^{-\beta}. \tag{8}$$

3.3.1 The plot size

The demand function of the plot size is

$$S = \alpha^{-\frac{\alpha}{\beta}} \left(\frac{\gamma}{w}\right)^{-\frac{\gamma}{\beta}} [I - (w\theta_c + t)x]^{\frac{\beta-1}{\beta}} A^{-\frac{\delta}{\beta}} \left\{\alpha^{\alpha} \beta^{\beta} \left(\frac{\gamma}{w}\right)^{\gamma} A^{\delta} \left[\frac{N}{2}(w\theta_c + t) + R_A\right]^{-\beta}\right\}^{\frac{1}{\beta}},$$

---

[7] In the empirical analysis, these derivatives cannot be interpreted as the effects of a climatic change over time (for example, global warming). Rather, they must be interpreted as the effects of spatial differences between regions that experience different climates.

$$S=\beta\left[I-(w\theta_c+t)x\right]^{\frac{\beta-1}{\beta}}\left[\frac{N}{2}(w\theta_c+t)+R_A\right]^{-1}. \tag{9}$$

This function does not depend on the direct consumption of the climatic amenity because we assume that climate is the same at all locations in the city. However, the taste for an outdoor way-of-life or the commuting cost may depend on A and therefore we study $\partial S/\partial A$.

$$\frac{\partial S}{\partial A}=\left[\frac{N}{2}(w\theta_c+t)+R_A\right]^{-1}\left[I-(w\theta_c+t)x\right]^{\frac{\beta-1}{\beta}}$$

$$\left\{\underset{>0}{\frac{\partial\beta}{\partial A}}+\underset{>0}{\frac{1}{\beta}\frac{\partial\beta}{\partial A}\ln\left[I-(w\theta_c+t)x\right]}-\underset{<0}{\frac{(\beta-1)wx}{I-(w\theta_c+t)x}\frac{\partial\theta_c}{\partial A}}-\underset{>0}{\frac{\beta wN}{(w\theta_c+t)N+2R_A}\frac{\partial\theta_c}{\partial A}}\right\}. \tag{10}$$

The term in the top line of (10) is positive. We have assumed that $I-(w\theta_c+t)x>1$, so that its logarithm is positive. The two terms in the first line, and the last term in the bottom line of (10) are positive, and the third is negative: therefore, the effect of an increase in the climatic amenity is indeterminate. The effect of the third term on $\partial S/\partial A$ is greater in magnitude for larger values of x. The greater is x, the more negative is $\partial S/\partial A$. At $x=0$, the effect of A on S is positive if $\partial\beta/\partial A>0$; it is everywhere positive if $\partial\theta_c/\partial A=0$ and $\partial\beta/\partial A>0$.

In the general case if $(\partial\theta_c/\partial A<0, \partial\beta/\partial A>0)$ the lot size increases with the climatic amenity close to $x=0$ if $\partial S/\partial A>0_{|x\ small})$; for larger x the lot size may decrease with the climatic amenity $(\partial S/\partial A<0_{|x\ large})$. Indeed, if $\partial\theta_c/\partial A<0$ more people locate in the periphery because commuting is cheap when A is large, leading here to more expensive land rent and a reduction of the peripheral lot size compared to the same location in a city where A is lower.

In warmer cities, the peripheral lot size is larger or smaller than in cold cities depending on the value of the parameters: whether the magnitude of a relative increase dominates in the periphery is an empirical question. The smaller the absolute value of $\partial\theta_c/\partial A$ compared to $\partial\beta/\partial A$, or the larger is β, the more a particular peripheral lot size in a warm city trends upwards compared to a cold city. In other words, when the consumer is more sensitive to climate's effect on lot size than she is to its effect on commuting cost, an increase in the climatic amenity entails an increase in the lot size everywhere; otherwise, a decrease in the lot size occurs in the periphery. This latter case is represented by "Warmer city 2" in Figure 1 below.

Proposition 1. *Close to the city center lot size increases with the climatic amenity. The magnitude of this increase diminishes with distance to the CBD, leading, beyond a certain distance, to a possible decrease of the lot size when the consumer is more sensitive to the effect of climate on commuting cost than to the effect of climate on lot size.*

3.3.2 Leisure, and the optimal distance x^*

The demand for leisure time is given by (5). If the unit commuting cost depends on the amenity $(\partial\theta_c/\partial A<0)$, the derivative of the leisure time is:

$$\frac{\partial \theta_l}{\partial A} = -x\gamma \frac{\partial \theta_c}{\partial A} > 0. \tag{11}$$

Proposition 2. *At a given distance, leisure time increases with the climatic amenity.*

For a given leisure time θ_l^* the optimal distance x^* is

$$x^* = \frac{I - \frac{w}{\gamma}\theta_l^*}{w\theta_c + t} \tag{12}$$

with

$$\frac{\partial x^*}{\partial A} = -\frac{w}{\left(w\theta_c + t\right)^2}\left(I - \frac{w}{\gamma}\theta_l^*\right)\frac{\partial \theta_c}{\partial A} > 0\,.$$

Proposition 3. *For a given leisure time, the optimal distance increases with the climatic amenity.*

The primary outcomes of interest of the microeconomic program can be summarized by Figure 1 where for simplification, the curves of the plot size functions are stylized as straight lines.

[FIGURE 1 HERE]

4. The empirical model

The theoretical model of the previous section presents insights into the household tradeoffs between proximity to urban centers and the consumption of residential goods, and how those tradeoffs may be influenced by the level of a climatic amenity. From an empirical perspective, the size of the residential good, the distance from the city center (or leisure), and the quantity of the composite good constitute a system of three equations, one of which can be dropped. In this case, we drop the composite good. The two other goods are simultaneously chosen by each household, leading to endogenous variables in the model to be estimated and additional variables must be introduced into the model to control for observed household heterogeneity. Therefore, we estimate a system of simultaneous equations:

$$\begin{aligned} S &= f(D, w, \alpha_1, C_1, Y_1) + \varepsilon_1, \\ D &= g(S, w, \alpha_2, C_2, Y_2) + \varepsilon_2, \end{aligned} \tag{13}$$

where S is the size of the residential plot, D is distance to the center of the urban area (see below), w is the income of the household, C are climatic measures, and Y are local variables (e.g., population, mean income of the locality, and zoning). The α variables characterize the household. An important difference between the theoretical model and the empirical model is that the former assumes each household to be identical and distributed across a featureless plain.

The introduction of covariates is to control for the fact that in the real world households and landscapes are heterogeneous.

