UNIVERSIDAD DE CANTABRIA

E.T.S. de Ingenieros de Caminos, Canales y Puertos

Tesis Doctoral

Modelos econométrico – espaciales para el estudio de los impactos del transporte en los usos del suelo

Autor

RUBÉN CORDERA PIÑERA

Directores

LUIGI DELL´OLIO

ÁNGEL IBEAS PORTILLA

Santander, 2014

Ipsa scientia potestas est

RESUMEN

Aumentar el nivel de conocimiento referente a las interrelaciones entre los sistemas de transporte y de usos del suelo es un reto fundamental de cara a impulsar políticas públicas encaminadas a promover la sostenibilidad de los sistemas urbanos. En muchas ciudades se ha desarrollado un equilibrio no deseado formado por un círculo vicioso entre los usos del suelo y el transporte donde el empleo creciente del automóvil ha ido acompañado de una mayor dispersión urbana. Ambos fenómenos generan problemas complejos de resolver como un aumento de los tiempos de viaje, mayor congestión, contaminación y un uso ineficiente del suelo y de los recursos energéticos.

Para resolver este tipo de problemas se han propuestos diversas medidas como la inversión en nuevas infraestructuras de transporte público. Este tipo de medidas además de afectar a los costes de viaje también pueden tener otros impactos sobre el sistema de usos del suelo en materia de localización de los hogares y las empresas y del valor de los bienes inmobiliarios. Estos impactos pueden ayudar a alcanzar los objetivos de una mayor sostenibilidad económica y ambiental si los aumentos de accesibilidad que proporcionen las medidas implantadas llevan aparejados incrementos en el valor de los bienes inmobiliarios y una mayor preferencia de hogares y empresas por localizarse de forma menos dispersa. Sin embargo la medición de estos impactos es compleja por lo que se necesitan modelos matemáticos que ayuden a determinarlos.

El objetivo de este estudio es por lo tanto especificar una serie de modelos econométrico – espaciales que permitan estimar los impactos que distintas políticas y proyectos de transporte pueden generar sobre los usos del suelo en el medio y largo plazo. La zona de estudio utilizada es el área metropolitana de Santander, una ciudad media localizada al norte de España. Los modelos estimados considerarán en todos los casos la fuerte componente espacial de los fenómenos que se intenta modelizar. Para ello se han utilizado una serie de técnicas econométrico – espaciales que permiten tener en cuenta las relaciones de dependencia que pueden darse entre observaciones.

Los resultados obtenidos permiten afirmar que una mayor accesibilidad es un factor significativo a la hora de hacer más probable la localización de hogares y empresas y de aumentar el valor de los bienes inmobiliarios. Sin embargo no todos los indicadores de accesibilidad utilizados resultaron significativos y la cercanía a las estaciones de tren puede disminuir más que aumentar el valor de los inmuebles.

Las técnicas econométricas – espaciales demostraron mejorar el ajuste de los modelos de forma significativa y evitaron problemas derivados de posibles sesgos o ineficiencias en los parámetros estimados. Además las distintas matrices de contigüidad utilizadas en los modelos de regresión espacial no modificaron en gran medida las estimaciones.

Hay que resaltar que esta investigación se basa en la compilación de tres artículos revisados por pares y publicados en revistas indexadas (JCR), lo cual apoya los resultados obtenidos.

ABSTRACT

Increasing the knowledge on the relationships between transportation and land-use systems is a major concern for the implementation of policies aimed to promoting the urban systems sustainability. Many cities have been caught in a vicious circle of a non-desired equilibrium between the land-use and transportation systems, where the increasing demand of the automobile has been accompanied by an increasing urban sprawl. Both these phenomena usually generate complex problems that are sometimes difficult to solve, such as increased travel times, higher congestion, pollution and an inefficient use of land and energy resources.

To solve such problems several policies have been proposed, such as new public transportation infrastructures. This kind of measures not only affects travel costs but also may have other impacts on the land-use system, in terms of households and firms location and real estate values. These impacts can help to achieve an increased economic and environmental sustainability if the increase in accessibility due to the implemented measures also carries a raise in the real estate values and a greater preference on the part of households and firms for settling according to a less dispersed structure. However, it's complex to measure these impacts, being necessary to use mathematical models.

The main objective of this study is to specify a series of econometric spatial models to estimate the impacts of different transport policies and projects on land-use in the medium- and long-term. The study area used in this study is the metropolitan area of Santander, a medium-sized city located in northern Spain. The estimated models are sensitive to the spatial component of the phenomena allowing modelling the dependency relationships between observations.

The results confirm that a greater accessibility in a given area is a significant factor to increase the probability of households and firms to locate themselves in that particular area, as well as to increase the real estate values. Notwithstanding, not every accessibility indicator was found significant and, in fact, the greater accessibility to train stations the lower the real estate values.

The use of spatial econometric techniques improves significantly the goodness of fit of the models and avoids bias or inefficiency problems in the estimated parameters. It should be noted that the different contiguity matrices used in the spatial regression models do not considerable change the estimations.

This research is based on the compilation of three peer-reviewed papers published on indexed journals (JCR), which supports the results obtained.

AGRADECIMIENTOS

Quisiera agredecer en primer lugar a mis directores de tesis, D. Luigi dell´Olio y D. Ángel Ibeas, el haberme dado su apoyo tanto a nivel científico como profesional y personal para la realización de esta tesis.

En segundo lugar me gustaría agradecer también a mis compañeros del Área de Transportes de la Universidad de Cantabria su ayuda y paciencia. Ellos han contribuido decisivamente al trabajo de investigación y de ellos he aprendido y sigo aprendiendo cada día.

Mi agradecimiento también al Ministerio de Ciencia e Innovación por la financiación del proyecto E 21/08, INTERLAND: interacción entre usos del suelo y nuevos modos de transporte sostenibles, que ha hecho posible en gran medida la investigación de la cual ha surgido esta tesis.

Finalmente quiero agradecer a mi familia y amigos por su apoyo, paciencia y comprensión: GRACIAS.

ÍNDICE DE CONTENIDO

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

Capítulo 1
INTRODUCCIÓN

1. INTRODUCCIÓN Y OBJETIVOS

1.1. Motivación

Los sistemas urbanos contemporáneos, tanto en los países desarrollados como en los países en desarrollo, están experimentando fuertes procesos de cambio. Estos procesos llevan asociados problemas derivados que, en un gran número de casos, apartan a las áreas urbanas de las directrices de crecimiento sostenible que, normativamente, la sociedad considera beneficiosos. Entre estos problemas, varios de los más importantes están vinculados a la interrelación entre el subsistema de transporte y el subsistema de usos del suelo de las áreas urbanas. En el sector del transporte fenómenos como la congestión, el elevado y creciente consumo de tiempo de viaje, la accidentalidad y los impactos en el medio ambiente suelen citarse como los retos colectivos más notables (Ortúzar y Willumsen, 2001). En el ámbito de los usos del territorio, la dispersión urbana (urban sprawl), la creciente segregación espacial de las actividades o la mayor cantidad de infraestructuras necesarias para abastecer a los hogares y empresas suelen señalarse también como dinámicas generadoras de diversos problemas (O'Sullivan, 2007). Una dinámica que no sólo está presente en las grandes áreas metropolitanas de EE.UU. sino también de Europa y de España (García Palomares y Gutiérrez Puebla, 2007). Tanto los procesos no deseados derivados del subsistema de transporte como los de los usos del suelo están claramente interrelacionados y de hecho podría decirse que se retroalimentan hasta formar un auténtico círculo vicioso en la dinámica de los sistemas urbanos (véase Fig 1-1). Las dinámicas asociadas a los procesos de crecimiento económico han llevado asociadas generalmente un aumento de la tasa de motorización de las sociedades (Joyce et al., 2007). Esto a su vez ha conducido a un proceso en el que el aumento de la generación de viajes en vehículo privado incrementa los tiempos de viaje, la congestión y los impactos medioambientales. Además la reducción de la accesibilidad y el aumento de la congestión en las áreas urbanas centrales han incentivado la dispersión urbana, tanto en lo que se refiere a la localización de los hogares como de las actividades económicas, con lo que la elección del vehículo privado como modo de transporte preferente se ha reforzado. La planificación urbana y del transporte en caso de orientarse a un tipo de estrategia del tipo "predict and provide" pueden reforzar el circulo vicioso mediante el aumento de las infraestructuras útiles al vehículo privado (nuevas carreteras, espacios de parking y otros), la disminución de la densidad urbana y la debilitación de la competitividad del transporte público.

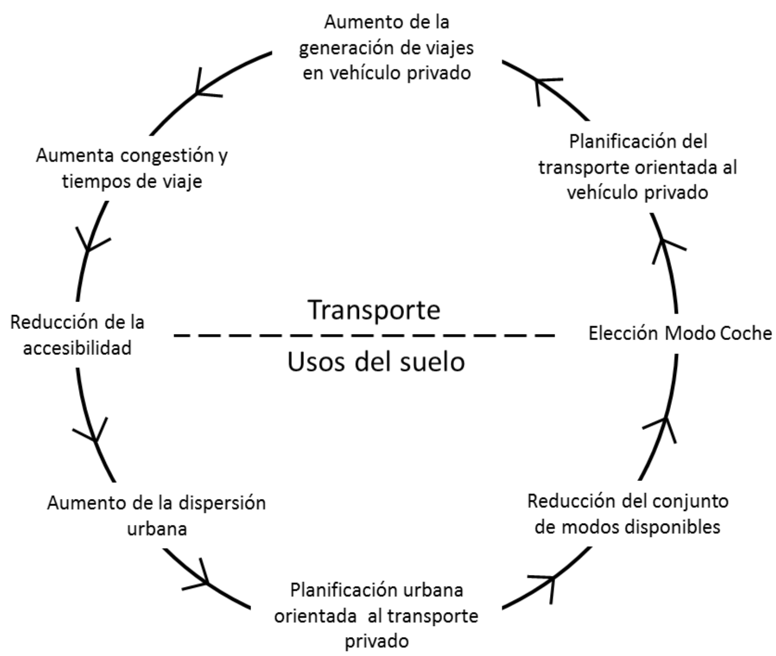

Fig 1-1. Circulo vicioso del transporte y los usos del suelo. Basado en Litman (2010)

Este tipo de equilibrios indeseados son típicos de los problemas de acción colectiva donde la acumulación de decisiones individuales racionales puede conducir a un resultado subóptimo a nivel social. En el sector del transporte, este fenómeno puede darse debido a distintos fallos del mercado como la presencia de fuertes externalidades positivas y negativas y de problemas generalizados de free – riding típicos de los servicios públicos. Esto justifica la intervención a través de la planificación urbana y del transporte como herramientas fundamentales para impedir que este tipo de equilibrios indeseados se consoliden (Wright y Rogers, 2011).

Como guía normativa que sirva de referencia a una planificación y gestión de los sistemas urbanos que pueda romper el círculo vicioso del transporte y los usos del suelo, se ha propuesto tentativamente el concepto de desarrollo urbano y del transporte sostenible. El desarrollo sostenible de estos sistemas implica al menos tres dimensiones fundamentales (May et al., 2003):

- La provisión eficiente de bienes y servicios para todos los habitantes del área urbana.
- La protección del medio ambiente, de la herencia cultural y de los ecosistemas para la presente generación.
- La conservación para las futuras generaciones de al menos el mismo bienestar que el disfrutado por las presentes, incluyendo el derivado del disfrute del medio ambiente natural y cultural.

La importancia del concepto de desarrollo sostenible radica en que permite fijar los problemas y los objetivos generales de la planificación indicando también qué efectos deberían ser evitados. Sin embargo es necesario subrayar que como todo objetivo general el concepto de desarrollo sostenible tiene un alto grado de abstracción por lo que es necesario operacionalizarlo en políticas y objetivos concretos que sean medibles en sus resultados y ayuden a romper el círculo vicioso del transporte y los usos del suelo de forma coordinada.

Se han propuesto diversas políticas tanto por el lado de la oferta como por el lado de la demanda para potenciar el desarrollo sostenible de los sistemas urbanos. Entre las más importantes desde el punto de vista de su impacto en los usos del suelo se encuentran las de inversión en nuevas infraestructuras de transporte. Proyectos de construcción o ampliación de infraestructuras tanto dirigidas principalmente al transporte privado (autopistas) como al transporte público (metro, metro ligero o Bus Rapid Transit) pueden tener impactos significativos en las áreas urbanas tanto en términos de sostenibilidad económica como medioambiental (Fig 1-2).