The urban economic model's assumption that the land market is in equilibrium is also relaxed. This assumption is used by Rosen (1974) to estimate equilibrium prices $\hat{p}$ (in a first step) and a demand function of attributes, such as the residential lot size (in a second step): $S = S(\hat{p}, M)$, where M are characteristics of the household. Here, we do not estimate demand functions. We estimate a function $S = S(X, M)$, where X is a vector of attributes that influence the consumption of the residential lot size. The advantage is that we need not impose equilibrium on the residential market (through $\hat{p}$), and we avoid the assumption of the hedonic price method: uniqueness of the residential market (which is not credible at the country level). Hence, because X does not capture all the effects that determine the prices, the function $S = S(X, M)$ is less informative than a demand function. A way to compensate for this is to use a local price of developable land as a covariate, which captures a number of omitted variables. The other covariates of X are then interpreted as factors of local deviations from this price of developable land (see below).

A final distinction to make between the theoretical and empirical models is that while the theoretical model considers one urban area under various climatic conditions; empirically we observe climatic variation across urban areas. Thus, we must take into account differences among the urban areas each household is located in, particularly the variation in urban area population.

In the French statistical definitions created by INSEE, an 'aire urbaine' (urban area) consists of a 'unité urbaine' (central city and suburbs defined by the continuity of the built up area and the hosting of at least 5,000 jobs) and a periurban belt (municipalities with discontinuous built-up land where at least 40% of active residents commute to work outside the commune, but within the urban area). This characterization of urban areas is similar to that of MSAs in the United States, but the population thresholds are lower. We have associated rural communes, which are beyond the periurban area border, to their nearest urban area.[8] In 1999, the French territory contained 354 urban areas, each of which corresponds to the 'city' of our theoretical model.

In spite of the differences between the theoretical and empirical models, the characterization of the household's microeconomic program and the hypothesis to be tested are essentially the same. Both models feature the tradeoff between accessibility, i.e. proximity to an urban center (where jobs and services are concentrated) and the level of consumption of a residential good (we focus on residential plot size in the empirical model), where each household simultaneously chooses quantities of these two goods. The hypothesis in both models is that this tradeoff between accessibility and residential consumption is affected by climatic amenities.

---

[8] Determined by minimizing the distance to the closest peripheral *aire urbaine* border.

5. Data

The individual socioeconomic data are from housing surveys conducted by INSEE in 1984, 1988, 1992, 1996, 2002, and 2006, which are stacked and merged with spatial data (including periodic INSEE censuses and the delineation of urban areas) and climatic data.[9] We describe below the household, local, urban area, and climatic variables we use in our two equations of interest (with S and D as dependent variables).

Endogenous variables

Residential lot size and the distance to the city center of the urban area are the two dependent variables (in logs).

- Lot size is obtained from the INSEE Housing Surveys and is defined only for single detached houses.
- Distance is the driving time in minutes to the closest urban center via the road network. Because location is defined at the commune level (the commune is the smallest level of French administrative delimitation), this distance equals zero if a household is located in the central commune of an urban area.

Household characteristics

- **Income.** In the theoretical model, the household's budget constraint is determined in part by its income level. *Ln_income* is measured as the natural log of total household income as reported in the housing surveys, and deflated by the GDP index (year 2006).
- **Age.** Different age cohorts are likely to have different preferences for accessibility to central city. We assume that younger households are willing to substitute greater accessibility for less space. The variable *age* is defined as the age of the head of household as reported in the housing surveys.
- **Children.** Households with children are likely to demand larger residential plot sizes, but also locations closer to central cities for the purpose of access to schools. *Children* is a dummy variable indicating the presence of children as reported in the housing surveys.
- **Two commuters.** Households with two commuters should place a higher value on proximity to job centers. Their leisure time is lower than when only one individual commutes to work, thus these households may differently value residential plot size. *Two commuters* is a dummy variable equal to one in the presence of two commuters in the household, as reported in the housing surveys.
- **Socio-professional status.** Household socio-professional status may be related to preferences for certain types of lifestyles. For example, in certain European cities (e.g., Paris), executives tend to locate in downtown because they value proximity to its cultural resources (Bruckner *et al.*, 1999); conversely, blue collar workers reside in the periphery: they may be gardeners or craftsmen, and thus value larger lot sizes. *Executive*, *Blue collar*, *Office worker*, and *Other* are dummy variables

---

[9] We thank the French Committee for Statistical Privacy for allowing us these matching operations and INSEE for providing us with the computing and technical facilities required for econometric modeling on a secured data server.

indicating the occupation of the head of household, as reported in the housing surveys, *Intermédiaire* being the reference.[10]

Urban area characteristics

- **Urban area land price.** Affordability of land is not the same across the urban areas of France. To account for regional variation in price and the affordability of the residential good, and to control for potential omitted variables, we calculate an average price for each 'aire urbaine'. We use data on the median price per square meter of developable land at the 'bassin de vie' level (a region of commercial activity defined by INSEE). The variable *land price* is the mean, averaged over the 'bassins de vie' located within the '*unité urbaine'* of each 'aire urbaine'.
- **Average price index.** A housing price index is used to control for national trends in the housing market (*index_logement*). The index is matched to the survey data based on the year a household purchased its dwelling.
- **'Aire urbaine' population.** In addition to the price of developable land, the size of an urban area will tend to affect the affordability of the residential good. Larger urban areas tend to offer more services, cultural amenities, job opportunities, and higher wages. Moreover, commute times tend to be longer in larger urban areas. The variables *pop area xxxx* are dummy variables indicating that the size of the household's aire urbaine falls within a particular population range.
- **Share difference.** A household may prefer to live in a part of its urban area where its neighbors are likely to be similar to itself. We attempt to account for this potential type of sorting using socio-professional data from the censuses. *Share difference* is calculated as the difference between two terms: *share own* is the percentage of workers (as defined at the household level in the housing surveys) in a household's commune that are in the same socio-professional category; *share avg* is the percentage of workers in a household's urban area that are in the same socio-professional category; *share difference = share own-share avg.*

Local characteristics

These commune-level characteristics may affect the local price of the residential good (other spatial variables were tested but were excluded as their parameters were insignificant):

- **Forest rate.** The share of commune territory covered by forest is used as an indication of green amenities and competition on the land market.
- **Income per capita.** The mean income level of the households of a commune may be indicative of the presence of local amenities/nuisances and/or peer effects, both capitalized into the land price, and thus affecting the lot size. The variable *ln local income* is the natural log of mean per-capita income of the commune.
- **Coastal commune.** *Coastal Commune* is a dummy variable indicating that the commune is located on a coastline.