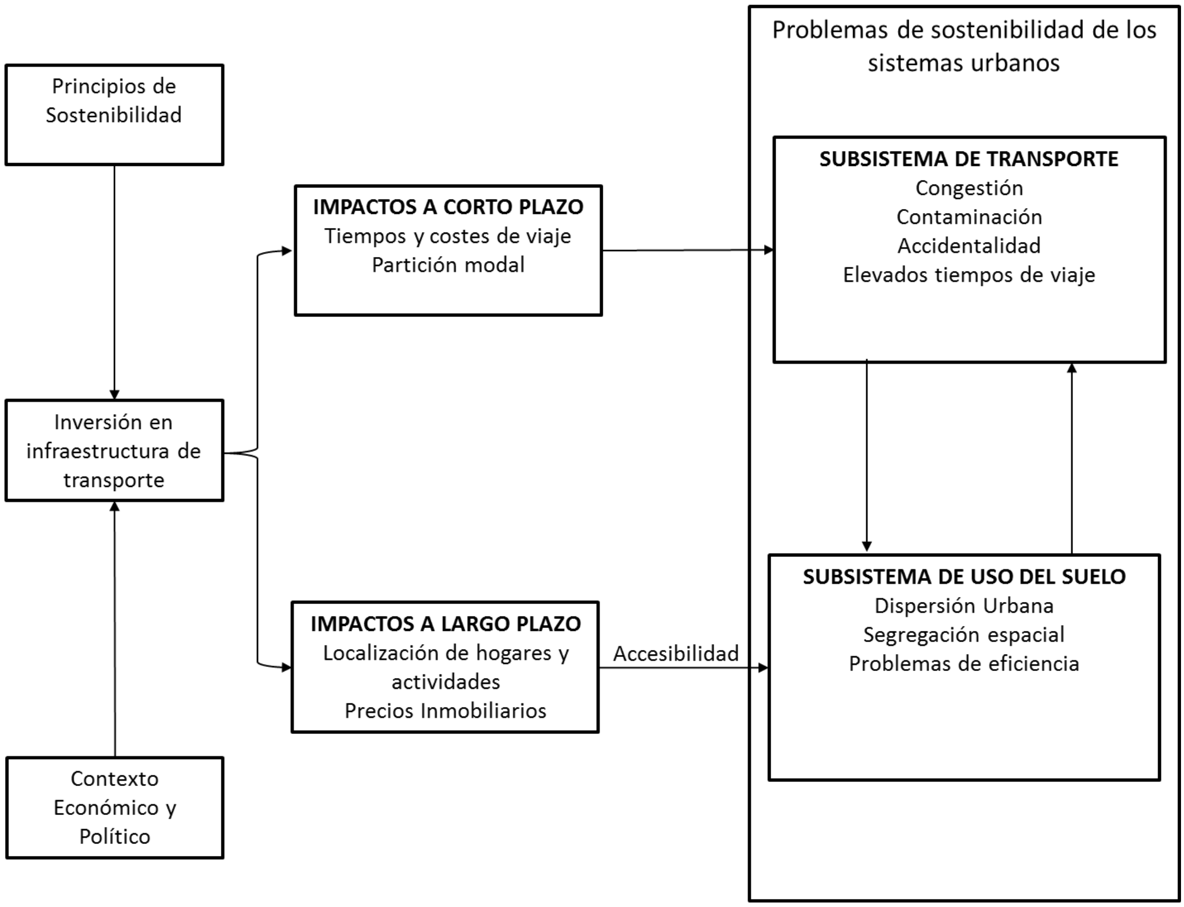

Fig 1-2. Esquema de los impactos de la inversión en infraestructura de transporte en el sistema urbano

A corto plazo la potenciación del transporte a través de la construcción de una nueva infraestructura o de la aplicación de otras políticas pueden provocar impactos positivos en los usuarios a través de la reducción de los costes y tiempos de viaje, lo que puede llevar asociado un cambio en la partición modal.

Adicionalmente, las políticas de transporte pueden generar también otra serie de impactos sobre los usos del suelo a medio y largo plazo (Badoe y Miller, 2000). La variación en las condiciones de accesibilidad de las distintas áreas de un sistema urbano puede afectar significativamente a las decisiones de localización de hogares y empresas así como a los precios del mercado inmobiliario. En el caso de que las inversiones se centren en el transporte público los efectos coordinados pueden incentivar el desarrollo compacto y la densidad urbana ayudando a generar lo que se ha denominado como Smart Growth (Litman, 2010). En cambio la inversión en infraestructura para el vehículo privado puede incentivar un desarrollo más difuso al aumentar la accesibilidad de forma lineal (Handy, 2005).

En todo caso las condiciones y grado en los que se pueden producir estos impactos a medio y largo plazo son aún un asunto de debate siendo además compleja su estimación (Badoe y Miller, 2000; Banister y Berechman, 2000).

Es en el marco de esta problemática donde los modelos econométrico – espaciales pueden ser aplicados tanto al estudio ex – ante como ex – post de los impactos provocados por el transporte especialmente en sus consecuencias a medio y largo plazo. Este tipo de modelos pueden utilizarse para ayudar a responder preguntas del tipo (Martínez, 2000):

- ¿A qué nivel y bajo qué circunstancias los proyectos de transporte provocan el desarrollo urbano?
- ¿Qué proporción de los beneficios generados por los proyectos de transporte son capitalizados por los propietarios inmobiliarios?
- ¿De qué dependen las elecciones de localización de los agentes urbanos y hasta qué punto influyen en éstas las condiciones de accesibilidad de los lugares?
- ¿Cómo va a distribuirse el crecimiento futuro de población y actividades en el conjunto del sistema territorial si se implanta una nueva infraestructura de transporte?

Una vez que un modelo ha sido calibrado frente a un escenario conocido, puede ser usado para hacer predicciones ex –ante sobre el estado futuro del sistema. Por lo tanto el uso de modelos como los presentados en esta tesis puede ser una herramienta de ayuda a la toma de decisiones en el campo de la planificación del transporte difícilmente sustituible una vez comprendida la avanzada y creciente complejidad de los sistemas urbanos.

1.2. Objetivos

El objetivo principal de esta tesis doctoral es el de proponer, especificar, estimar e implementar una serie de modelos econométrico – espaciales que permitan simular los impactos a medio y largo plazo de nuevas infraestructuras y políticas de transporte en el subsistema de usos del suelo.

Dentro de este objetivo general pueden diferenciarse tres grandes grupos de objetivos:

- En un plano teórico mejorar la comprensión de la influencia del sistema de transporte sobre el sistema de usos del suelo. Para ello se definirán las variables consideradas como más relevantes (distintos indicadores de accesibilidad), se establecerá la influencia de las mismas y se identificará su grado de significatividad explicativa. Adicionalmente la especificación de los modelos se realizará teniendo en cuenta las hipótesis que se derivan de la teoría de la economía urbana y del transporte con lo que los resultados obtenidos pueden ayudar a determinar el grado de apoyo empírico de la teoría.
- En un plano metodológico los modelos econométricos utilizados, de regresión múltiple y de elección discreta, se especificaran sin e incorporando la presencia de efectos espaciales. La consideración de estos efectos derivados de la dependencia espacial entre observaciones a través de fenómenos como la difusión y la contigüidad permitirá establecer si los modelos econométricos – espaciales mejoran de forma significativa el ajuste a los datos y por lo tanto su capacidad explicativa y predictiva.
- Por último los modelos especificados buscan ser útiles para un fin aplicado como es la evaluación de proyectos y políticas de transporte. Por eso los distintos modelos se han estimado con datos estadísticos obtenidos del núcleo y el área metropolitana de Santander por lo que podrían ser directamente aplicables a esta zona de estudio.

Estos tres grandes grupos de objetivos son comunes a los tres artículos en los que se basa esta tesis. Todos ellos se han desarrollado con el fin de realizar aportaciones en los planos científico y metodológico además de estar motivados por el interés en facilitar la evaluación de las políticas de transporte.

La aproximación empleada tiene una fuerte componente espacial, en la que conceptos como cercanía, contigüidad o difusión son claves. Se parte de la idea de que tanto en el ámbito de los usos del suelo como en el del transporte, la mayor parte de los problemas deben ser tratados considerando explícitamente la componente espacial de los fenómenos estudiados. Por lo tanto a lo largo de la presente tesis se ha empleado software especializado en el tratamiento de datos espaciales.

De forma más específica los objetivos que se persigue alcanzar en cada uno de los tres estudios que se compilan en esta tesis serán los siguientes:

En el caso del modelo integrado de usos del suelo y transporte del área metropolitana de Santander:

- Analizar los distintos tipos de modelos de interacción entre los usos del suelo y el transporte (modelos LUTI) desarrollados en la literatura.
- Establecer la estructura básica del modelo STIT (System of mathematical models for the simulation of land use and transport interaction) la cual estará compuesta por cuatro submodelos básicos: de transporte, de localización residencial, de localización de actividades económicas y de precios inmobiliarios.
- Especificar y estimar cada uno de los submodelos con datos provenientes del área metropolitana de Santander prestando especial atención a los parámetros relacionados con las condiciones de transporte.
- Determinar la bondad de ajuste del modelo a los datos del año base.

En el caso del modelo que simula la relación entre la elección de lugar de trabajo, la elección de lugar de residencia y el transporte:

- Revisar la bibliografía existente sobre modelización de la localización residencial y su relación con las condiciones de transporte y accesibilidad de las distintas áreas de un sistema urbano.
- Estimar el grado de influencia en la localización residencial de aspectos relacionados con la accesibilidad al empleo mediante modelos de elección discreta.

- Especificar y estimar modelos de elección discreta que consideren efectos espaciales y que por lo tanto evalúen la posible existencia de correlación entre alternativas.
- Determinar si los modelos que consideran dependencia espacial entre alternativas presentan un ajuste significativamente mejor a los datos.
- Estimar los modelos elección discreta considerando y sin considerar relaciones espaciales con datos provenientes del área urbana de Santander.

En el caso del modelo que simula la interacción entre el transporte y los precios inmobiliarios:

- Realizar un estado del arte sobre las principales aproximaciones en la literatura a la hora de modelizar cómo las condiciones de transporte influyen en los precios inmobiliarios.
- Estimar el grado de capitalización de los beneficios derivados de mejoras en la accesibilidad en los bienes inmobiliarios mediante métodos econométricos convencionales (regresión hedónica).
- Evaluar el grado de dependencia espacial entre las observaciones de la muestra y estimar modelos que sean capaces de evitar los sesgos que este fenómeno puede producir en los parámetros estimados por las técnicas de regresión.
- Evaluar si los modelos que consideran dependencia espacial entre observaciones presentan un mejor ajuste a los datos y permiten estimar con un grado mayor de confianza la capitalización de beneficios inmobiliarios derivados del transporte.
- Determinar el grado de influencia de las distintas especificaciones de las relaciones de vecindad entre observaciones en los parámetros estimados.
- Calibrar los distintos modelos con datos provenientes del área metropolitana de Santander.

1.3. Estructura de la tesis

La presente tesis doctoral se divide en cinco capítulos más dos anexos a lo largo de los cuales se profundiza en el conocimiento y la aplicación de distintos modelos econométricos – espaciales al estudio de los impactos del transporte en los usos del suelo. En cada apartado se revisará el estado del arte y se ofrecerá una exposición de los distintos tipos de modelos aplicados: regresión lineal múltiple, elección discreta, regresión lineal múltiple considerando efectos espaciales y elección discreta considerando correlación espacial entre alternativas. Estos métodos econométricos se han aplicado a un área de estudio, el sistema urbano de la ciudad de Santander, donde se analiza el valor de los parámetros de los modelos, su significatividad y la bondad de ajuste de cada modelo en su conjunto a los datos muestrales. La estructura de este trabajo es la siguiente.

En el presente capítulo se describen los objetivos de la tesis y las aportaciones de ésta al estado del arte.

En el capítulo 2 se presenta un modelo LUTI conjunto para toda el área metropolitana de Santander. Este modelo está formado por cuatro submodelos básicos interconectados que se describirán uno a uno: de transporte, de localización residencial, de localización de actividades económicas y de precios inmobiliarios. Una vez estimados los parámetros de los distintos submodelos, se evalúa la bondad de ajuste a los datos muestrales del modelo LUTI en su conjunto.

En el capítulo 3 se profundiza en la interacción entre el transporte y la elección de lugar de residencia de los hogares. En este caso la técnica elegida para realizar la modelización será la elección discreta y se propondrá mejorar los modelos considerando la posible existencia de correlación entre alternativas (las zonas más cercanas y con características similares pueden ser sustitutas más probables que las más alejadas y diferentes). Los modelos que consideran correlación entre alternativas presentan un mayor realismo a la hora de modelizar el proceso de elección y ayudarán también a evaluar en qué grado depende la localización residencial de los costes y accesibilidad al empleo.

En el capítulo 4 se investiga más en detalle la interacción entre el transporte y los precios inmobiliarios en el área metropolitana de cara a obtener un modelo más preciso que el estimado en el apartado 2. Una vez estimadas las ecuaciones de regresión lineal múltiple se evaluará la existencia de relaciones espaciales entre observaciones y se propondrán modelos mejorados que tengan en cuenta la fuerte componente espacial del mercado inmobiliario. Así mismo se cuantificará la capitalización de las mejoras de transporte en el mercado inmobiliario y se ofrecerán conclusiones sobre si estos resultados apoyan o no la teoría de la economía urbana y del transporte.

En el capítulo 5 se recoge las conclusiones finales del trabajo y la investigación realizada y se proponen las futuras líneas de investigación que complementarán los resultados obtenidos hasta el momento.

Finalmente se han añadido dos anexos con información complementaria a la aportada por los tres artículos publicados que forman el cuerpo principal de esta tesis. En el Anexo A se presentan los parámetros estimados en los modelos de generación / atracción de viajes y de elección modal del modelo de movilidad empleado. En el Anexo B se detalla una aplicación práctica realizada mediante el modelo LUTI presentado en el apartado 2 con una simulación de los posibles impactos que tendría la implantación de un metro ligero en el área metropolitana de Santander. Esta información complementaría amplia los resultados obtenidos y subraya las posibilidades de emplear este tipo de modelos de cara a simular los impactos de proyectos concretos y ayudar a una toma de decisiones más apoyada en criterios técnicos.

1.4. Aportaciones

A lo largo de esta tesis se tratará un problema importante para el funcionamiento y la sostenibilidad de los sistemas territoriales: la medición de los impactos del transporte en los usos del suelo así como la relevancia de considerar la componente espacial de estos fenómenos. Ambos aspectos se han considerado como muy relevantes a la hora de evaluar políticas como la inversión en nuevas infraestructuras de transporte que afecten a aspectos como el valor de los bienes inmobiliarios, la localización de los hogares o el desarrollo urbano en general.