Climatic variables

Climatic data were obtained from Météo-France over the period 1971–2000. They consist of average temperature for the months of January and July, and the number of rainy days in

---

[10] There is not a good translation for this category which includes, intermediate professions in education, health, public service, and corporate administration, as well as technicians, foremen, and supervisors.

January[11]. These data are recorded on a network of weather stations, and interpolation is used to reconstruct a spatial continuum of weather data using a GIS (Joly et al., 2011). Regressions between temperature/rainfall and explanatory variables suggested by climatology (altitude, land-cover, orientation, etc.) were estimated, followed by kriging of residuals from the regressions. As the models and parameters estimated are not identical over an area of the size of France, interpolation is done for small polygons including the 30 closest stations. The predicted values are computed for each French commune. The selected variables are:

- **July temperature** The variable *ln temperature july* is the natural log of the average temperature in the household's commune during the month of July (degrees Celcius).
- **January temperature** The variable *ln temperature january* is the natural log of the average temperature in the household's commune during the month of January (degrees Celcius).
- **January days of precipitation** The variable *ln rainydays january* is the natural log of the average number of days of rain in the household's commune during the month of January.

6. Model specification and estimation methods

6.1 The population of interest

We assume each household has freely chosen its location and the characteristics of the parcel (in particular, its size) on a competitive market without imperfections or failures. The competition between numerous development companies operating across the whole country makes such a competitive market assumption plausible. As in our theoretical model, we assume that the movers drive the size and the location of new development. In other words, households are price takers because they are small agents, but they are 'quantity makers' because the highest bidder on the market can choose her desired residential plot size. It is with these households that we are able to observe a residential location "decision" that reflects preferences for proximity and indoor/outdoor space.

While our research question could in principle be addressed by using all residential development in France, it would be inappropriate to apply our model to the entire housing market, which is segmented along several dimensions. First, the housing market is made of newly constructed and older dwellings. Our focus is on marginal rather than cumulative patterns, so we select relatively new development, after 1974, result of contemporaneous state of the word. Second, we consider that renters and owners belong to separate residential markets, with different preference structures, and restrict our analysis on owner-occupier behavior. Third, we limit our analysis to households who have recently moved (in relation to the Household Survey date). We are interested in households' residential decisions, and household attributes obtained from the Surveys will tend to more accurately describe these households at the time of its moving decision.

---

[11] Due to correlation between climatic variables, we drop the measure of days of precipitation in July.

We exclude farmers and retired people from consideration because they do not have the same relationship to the central cities of urban areas in terms of proximity to job centers as other households. Finally, we drop the urban area of Paris, which is distinct due to its large population (around 20% of the French populace): its scale is very different from every other urban area and would be difficult to accommodate under a single econometric model.

To be more specific, the population of interest consists of the 16,947 households that satisfy the following criteria:

1. Residence was constructed after 1974
2. Residence is owner-occupied
3. Household moved to its current residence within 8 years of being surveyed
4. Household is not a farmer
5. At least one adult member of the household is not retired
6. Household is located outside the urban area of Paris

6.2 The selection problem

We assume that the selection associated with defining the population of interest does not introduce bias, as all criteria are intended to delineate a specific market. However, remaining are two major econometric challenges related to the definition of the two dependent variables, residential plot size S and distance D. The first challenge is that residential plot size is equal to zero for a significant number of dwellings in the population of interest. These dwellings are generally apartments, and the meaning of $S = 0$ is not the same for a single detached house and an apartment. Moreover, some unobserved factors can influence the choice between flat and house, the lot size and the distance. As such, we introduce a selection equation explaining a household's choice between a single detached house and an apartment, and then control for this potential bias.

The second challenge arises from our measurement of distance. Distance is equal to the driving time in minutes from the center of a household's local commune to the center of the commune that is the central city of the nearest *Aire Urbaine*. It follows that distance is equal to zero for all households located in the commune that is the central city of its *Aire Urbaine* with is not the true distance. Moreover, these zero values indicate more than proximity. They indicate the decision to live in very particular type of commune: the heart of an *Aire Urbaine*, which is a large city, with ancient housing, and many structural constraints. In this context, the assumption of 'quantity maker' is questionable. Hence, we introduce a second selection equation, explaining the choice between location in the core of the *Aire Urbaine* and a location in its periphery.

Among the 16,947 observations in the population of interest, 1,775 (10.4%) are households living in apartments and 2,718 (16.0%) of the dwellings are in the central commune of their respective *Aire Urbaine*. A total of 13,377 (78.9%) households reside in single-detached houses outside of urban area city centers.

There exist standard econometric procedures to deal with selection issues in multivariate settings. In a single-equation setting, parametric or semiparametric procedures are available to correct for endogenous selection. Under an assumption of normality, multiple sources of selection can also be dealt with (see Yen, Lin and Smallwood (2003) for a survey of multivariate selection methods). However, the fact that our model consists of simultaneous equations requires a slightly more involved estimation approach to deal both with selection and simultaneity bias.

6.3 Estimation Strategy

The underlying econometric structural model can be written as a system of latent (unobserved) variables:

$$\begin{cases} \ln SIZE^{*} = X_{1}\beta_{1} + \ln DIST^{*}\gamma_{1} + \varepsilon_{1}, \\ \ln DIST^{*} = X_{2}\beta_{2} + \ln SIZE^{*}\gamma_{2} + \varepsilon_{2}, \end{cases} \tag{14}$$

where $SIZE^{*}$ is the latent size of the residential plot, and $DIST^{*}$ is the latent driving distance in minutes to the central city, and ε_{1} and ε_{2} are random terms. X_{1} and X_{2} are two vectors of observed explanatory variables. The corresponding reduced-form model is:

$$\begin{cases} \ln SIZE^{*} = X_{1}\pi_{11} + X_{2}\pi_{12} + u_{1}, \\ \ln DIST^{*} = X_{1}\pi_{21} + X_{2}\pi_{22} + u_{2}, \end{cases} \tag{15}$$

where (u_{1}, u_{2}) are assumed jointly Normal $N(0_{2}, \Sigma)$.

As mentioned above, we need to control for possible selection biases. In terms of observed variables, the following selection and censoring conditions apply to the reduced-form model:

$$\begin{cases} HOUSE = 1 \text{ if } HOUSE^{*} > 0 \Leftrightarrow \varepsilon_{3} > -X_{3}\beta_{3}, \\ HOUSE = \ln SIZE^{*} \text{ if } HOUSE^{*} > 0 \Leftrightarrow \varepsilon_{3} > -X_{3}\beta_{3}, \\ \ln DIST = \ln DIST^{*} \text{ if } \ln DIST^{*} > 0 \text{ and } HOUSE^{*} > 0 \\ \quad \Leftrightarrow u_{2} > -X_{1}\pi_{21} - X_{2}\pi_{22} \; and \; \varepsilon_{3} > -X_{3}\beta_{3} \end{cases} \tag{16}$$

where $(u_{1}, u_{2}, \varepsilon_{3})$ are jointly Normal $N(0_{3}, \Sigma)$.