Para lograr dicho objetivo, en la presente tesis se proponen avances en los siguientes aspectos:

a) La evolución y mejora de un modelo espacial LUTI ya existente (STIT) para incorporarle un submodelo de precios hedónicos exógeno. El modelo en su conjunto, capaz de simular también la localización de hogares y empresas, se ha implementado mediante código MATLAB y se ha integrado con un modelo de transporte al nivel del estado del arte (ESTRAUS).

b) La calibración del modelo LUTI utilizando datos del Área metropolitana de Santander. Este modelo podría utilizarse inmediatamente para evaluar las consecuencias de la implantación de nuevas infraestructuras de transporte en el área urbana (véase Anexo B).

c) El desarrollo de un modelo de elección residencial que permita medir la importancia de variables relacionadas con el transporte como los costes de viaje casa – trabajo o la accesibilidad a empleos en las decisiones de localización de los hogares.

d) La mejora del modelo de localización residencial desarrollado en el apartado 2 considerando la importancia de incorporar la existencia de correlación espacial entre las distintas alternativas de localización residencial.

e) El desarrollo de un modelo de precios implícitos capaz de medir la importancia de los impactos del transporte en el valor de los bienes inmobiliarios. Para ello se han evaluado varias variables como los costes de viaje al centro urbano, la accesibilidad a empleos o el número de líneas de transporte público presentes en cada zona. Este modelo supone una mejora respecto al desarrollado en el modelo LUTI del apartado 2 al presentar una especificación más realista.

f) Considerar la importancia de incorporar a un modelo de precios implícitos convencional la componente espacial, dado que el mercado inmobiliario se caracteriza por efectos de difusión y autocorrelación espacial importantes.

g) Es importante mencionar que esta tesis es producto del compendio de tres artículos publicados en revistas internacionales (JCR), los cuales validan el aporte de las investigaciones desarrolladas al estado del arte. Las referencias de los tres artículos mencionados son las siguientes:

I. Coppola, P., Ibeas, Á., dell'Olio, L., Cordera, R. (2013) A LUTI Model for the Metropolitan Area of Santander. *Journal of Urban Planning and Development,* 139, 3, 153-165. http://dx.doi.org/10.1061/(ASCE)UP.1943-5444.0000146

II. Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P. (2013) Modelling the spatial interactions between workplace and residential location. *Transportation Research Part A: Policy and Practice,* 49, 110-122. http://dx.doi.org/10.1016/j.tra.2013.01.008

III. Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P., Dominguez, A. (2012) Modelling transport and real-estate values interactions in urban systems. *Journal of Transport Geography,* 24, 370-382. http://dx.doi.org/10.1016/j.jtrangeo.2012.04.012

Capítulo 2
A LUTI MODEL FOR THE METROPOLITAN AREA OF SANTANDER

2. A LUTI MODEL FOR THE METROPOLITAN AREA OF SANTANDER[1]

2.1. Resumen

En este apartado se presenta un modelo conjunto de interacción entre el subsistema de usos del suelo y el subsistema de transporte con el objetivo de simular el equilibrio entre la oferta y la demanda de localización y movilidad en un sistema urbano. El modelo propuesto está basado en STIT (System of mathematical models for the simulation of land – use and transport interaction) desarrollado por Nuzzolo y Coppola (2005). De éste hereda parte de su estructura (teoría de la utilidad aleatoria) aunque incorpora como características novedosas la conexión con un modelo de transporte complejo como es ESTRAUS y un submodelo de precios implícitos. Este modelo es por lo tanto capaz de estimar la localización de población, actividades económicas y

[1] Coppola, P., Ibeas, Á., dell'Olio, L., Cordera, R. (2013) A LUTI Model for the Metropolitan Area of Santander. *Journal of Urban Planning and Development*, 139, 3, 153-165. http://dx.doi.org/10.1061/(ASCE)UP.1943-5444.0000146

precios inmobiliarios medios en las diferentes zonas en que puede dividirse un sistema urbano ante cambios en el sistema de transporte.

El modelo puede caracterizarse como un simulador a caballo entre los de primera y segunda generación. El sistema presenta una estructura similar a la del modelo Lowry (1964), combinando la teoría de la base económica como factor explicativo de la dinámica entre empleo y población, con la teoría de la interacción espacial, ya que considera los patrones de localización como dependientes de los costes de transporte entre zonas. Además incorpora la teoría de la utilidad aleatoria como modelo de predicción de las elecciones de los agentes (empresas y hogares).

En cuanto a la interacción entre los distintos submodelos, el sistema LUTI desarrollado puede caracterizarse como un modelo de equilibrio estático. Así cualquier cambio que se produzca en un elemento del sistema territorial (e.g. el subsistema de transportes) conduce a una nueva solución de equilibrio que expresa el nuevo estado del sistema.

El modelo en conjunto presenta cuatro grandes submodelos integrados:

- Un modelo de transporte el cual dado un patrón de localización de residentes y actividades simula las fases de generación – atracción de viajes, distribución de viajes, elección modal y asignación a la red.
- Un modelo de localización residencial el cual dado un tiempo de viaje entre zonas, un patrón de localización de actividades y una oferta residencial, simula la localización de los trabajadores y residentes del área de estudio desagregados en dos niveles de ingresos: superiores e inferiores a los 2500 €.
- Un modelo de localización de actividades económicas, el cual, dados la accesibilidad de cada zona y el patrón de localización residencial, simula la distribución de la localización de las actividades desagregadas en dos tipos: comercial minorista y de servicios.
- Un modelo de precios implícitos el cual dadas las características estructurales de las viviendas, las características del entorno y la oferta de transporte de cada zona, simula los valores inmobiliarios medios.

El modelo LUTI fue aplicado al área metropolitana de Santander (España) para testear su bondad de ajuste a los datos del año base y por lo tanto su capacidad para predecir los impactos de introducir nuevas políticas de transporte como por ejemplo una nueva infraestructura de transporte público. En el caso de Santander, el modelo de precios inmobiliarios así como el modelo de localización de población y actividades mostraron ser sensibles a los cambios en las condiciones de accesibilidad y transporte de cada zona y el modelo como un conjunto mostró una buena bondad de ajuste a los datos.

2.2. Introduction and objectives

Classic urban economic theory states that the accessibility conditions of different places are a key factor in explaining the location of the population and the economic activities in an urban system. This idea began in early studies on agricultural land use patterns by Von Thünen (1826), which were later applied to urban spaces by Alonso (1964), Muth (1969) and Mills (1972b). The theory is based on the existence of a trade – off between accessibility and occupied space which needs to be resolved by the different urban agents. Thus the locations with better access to the Central Business District (CBD) will have, ceteris paribus, higher population and activity densities, as well as higher land rents per unit area.

One of the consequences of this theory is that improvements in accessibility may increase the attraction of an area for people and economic activities as well as increase real estate value. However, there are other factors which contribute to the complexity of urban systems. Fujita (1989), for example, provides a comprehensive exposition of urban location theory and considers the existence of a third essential factor which is the presence of positive and negative externalities caused by certain environmental conditions such as public goods, high population densities and traffic congestion among others.

Land use and transport interaction models (LUTI models) have been designed to simulate these complex relationships, starting from assumptions similar to those of urban economic theory although adapting them to the conditioning factors of real planning situations. These models have traditionally been used to simulate the possible effects of introducing new policies and projects into existing urban systems and, especially, those related to transport (Foot, 1981). LUTI models have therefore been used as a complementary tool to transport models which consider a fixed pattern of mobility demand location and are, therefore, more adapted for making shorter term predictions (Ortúzar and Willumsen, 2001).