One possibility is to estimate the model in reduced form while accounting for selection and the censored dependent variable, and then to solve for the structural parameters. However, this Indirect Least Squares approach requires the model to be exactly identified, which reduces the number of exogenous regressors that can be candidates in the size and distance equations. Another possibility is to construct the likelihood of the model based on the normality assumption above. We consider a simpler approach that corrects for the selection and simultaneity bias in a multi-step procedure based entirely on Probit and 3SLS models.

Using a linear regression framework (2SLS, 3SLS), requires the following conditional expectations:

$$E(\ln SIZE^{*}|\varepsilon_{3} > -X_{3}\beta_{3}) = E(\ln SIZE|\varepsilon_{3} > -X_{3}\beta_{3}) = X_{1}\pi_{11} + X_{2}\pi_{12} + E(u_{1}|\varepsilon_{3} > -X_{3}\beta_{3}), \tag{17}$$

and

$$\begin{aligned} & E(\ln DIST^{*}|\varepsilon_{3} - X_{3}\beta_{3}, u_{2} > -X_{1}\pi_{21} - X_{2}\pi_{22}) \\ & = E(\ln \mathrm{DIST}\,|\varepsilon_{3} > -X_{3}\beta_{3}, u_{2} > -X_{1}\pi_{21} - X_{2}\pi_{22}) \\ & = X_{1}\pi_{21} + X_{2}\pi_{22} + E(u_{2}|\varepsilon_{3} > -X_{3}\beta_{3}, u_{2} > -X_{1}\pi_{21} - X_{2}\pi_{22}) \end{aligned} \tag{18}$$

The model is estimated in four steps. First, we estimate a bivariate probit model on $HOUSE = 1$ (explained by the covariates in X_3) and $lnDISTANCE > 0$ (explained by X_1 and X_2). We use three variables as instruments in the house equation: a variable indicating the household undertook energy conserving renovations; a variable indicating the home loan did not cover 100% of the investment; and a variable indicating the household received social help to pay the loan. We find no correlation between these variables and the size of the garden or distance to the city center. The bivariate probit (House=1, distance>0) shows a strong correlation between the two equations (as expected): the decisions to live in a house and outside of the central commune of the urban area are positively correlated ($\rho = 0.62\ (0.02)$).

In a second step, the equation for $lnSIZE$ (explained by X_1 and X_2) is estimated by Heckman's method on the subsample of observations where $HOUSE = 1$, and we compute a vector of predicted values for $lnSIZE$. These two estimations enable the construction of a bivariate selection term as follows:[12]

$$E(u_2 | \varepsilon_3 > -X_3\beta_3, u_2 > -X_1\pi_{21} - X_2\pi_{22}) = \theta_1\lambda_1 + \theta_2\lambda_2,$$

where

$$\begin{aligned}
\lambda_1 &= \phi(X_3\beta_3) \times \Phi(Z_1^*) / \Phi_2(X_3\beta_3, X_1\pi_{21} + X_2\pi_{22}, \rho),\\
\lambda_2 &= \phi(X_1\pi_{21} + X_2\pi_{22}) \times \Phi(Z_2^*) / \Phi_2(X_3\beta_3, X_1\pi_{21} + X_2\pi_{22}, \rho),\\
Z_1^* &= \left[(X_1\pi_{21} + X_2\pi_{22}) - \rho(X_3\beta_3)\right] / \sqrt{(1-\rho^2)},\\
Z_2^* &= \left[X_3\beta_3 - \rho(X_1\pi_{21} + X_2\pi_{22})\right] / \sqrt{(1-\rho^2)},
\end{aligned} \tag{19}$$

$\phi(.)$ and $\Phi(.)$: density and c.d.f. of the Normal distribution $N(0,1)$,

$\Phi_2(.,.,.)$: bivariate c.d.f. (probability estimated from the bivariate Probit),

ρ: correlation coefficient between ε_3 and u_2 (obtained from Bivariate Probit),

θ_1 and θ_2: parameters to be estimated.

Step 3 estimates the equation for $lnDIST$ (explained by X_1 and X_2) on the subsample of observations where $HOUSE = 1$ and $lnDIST > 0$, inserting the bivariate selection terms λ_1 and λ_2 with the corresponding parameters θ_1 and θ_2. A vector of predicted values is computed for $lnDIST$. Step 4 returns to the system of structural equations, and replaces $lnSIZE$ and $lnDIST$ with their predicted values. The system is estimated on the subsample of observations where $HOUSE = 1$ and $lnDIST > 0$, and includes the bivariate selection terms (λ_1 and λ_2) as before. Standard errors are estimated using the bootstrap method.

7. Results

Results of the fourth step are presented in Table 1. Plot size and distance are positively linked and simultaneously determined, with larger lot sizes associated with location further from the

---

[12] For details about the derivation of these correction terms, see Ham (1982), Gouriéroux (2000), and Lacroix and Thomas (2011).

city center. A positive selection process occurs: people who choose to live in a house have unobserved characteristics that lead them to choose a bigger garden (the Inverse Mills Ratio is positive and significant in the lot size equation). In the distance equation, λ_1 is positive but insignificant and $\lambda_2$is positive and significant, controlling for the fact that people who choose to live outside the city center have unobserved characteristics associated with preference for more distant location. Due to the selection treatments, the parameters estimated in Step 4 cannot be interpreted as elasticities. The computation of elasticities is straightforward in the case of the lot size equation, but quite cumbersome in the case of the distance equation as λ_1 and λ_2 must be derived (Greenwood, 1996, 1998). We compute elasticities for the climate variables over whole sample.[13]

7.1 Control variables

Most of the variables controlling for household characteristics behave as expected. Wealthier households locate on larger plots and closer to the central commune of the urban area, which is consistent with the behavior of French households (see Brueckner *et al.*, 1999). A household's socio-professional status adds supplementary information: as expected, blue collar workers locate on smaller lots and at greater distance than executive and office workers (also consistent with French characteristics).