The research presented here develops an integrated model to evaluate the interaction between land use and transport which includes the structure developed in the STIT (System of mathematical models for the simulation of land – use and transport interactions) model (Nuzzolo and Coppola, 2005) and an exogenous real estate price model based on hedonic pricing theory. STIT is a LUTI model developed by Nuzzolo and Coppola (2005, 2007) with a similar structure to the Lowry model, combining economic base theory as an explanatory factor for the dynamics between employment and population with spatial interaction theory, as it considers location patterns as dependent on the cost of transport between zones. The model also incorporates random utility theory as a choice prediction model for the agents involved (companies and households) which provides STIT with a stronger base as it can be supported by microeconomic theory. The solution system for STIT can be characterized as that of a static equilibrium model. Any change occurring to an element in the territorial system (e.g. the transport sub-system) leads to a new equilibrium solution expressed by the new state of the system. The comparative equilibrium approach is currently the most viable when performing practical applications given the greater simplicity of these types of models during the calibration and implementation phases (Nuzzolo and Coppola, 2005).

The developed LUTI model is made up of four interrelated submodels: a residential location submodel, an economic activities location submodel, a real estate price predicting submodel and a transport submodel. The four submodels interact through various information flows, mainly the indicators of accessibility and journey times between areas, in order to simulate the equilibrium of the urban system being studied.

The use of random utility theory in several of these submodels provides a solid micro-economic framework based on the maximization of utility by the agents involved compared with the analogies of gravity models based on aggregated data (Lowry model (Lowry, 1964), DRAM/EMPAL (Putman, 1996)). The current STIT model is more complete as it is able to endogenously simulate average real estate values in the different zones of the study area. Finally, the land use model is completed by its interaction with a modern transport model, ESTRAUS, which is able to simulate the relationships between mobility supply and demand as a problem of simultaneous equilibrium (De Cea and Fernandez, 1993; De Cea et al., 2003). This feature allows the model to simulate the workings of the urban system in a consistent way and at the same time consider the presence of network congestion, which previous LUTI models failed to do and sometimes tended to use transport models from outside the established literature (Wegener, 2004).

The integrated system of models will be applied to a real case study, the metropolitan area of Santander (Spain), although it can be applied to any other study area. This practical application will then be used to check the model's goodness of fit with the aggregated and disaggregated data collected for the area.

The presentation of the research will be organised in the following way. A state of the art review will be presented about the field of LUTI modelling. This will be followed by an explanation on how the integrated system works, providing details on its overall structure and the composition of the different submodels. The study area being analysed will then be introduced along with the data being used and the estimation of the parameters for each submodel. The model's goodness of fit to the base year data will be explained and this will be followed by a series of conclusions.

2.3. Bibliographic review

LUTI mathematical simulation models combine theory, data and algorithms to provide an abstract representation of the interaction between the two main components of urban areas: the transport and land use subsystems (Torrens, 2000). The first important theoretical contribution which related transport and land use was made by Von Thünen (1826) in the first half of the 19th century. Von Thünen related the agricultural land use system of an area with the costs of transporting goods to the central market. These contributions were later applied to urban areas by Alonso (1964) and in agreement with the Von Thünen theory, were based on the existence of a single Central Business District (CBD) where all employment would be concentrated. Households and companies would therefore decide on their locations based on budgetary restrictions and their preferences when making the trade – off between accessibility and occupied space. Fujita (1989) provides an excellent comprehensive exposition of urban economic theory by considering the contributions made by a great many authors based around three basic location factors: accessibility, occupied space and environmental conditions.

However, the application of these theories to the world of urban planning has not been without its difficulties, which has led to the development of a wide range of apt models for carrying out planning exercises but not necessarily coherent with economic theory. This phenomenon is what Harris (1985) referred to as the tensions between economic theory and the practice of simulation and planning.

Many models are currently available for carrying out simulation and planning exercises and several authors have proposed classifications which group together different models according to diverse criteria. Wegener (2004) classified more than twenty models developed in the literature using nine main criteria: comprehensibility, structure, theoretical basis, techniques used, dynamics, data required, calibration, operationality and applicability. Waddell and Ulfarsson (2004) made a classification based on the theoretical approaches that have appeared in the field of LUTI modelling over the last 50 years, whereas Iacono et al. (2008) classified the models found in the literature according to the historical development of the great theoretical paradigms of modelling. The typology presented below classifies the models in accordance with their basic theoretical nucleus and the chronological generation they belong to. Three generations and five basic types of models have been differentiated (see also Fig 2-1).

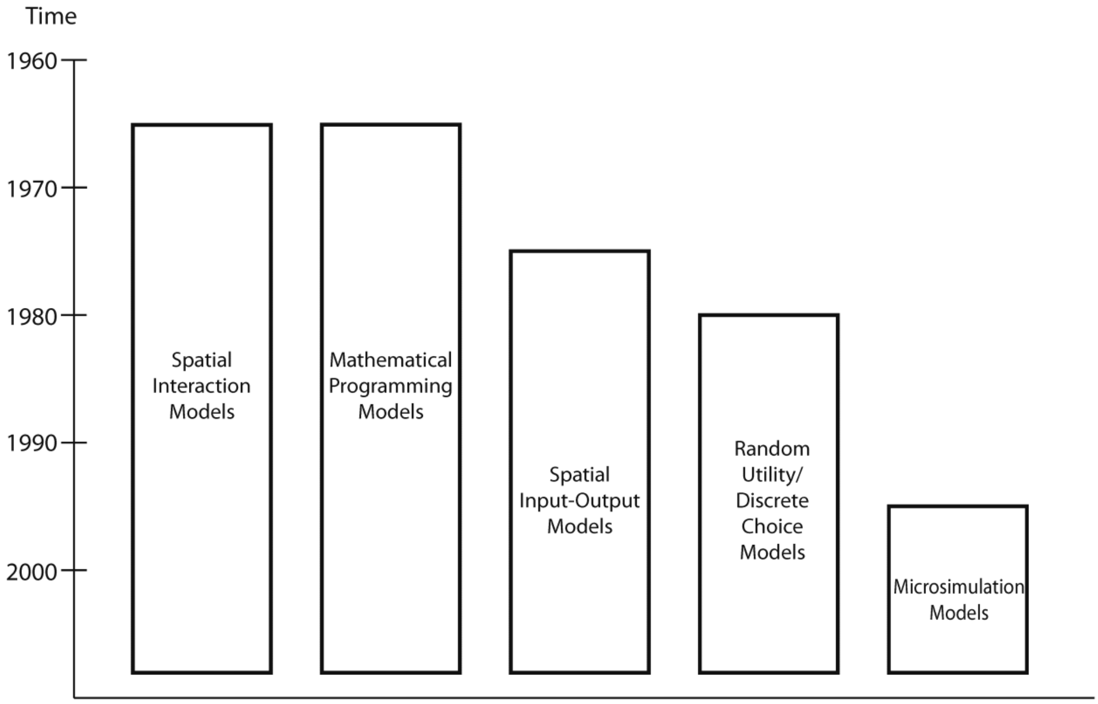

Fig 2-1. Chronological development of LUTI models. Own elaboration based on data from Iacono et al. (2008)

i. First generation models: the models that appeared during the 1960s and 1970s. They can be divided into three main types according to their theoretical basis for performing simulations.
    a. Spatial and gravity models: based on the theory of spatial interaction or on the statistical mechanics generalisation performed by Wilson (1970). The classic example is the interaction model developed by Lowry (1964) which was also supported by economic base theory for simulating the locations of population and economic activities (Andrews, 1953).
    b. Mathematical programming models: based on optimisation techniques. This type of model is based on a simulation of agent behaviour through the minimisation or maximisation of a certain objective. The classic model of this type was developed by Herbert and Stevens (1960) and simulated the operation of the residential location market following the theory of Alonso through maximisation of aggregated rents. Another example of this type of model is TOPAZ (Technique for Optimum Placement of Activities into