The presence of two working people in the household is associated with larger lot sizes (income effect), and the indicator's coefficient in the distance equation is of the expected sign (negative due higher commute time constraints), but not statistically significant. Lot size increases with the age of the head of household and distance to the city center slightly decreases. In interpreting the results, keep in mind the age distribution in our sample: only 10% are younger than 30 and 80% are between 30 and 53 and as a result, the effect of age is weak. Older (i.e. senior) households presumably have income allowing them to locate closer to the city center. The presence of children has a positive effect on lot size (significant at the 10% level) and no effect on the distance; variables accounting for age and the presence of multiple commuters may be capturing the effects related to household composition.

Variables controlling for regional and local variation are also well behaved. Lot sizes decrease with the population level of the urban area because densities (and thus land values) increase, conducing to smaller plots. Distance to the center of the urban area increases with urban area population: the larger the urban area the more distant its edge, thus increasing the commuting distance to its center. Higher regional land prices and coastal communes are associated with smaller lot sizes. Moreover, lot size is positively linked with the average income of the commune inhabitants, which may be due to minimum-lot-size zoning regulations designed to exclude poorer households from affluent communes.

[TABLE 1 ABOUT HERE]

---

[13] Details of the computation are available from the authors upon request.

7.2 Climate variables

The theoretical model implies that near the CBD, lot size is unambiguously greater under a warmer climate. Away from the CBD, the effect of warming on lot size is dependent on (1) the influence of climate on households' preferences for outdoor lifestyles, and (2) the climate's influence on commuting cost. When the first outweighs the second, lot size increases with the climatic amenity, and the net effect of climate is ambiguous when the dominant relation is reversed. The empirical results suggest that (1) predominates: the parameter on the log of July temperature is 0.345 ± 0.1 (95% level) in the lot size equation.

We calculate the elasticity of the lot size in response to a change in climate for a representative individual defined by average values (continuous variables) and reference categories (discrete variables).[14] A 10% increase in average July temperature (2 °C) leads to a 4.9% increase in lot size (54 m² on average). For example, lot sizes in communes with temperatures in the 90th percentile, as in much of the south of France (21°C), are 10.5% larger than those in the 10th percentile as in most areas north of the Paris-Strasbourg line (17.3°C), an increase of approximately 115 m² on average.[15]

The theoretical model also predicts that people locate farther from the center when the climatic amenity increases. Our results are consistent with this prediction: higher January temperatures correspond to more distant locations (the calculated elasticity for the same representative individual is 0.08), but days of January precipitation has no significant effect on proximity. The variation in average January temperatures in France is substantial: the average temperature is equal to 1.4°C at the 10th percentile and 5.8°C at the 90th percentile, with a median of 3.6°C. Distance to the city center is 4% higher (resp. lower) in a commune in the 90th percentile (resp. 10th percentile) than in a commune where the temperature is at the median value. The average distance to the urban area center is approximately 23 minutes, with the variation at this average is close to 1 minute. The insignificant effect of precipitation may be due to the small variation across the French territory in January. Indeed, 80% of the communes record between 9 and 13 rainy days in January and more than 40% of the communes around 12 days.

An advantage of using household-level data is that it allows estimations according to the personal characteristics of the households. Many possibilities may be explored and may be extensions for future work: for example, examination of younger vs. older households, or singles vs. families with children. To illustrate these possibilities we estimate the system of equations for two specific social categories: households with a blue collar worker(s) and those with an executive(s). We do not observe significant differences between the two sub-populations in the behavior of the control variables (at most, some variation in magnitude). The only noticeable difference concerns the Inverse Mills Ration (IMR): the parameters are no longer significant,

---

[14] Age = 38; Children = Yes; Urban Area population = 50,000 to 100,000; Average July Temperature = 20°C; Average January Temperature = 5°C; Rainy January Days = 11. Formulae are available on request. Standard errors are bootstrapped.

[15] Applying the correction for selection to the parameter on average July temperature in the lot size equations leads to a slight increase of the elasticity.

suggesting that social sorting occurs between these two populations rather than within them (there is no variation in the Lambdas).

[TABLE 2 ABOUT HERE]

On the one hand, blue collar workers are sensitive to accessibility: all things being equal, they locate closer to the center than other social categories when winter climatic conditions are harsher. On the other hand, they are less sensitive to lot size. Indeed, the parameter on average July temperature is insignificant. In contrast, executives demand for land appears to increase under warmer summer climates, but climatic conditions in winter do not influence their preferences for proximity.

These results are consistent with the standard case in urban economics. As Wheaton shows (1974, p. 620-621): "With an income inelastic land demand and noticeably greater commuting costs for the wealthy, greater income leads to more central locations. If land demands are income elastic and commuting expenses relatively fixed, income should increase with distance from work". The second case is the more common, even if the first occurs in some European cities due to downtown amenities sought by the wealthy (Paris is a classic example: see Brueckner *et al*., 1999). Overall, the income elasticity of land demand in France is 0.27 for executive households and 0.22 for blue collar households, while the income elasticity of demand for accessibility is 0.11 for the former and 0.30 for the latter (Cavailhès, 2005). Therefore, executives (blue collars) are more (less) "land elastic" and less (more) "distance elastic": where July temperature is high (January temperature cold), they increase land consumption (decrease commuting distance).

7.3 Robustness checks

As previously noted, distance increases with urban-area population, and lot size decreases. On average, lot sizes (distances) are 20% smaller (larger) in urban areas with at least 200,000 inhabitants than in urban areas with less than 200,000 inhabitants. To test if the mechanisms driving preferences for lot size and proximity behave differently in larger and smaller urban areas, we estimate models using two sub-samples defined by population level (see **Table 3**).

Selection mechanisms in the two sub-samples behave as in the whole sample, with the exception of the IMR in large urban areas (which control for the fact that we select households who choose to live in a single-detached-house vs. an apartment), which is insignificant. In the large urban areas, lot size appears to be driven by market constraints more than by household preferences. While wealthier households do locate on larger lots, the influences of other household characteristics are no longer significant. Lot size increases with the age of the head of household, but not with children or with the fact that two workers commute in the household. The effect of distance on lot size behaves in the same way, but households must locate much farther from the urban center to obtain larger lots.

Regarding the climatic variables, we find some differences in the magnitude of the effect of temperature: the temperature in July has a positive effect on lot size in the smaller urban areas, but the effect is negative and insignificant in the larger urban areas. The distance to the city center increases with the temperature in January in both sub-samples, but at a lower rate in the smaller urban areas. The distance to city center decreases with the number of days of precipitation in January in large urban areas, but the effect is not significantly different from zero for households living in smaller urban areas. These two last results are consistent with the notion that commuting conditions are of greater concern to households in a large metropolis than to those in smaller cities.