Zones) which determined the locations of activities as a function of the minimisation of transport costs and urban development (Brotchie et al., 1980).

c. Models based on INPUT/OUTPUT matrices: this type of model simulates the urban or regional economy using the technique of input/output matrices developed from the work of Leontief (1966). An example of this type of model is MEPLAN (Echenique, 1994, 2011).

ii. Second generation models appearing in the 1980s and 1990s. This type of model is based on random utility theory developed from the work of McFadden (1974). This generic type can be further differentiated into simulation of land markets using random utility theory based on the work of Anas (1982). The land use model of Santiago (MUSSA) developed by Martinez (1997) is an example of a second generation model.

iii. Third generation models appearing more recently around the second half of the 1990s. These are highly disaggregate models known in some cases as micro simulation models (Iacono et al., 2008). They have a dynamic nature meaning they don't reach complete market equilibrium as the solution to their simulations. One of the better known models of this type, very widely applied, is URBANSIM developed by Waddell and collaborators at the University of Washington (Waddell et al., 2007b).

It is important to note that research is moving forward with the three generations of models and none of them has successfully managed to replace any of the others. Random utility theory is currently the most commonly used paradigm in location choice modelling for different urban agents. This theory has, to a large extent, substituted the location models based on spatial interaction theory which offered a smaller behavioural base even though various researchers have shown that both approaches may provide similar results under certain assumptions (Anas, 1983).

In the field of simulating the transport subsystem, LUTI models have sometimes been criticised because they incorporate methods which are somewhat distant from the state of the art. Many LUTI models are still based around the classic four stage sequential approach leading some authors to recommend the use of more modern models which could be either endogenous or exogenous to the rest of the LUTI simulator (Wegener, 2004).

The ability to predict changes in real estate values is an aspect of LUTI modelling which has received increased interest for evaluating and financing planning policies and public projects. However, real estate prices are not predicted by all the models and when they are, the predictions are not all based on the same procedures. The techniques which agree most with established economic theory are those which predict the prices of the different properties based on a system of market equilibrium. The MUSSA model developed by Martinez (1992) incorporates a bid – choice system of modelling similar to that proposed in the theory of Alonso with the inclusion of a stochastic component derived from random utility theory. Other models have applied hedonic regression techniques based on the market simulation of heterogeneous goods formalised after the work of Rosen (1974). This type of model uses multiple regression to estimate variations in real estate prices and represents the envelope function of the functions of supply and valuations of producers and consumers. The estimated prices depend on a series of structural attributes related to the properties themselves as well as to other environmental and location characteristics. This class of models are able to exogenously estimate prices, in other words, by only taking into account property characteristics specified using external data to the model, or endogenously, by simulating a supply and demand mechanism derived from measures of an area's location capacities and the number of agents willing to locate there (Coppola and Nuzzolo, 2011).

2.4. Description of the integrated models system

2.4.1. General structure

The main purpose of the LUTI model presented in this article is to estimate variations in the location of population, employment, and property prices in an urban system when faced with changes which are mainly associated to the transport system, such as the introduction of new transport modes or new travel demand management policies.

The system developed here can be framed between the first and second generation models in the classification presented in the previous section. It combines traits derived from the spatial interaction models, especially the Lowry type, including the distinction between basic and non-basic activities, with the inclusion of random utility theory in the simulation of household and economic activities location decisions. The model is also able to simulate average zonal real estate prices by including an endogenous hedonic pricing submodel. Finally, the integrated system incorporates a connection with a modern transport model, ESTRAUS, which simulates the relationships between the supply and demand of transport with simultaneous equilibrium solution (De Cea et al., 2003). This allows the model to consistently simulate urban systems with the presence of congestion.

The integrated system is made up of four interrelated submodels. The interaction between the submodels is solved through an equilibrium solution, meaning the model could be characterised as a comparative equilibrium model. This allows for any change occurring in the territorial system to lead to a new equilibrium solution representing the new state of the system. The four submodels making up the main structure of the LUTI model are as follows:

a. A transport model which, given a location pattern for residents and activities, simulates the simultaneous equilibrium between supply and demand in the transport system.
b. A residential location model which, given a journey time between zones, an activity location pattern, a set of real estate prices and a residential supply, simulates the location of workers and residents in the study area disaggregated by income classes.
c. An economic activities location model which, given each zone’s accessibility and the residential location pattern, simulates the distribution of these activities disaggregated by economic sector.
d. An implicit prices model which, given the structural characteristics of the properties, the environmental characteristics, the demand/supply of location and the transport conditions of each zone, simulates the property prices.

Depending on the theoretical hypothesis being proposed, multiple interactions and equilibrium problems can occur between the different submodels. Firstly, the transport demand in the different zones making up the study area depends on the location of residents and economic activities. This demand involves a series of journey and mode choices which, depending on the available transport supply (network capacity and public transport services), generate a cost matrix between the zones (expressed in either journey times or generalised transport costs) which, in turn, influence the accessibility of each zone.

The implicit prices model calculates the average property prices in each zone as a function of the supply and demand for locating in each zone as well as the structural and environmental characteristics of the area.

The residential location model starts from the hypothesis that workers locate depending on different zonal characteristics, including the distance from work places, journey times or the costs involved, obtained from the interaction between the transport and economic activities location submodels. Another important variable involved in residential location are the property prices of each zone, obtained using the hedonic pricing submodel.

The activity location model works in a similar way to the residential location model, considering the utility of each zone as a function of different variables. Among these variables is the accessibility of the population to each zone, depending on the travel costs between areas derived from the transport model, and the population of each zone derived from the residential location model. The flow diagram of the four interrelated submodels can be seen in Fig 2-2.