[TABLE 3 ABOUT HERE]

8. Conclusion

The process of urban sprawl and its primary determinants have been well identified by urban economics models and their empirical extensions. Primary among the determinants of urban sprawl are falling commuting costs and rising incomes. The tradeoff central to households' residential decisions is that between accessibility and the consumption of residential land. We build on this understanding of a household's residential behavior by introducing a climatic amenity into its micro-economic program. We consider how climate may affect this central tradeoff between accessibility and residential consumption. In particular, we hypothesize that in warmer climates, commuting costs are lower because the journey is faster and less hazardous, and that the preference for outdoor lifestyles is stronger. Consequently, households tend to locate farther from urban centers and on larger plots of land where climate is warmer. Our urban economics model is a closed city time-extended model, in which the household consumes a climatic amenity that affects its taste for the residential good and the unit commuting cost. The primary results derived from this model are as follows: First, more people locate farther from the CBD where climate is warmer. Second, in warm regions the lot size is unambiguously larger close to the CBD, but the effect may be reversed at the periphery depending on consumers' sensitivity to climate's effects on lot size and commuting cost: lot sizes may be larger (smaller) in outlying regions if the lot-size effect is greater (smaller) than the commuting-cost effect.

Empirically, we maintain our focus on the trade-off between accessibility to the CBD and residential consumption. We use household-level data from French housing surveys to model the household's simultaneous choice of accessibility and lot size. In this way, we model a household's sprawling residential behavior rather than the level of an urban area's sprawl. This approach stands in contrast to previous empirical analyses of urban sprawl which use urban-area level measures of sprawl.

The empirical results have well behaved control variables that are consistent with expectations derived from urban economics studies. The significant temperature variables in both the distance and lot size equations suggest that the climatic amenity does influence a household's residential preferences. In particular, we find that households occupy larger lots in warmer summer climates and locate farther from urban centers in milder winter climates. We also find evidence that the behavioral effects of climate vary with household characteristics. We look at one type of characteristic (socioeconomic status), finding that households of executive status are more sensitive to the influence of summer climate on lot size, while blue collar households are more sensitive to the influence of winter climate on accessibility. The exploration of other factors is potentially grounds for future work. Economists traditionally study the effects of changes in the state of the world on economic equilibriums for a given set of consumer preferences, leaving the preferences themselves to be investigated by psychologists or sociologists. We show here how economists can improve knowledge of the behavioral consequences of consumers facing different climatic conditions.

As societies continue to face the environmental and social challenges posed by climate change and urban sprawl, it is important to improve our understanding between these two phenomena. The contribution of urban sprawl to climate change is well known. Here, we show a reverse relationship between climate and sprawling residential behavior: household's residential preferences are influenced *by* climate conditions. Our results are obtained in cross-section, but if they hold over time, they predict that global warming will increase households' demand for larger lot sizes farther from urban centers. Put another way, global warming and sprawling residential behavior strengthen one another in a vicious circle.

References

Alonso, W., 1964. Location and Land Use. Cambridge: Harvard University Press.

Anas, A., Arnott, R., 1997. Urban spatial structure, Journal of Economic Literature 36(3), 1426-1464.

Baccaïni B., 2005. Enquêtes annuelles de recensement: résultats de la collecte 2004, INSEE-Première no.1028

Bart, I.L., 2010. Urban sprawl and climate change: A statistical exploration of cause and effect, with policy options for the EU, Land Use Policy 27(2), 283-292.

Brueckner, J.K., 1987. The structure of urban equilibria: A unified treatment of the Muth-Mills model, Handbook of Regional and Urban Economics, vol. 2, pp. 821–845.

Brueckner, J.K., 2000. Urban sprawl: diagnosis and remedies, International Regional Science Review 23(2), 160-171.

Brueckner, J.K., 2001. Urban sprawl: Lessons from Urban Economics, Brookings-Wharton Papers on Urban Affairs 2001(1), 65-97.

Brueckner J.K., Thisse J.F., Zenou, Y., 1999. Why is central Paris rich and downtown Detroit poor? An amenity-based theory, European Economic Review, 43, 91-107.

Burchfield, M., Overman, H.G., Diego P., Turner, M.A., 2006. Causes of sprawl: A portrait from space. Quarterly Journal of Economics 121(2), 587-634.

Cavailhès J., 2005. Le prix des attributs du logement, Economie et Statistique 381(1), 91-123.

Cheshire, P.C., Magrini, S., 2006. Population growth in European cities: weather matters-but only nationally, Regional Studies 40, 23-37.

Christaller, W., 1933. Central places in Southern Germany. Englewood Cliffs: Prentice-Hall.

European Commission, 2006. Urban sprawl in Europe: The ignored challenge, Directorate General Joint Research Center, Copenhagen, Denmark.

Ewing, R.H., Pendall, R., Chen, D., 2002. Measuring sprawl and its impact.

Fujita, M., 1989. Urban Economic Theory: Land Use and City Size, Cambridge University Press, Cambridge, U.K.

Galster, G., Hanson, R., Ratcliffe, M.R., Wolman, H., Coleman, S., Freihage, J., 2001. Wrestling sprawl to the ground: defining and measuring an elusive concept, Housing Policy Debate 12(4), 681–718.

Glaeser, E.L., Kahn, M.E., 2004. Sprawl and urban growth, in: Mills, E. (Ed), Handbook of regional and urban economics, vol. 4, North-Holland, Elsevier, Netherlands, pp. 2481–2527.

Glaeser, E.L., Kahn, M.E., Chu, C., 2001. Job sprawl: Employment Location in US Metropolitan Areas. Brookings Institution, Center on Urban and Metropolitan Policy, Washington D.C.

Gouriéroux, C., 2000. Econometrics of Qualitative Dependent Variables. Cambridge University Press, Cambridge, U.K.

Graves, P.E., 1976. A reexamination of migration, economic opportunity, and the quality of life, Journal of Regional Science 16, 107-112.

Graves, P.E., 1980. Migration and climate, Journal of regional Science 20, 227-237.

Graves, P.E., Linneman, P.D., 1979. Household migration: theoretical and empirical results, Journal of Urban Economics 6, 383-404.

Ham, J.C., 1982. Estimation of a labour supply model with censoring due to unemployment and underemployment, Review of Economic Studies 49(3), 335-354.

Irwin, E.G., Bockstael, N.E., Cho, H.J., 2006. Measuring and modeling urban sprawl□: Data, scale and spatial dependencies, Invited paper prepared for the Urban Economic Sessions, 53rd Annual North American Regional Science Association Meetings of the Regional Science Association International, November 16-18, 2006, Toronto, Canada.

Irwin, E.G., Bockstael N.E., 2007. The evolution of urban sprawl: evidence of spatial heterogeneity and increasing land fragmentation, Proceedings of the National Academy of Sciences of the United States of America 104(52), 20672-20677.

Irwin, E.G., Bell, K.P., Bockstael, N.E., Newburn, D.A., Partridge, M.D., Wu, J., 2009. The economics of urban-rural space, Annual Review of Resource Economics 1(1), 435-459.

Joly, D., Brossard, T., Cardot, H., Cavailhès, J., Hilal, M., Wavresky, P., 2011. Temperature interpolation based on local information: the example of France, International Journal of Climatology 31, 2141-2153.

Kasanko, M., Barredo, J.I., Lavalle, C., McCormick, N., Dmicheli, L., Sagris, V., Brezger, A., 2006. Are European cities becoming dispersed? A comparative analysis of 15 European urban areas, Landscape and Urban Planning 77(1), 111-130.

McMillen, D.P., McDonald, J.F., 1991. Urban land value functions with endogenous zoning, Journal of Urban Economics, 19, 14-27.

Mieszkowski, P., Mills, E.S., 1993. The causes of metropolitan suburbanization, The Journal of Economic Perspectives 7(3), 135-147.

Mills, E.S., 1967. An aggregative model of resource allocation in a metropolitan area, The American Economic Review 57(2), 197-210.

Muth, R. 1969. Cities and Housing. University of Chicago Press, Chicago, IL.

Nechyba, T.J., Walsh, R.P., 2004. Urban sprawl, The Journal of Economic Perspectives 18(4), 177-200.

Newman, P., Kenworthy, J., 1999. Sustainability and Cities: Overcoming Automobile Dependence. Island Press, Washington D.C.

Patacchini, E., Zenou, Y., 2009. Urban Sprawl in Europe, Brookings-Wharton Papers on Urban Affairs 25(2), 125–149.

Rosen, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition, *Journal of Political Economy* 82: 34-55.

Smith, L.B., Rosen, K.T., Fallis, G., 1988. Recent developments in economic models of housing markets, Journal of Economic Literature 26, 29-64.

Song, Y., Knaap, G.J., 2004. Measuring urban form: Is Portland winning the war on sprawl?, Journal of the American Planning Association 70(2), 210-225.

Torres, P.M., Alberti, M., 2000. Measuring Sprawl, Center for Advanced Spatial Analysis, Working Paper (27).

Wheaton, W.C., 1974. A comparative static analysis of urban spatial structure, Journal of Economic Theory 9(2), 223-237.

Wheaton, W.C., 1977. Income and urban residence: An analysis of consumer demand for location, Urban Studies 67, 620-631.

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., Sarzyniski, A., 2005. The fundamental challenge in measuring sprawl: Which land should be considered?, The Professional Geographer 57(1), 94–105.

Wu, J., 2006. Environmental amenities, urban sprawl, and community characteristics, Journal of Environmental Economics and Management 52(2), 527-547.

Yen, S.T., Lin, B.H., Smallwood, D.M., 2003. Quasi- and simulated likelihood approaches to censored demand systems: food consumption by food stamp recipients in the United States, American Journal of Agricultural Economics 85(2), 458-478.

Tables and Figures

Table 1. Structural equation estimates (Step 4)

Variables	Coefficients	Bootstrapped Std. Errors
Lot size equation		
Ln_Distance	0.802	0.076
Ln_Income	0 .289	0.021
Age	0.002	0.0009
Children (yes)	0.042	0.023
Two-commuters (yes)	0.057	0.017
Blue collar worker	-0.038	0.018
Office worker	-0.065	0.024
Intermediare	Reference	Reference
Executive	0.036	0.023
Other	0.088	0.026
Land price	-0.005	0.0003
Land price trend	-0.254	0.063
Pop area less than 50 000	Reference	Reference
pop area 50 000 to 100 000	-0.127	0.022
pop area 100 000 to 200 000	-0.233	0.021
pop area 200 000 to 500 000	-0.401	0.028
pop area more than 500 000	-0.661	0.044
Ln_Local income	0.151	0.022
Coastal commune	-0.297	0.027
Ln_July temperature	0.345	0.098
Inverse Mills Ratio	0.211	0.099
Intercept	-0.456	0.523
Distance equation		
Ln_Lot size	0.132	0.0298
Ln_Income	-0.130	0.014
Age	-0.002	0.0006
Children (Yes)	-0.007	0.0128
Two-commuters (Yes)	-0.018	0.012
Blue collar worker	0.069	0.015
Office worker	0.007	0.018
Foreman	Reference	Reference
Executive	-0.027	0.017
other	0.066	0.020
Share difference	0.007	0.022
Pop area less than 50 000	Reference	Reference
pop area 50 000 to 100 000	0.153	0.018
pop area 100 000 to 200 000	0 .164	0.021
pop area 200 000 to 500 000	0.285	0.023
pop area more than 500 000	0.513	0.029
Forest rate	0.546	0.040
Ln_January temperature	0.075	0.013
Ln_January rainy days	0.031	0.022
Lambda 1	0.086	0.060

Variables	Coefficients	Bootstrapped Std. Errors
Lambda 2	0.163	0.064
Intercept	2.999	0.210

Table 2. Climate variable parameters on blue collar and executive sub-samples

	July temperature in lot size equation	January temperature in distance equation	January rainy days in distance equation
Blue collar workers	0.345	0.130	-0.104
	(0.252)	(0.028)	(0.050)
Executives	0.488	0.0051	0.159
	(0.164)	(0.018)	(0.141)

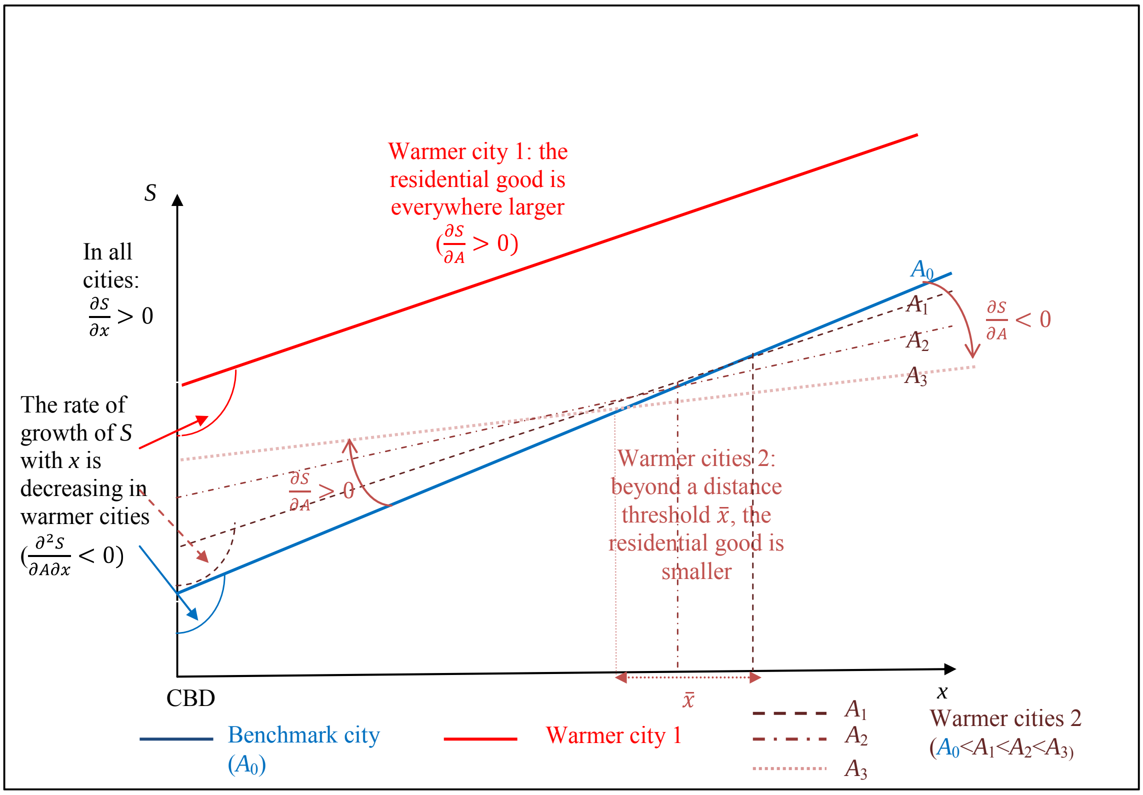
Note: Bootstrapped standard errors are in parentheses.

Table 3: Structural equation estimates (Step 4) for small and large urban areas

	Population <= 200 000 inhabitants		Population > 200 000 inhabitants	
Variables	Coefficients	Bootstrapped Std. Error	Coefficients	Bootstrapped Std. Error
Lot size equation				
Ln_Distance	0.579	0.086	1.11	0.090
Ln_Income	0.273	0.024	0.283	0.033
Age	0.000	0.001	0.005	0.001
Children (Yes)	0.045	0.027	0.031	0.035
Two-commuters (Yes)	0.080	0.022	0.022	0.028
Blue color worker	-0.047	0.020	-0.015	0.031
Office worker	-0.079	0.029	-0.033	0.035
Intermediare worker	Reference	Reference	Reference	Reference
Executive worker	0.056	0.035	0.029	0.033
Other worker	0.079	0.035	0.135	0.042
Land price	-0.007	0.0004	-0.004	0.0004
Land price trend	-0.281	0.087	-0.233	0.098
Ln_Local income	0.121	0.027	0.123	0.036
Coast	-0.284	0.030	-0.336	0.047
Ln_July temperature	0.475	0.119	-0.253	0.169
IMR	0.299	0.159	0.108	0.118
Intercept	0.337	0.679	0.022	0.679
Distance equation				
Ln-Lot size	0.159	0.031	0.017	0.050
Ln_Income	-0.154	0.022	-0.075	0.022
Age	-0.002	0.0009	-0.002	0.0008
Children (Yes)	-0.014	0.021	-0.010	0.022
Two-commuters (Yes)	-0.041	0.017	0.033	0.016
Blue color worker	0.074	0.020	0.042	0.019
Office worker	0.021	0.027	-0.029	0.023
Intermediare worker	Reference	Reference	Reference	Reference
Executive worker	-0.031	0.028	-0.007	0.019
Other worker	0.084	0.030	0.049	0.023

Variables	Population <= 200 000 inhabitants		Population > 200 000 inhabitants	
	Coefficients	Bootstrapped Std. Error	Coefficients	Bootstrapped Std. Error
Share difference	-0.062	0.098	0.029	0.028
Forest rate	0.483	0.046	0.599	0.068
Ln_January temperature	0.052	0.014	0.185	0.017
Ln_January rainy days	0.178	0.233	-0.096	0.028
Lambda 1	0.054	0.152	0.024	0.061
Lambda 2	0.349	0.093	0.012	0.079
Intercept	2.799	0.272	3.718	0.282

Figure 1. Rate of growth of S with x, according to the climatic amenity A



Appendix: Bivariate Probit Results (Step 1)

Variables	Coefficients	Std. Errors
House		
Ln_Income	0.074	0.042
Age	0.002	0.002
Marital status (single)	-0.987	0.055
Blue collar worker	0.134	0.047
Office worker	-0.126	0.049
Intermediare worker	Reference	Reference
Executive worker	-0.109	0.046
Other worker	0.051	0.055
Land price trend	-0.562	0.121
Pop area less than 50 000	Reference	Reference
pop area 50 000 to 100 000	-0.133	0.062
pop area 100 000 to 200 000	-0.344	0.058
pop area 200 000 to 500 000	-0.633	0.049
pop area more than 500 000	-0.877	0.048
Population of commune (1 000s)	-0.0007	0.0001
Coastal commune	-0.385	0.044
Ln_July temperature	-1.460	0.162
College degree (yes)	-0.107	0.034
Active in workforce	0.130	0.028
Energy conserving renovations	0.294	0.080
Social help for housing	0.264	0.043
Self-financing	-0.222	0.051
Intercept	5.908	0.677
suburb		
Ln income	-0.078	0.032
Age	-0.008	0.001
Children (Yes)	0.123	0.031
Two-commuters (Yes)	0.028	0.029
Blue collar worker	0.270	0.034
Office worker	-0.022	0.040
Intermediare worker	Reference	Reference
Executive worker	-0.135	0.038
Other worker	0.252	0.044
Share difference	0.113	0.052
Land price	-0.0003	0.0004
Land price trend	-0.060	0.107
Population less than 50 000	Reference	Reference
Population 50 000 to 100 000	0.327	0.040
Population 100 000 to 200 000	0.494	0.042
Population 200 000 to 500 000	0.395	0.038
Population more than 500 000	0.632	0.039
Forest rate	1.675	0.090
Ln_Local income	0.154	0.034
Coastal commune	-0.675	0.038
Ln_July temperature	-2.627	0.260

Variables	Coefficients	Std. Errors
Ln_January temperature	-0.019	0.030
Ln_January rainy days	-0.479	0.089
Intercept	9.077	1.076
rho	0.609	0.020