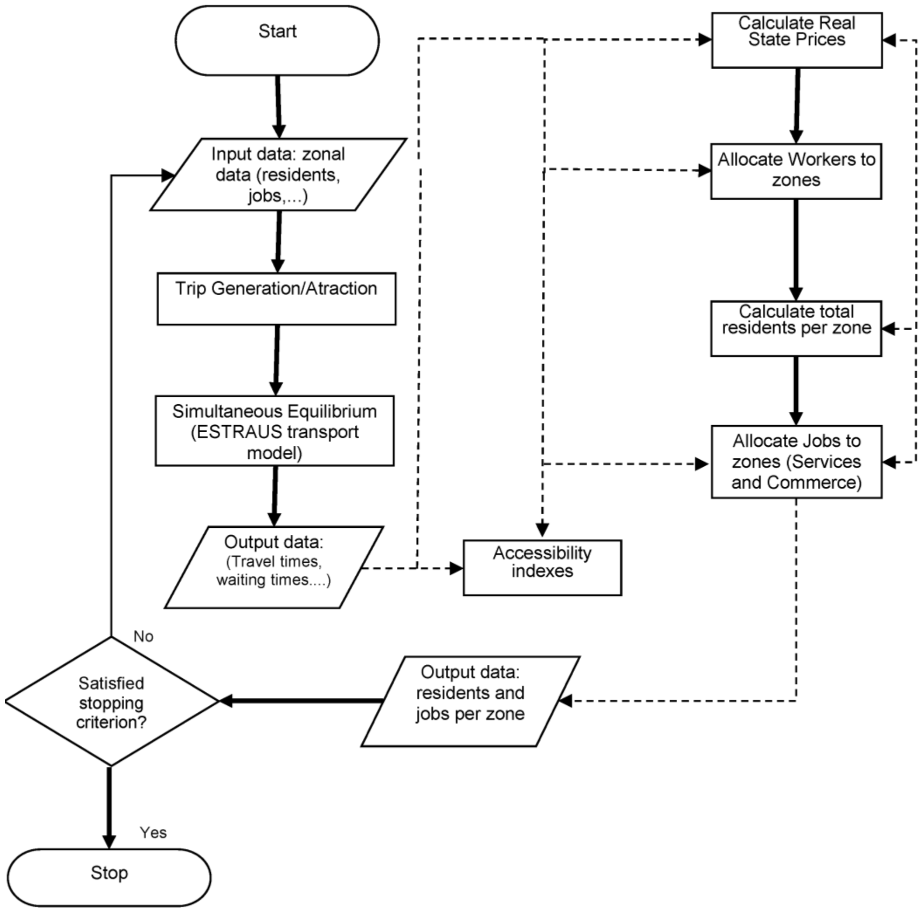

Fig 2-2. Flow diagram of the LUTI model

Based on the diagram shown in Fig 2-2, the residential and economic activity location models spatially distribute households and companies. These activities (residence, production, consumerism etc.) attract and generate journeys throughout the different places located in the territorial system which, in turn, are used as input for the transport model during the distribution, modal split and trip assignment phases. In traditional transport models the trip generation/attraction phase is normally considered to be exogenous or could even be calculated from current or planned land use. The use of the LUTI model which envelopes the household and company location decisions allows for more consistent estimations to be made on the dynamics of the urban system in reaction to different medium and long term planning policies.

The transport model used includes the modal choice phase which is estimated using a logit model. The simulation performed here was based on a choice group of 5 modes: private transport (car or motorbike), urban bus, light rail, inter-urban bus and a combined mode of light rail and urban bus.

Finally, the model as a whole starts from a series of theoretical hypotheses which contribute to the simplification of reality by excluding from the modelling process certain aspects considered secondary or which require data which are difficult to obtain. These hypotheses can be grouped into four basic suppositions:

1. The study area to be modelled is considered “closed” i.e. the jobs are occupied by the internal demand. From the modelling point of view, closing the territorial system avoids problems associated with worker immigration/emigration and commuter movements outside the study area. On the other hand, this assumption might lead to neglect the impact long distance commuters (those employed in the study area and residing outside of it, e.g. in other municipalities) have on the transport system being studied. To limit the drawbacks of such an assumption the study area should be metropolitan in nature and not exclusively local or municipal. In other words, it should include all the residential areas and municipalities where people working in the study area reside. For instance, it has been verified in the application presented in the following session that the labour force within the metropolitan area of Santander covers about the 97 % of the jobs in the study area. This makes the assumption of “closed” study area largely acceptable, since the percentage of long distance commuter trips over the total home-to-work demand is negligible. For the sake of completion, it must be said, however, that there could be cases in which this assumption does not hold, such as the case of some US and European metropolis (served by High Speed Railway and/or big airports) where people daily commute between cities at distances of hundreds of kilometres (i.e. the so called phenomenon of “super commuting”). In these cases, since it is not possible to enlarge the study area to such a great extent, an economic model simulating the migrations between different cities on a national or international scale should be adopted (see for instance the regional model within the DELTA package by Simmonds, 2001).
2. The real estate supply is considered fixed in the study area, i.e. the proposed model does not present any property supply submodel.

3. The location of basic sector activity is exogenous to the model following the economic base theory implemented in the Lowry model (Lowry, 1964). Basic sector economic activities are assumed to locate without taking into account the distribution of the population in the study area because they do not directly depend on the internal demand. The distribution of this type of employment can therefore be considered as fixed or change only because of external considerations to the working of the model.
4. There are no explicit capacity constraints on the location of economic activities and population. Although in principle this assumption allows the number of people living in a zone to exceed the housing capacity of that zone, this is unlikely in a practical application if the parameter of the housing demand and supply ratio (i.e. the attribute DS(o) in the real estate price model) is correctly calibrated. In fact, when the demand for housing in a given zone approaches capacity, the ratio between residential demand and supply ratio, i.e. DS(o), increases, leading the price of the houses in zone “o” to such a high value that it makes the probability of additional demand in the zone equal to zero. This is the same dynamic which arises in the traffic assignment equilibrium on road networks, when no explicit link capacity constraints are included in the model, but the link cost is a function of the flow-capacity ratio (e.g. through a BPR function). In other words, when the parameter of the attribute DS is properly calibrated the capacity constraint is respected, unless the total housing demand in the study area exceeds the total housing supply. On the other hand, from the mathematical point of view, the assumption allows us to prove there is a unique solution to the equilibrium problem of the spatial distribution of residents and economic activities (see Nuzzolo and Coppola, 2005).

2.4.2. Transport model

The transport model is briefly summarised because it does not form part of the main goal of this article which is more focussed on the simulation of land use, however, as mentioned earlier, the integrated system incorporates the ESTRAUS transport model (De Cea et al., 2003).

After receiving the input data on population and activity patterns, the transport model can simulate the generation, distribution, modal split and assignment of journeys. The transport model can, in turn, provide input data for the land use models based on travel costs by transport mode between zones and average waiting times in each area, among others. ESTRAUS uses a deterministic equilibrium model based on Wardrop's first principle for the route choice model and a hierarchical logit model for the other travel choices. In addition, ESTRAUS allows simultaneous equilibrium between transport supply and transport demand which refers to consistency in the levels of services and transport flows in each step of the model.

As the equilibrium between transport supply and demand is provided by ESTRAUS the accessibility indicators can be calculated, representing the fundamental link between the transport and land-use subsystems. From a theoretical point of view, accessibility has been defined as the ease by which any activity can be reached from a given location using a transport system (Geurs and van Wee, 2004). Handy and Niemeier (1997) classify the measures of accessibility into three large groups:

- Measures based on accumulated opportunities: these are taken from two points of view depending on whether the opportunities are measured from either the demand or supply centres. In the first instance they quantify the possibility of reaching a goal within a given area measured in journey time or distance (potential supply). In the latter, they measure the population size within the spatial reach of the equipment or service (potential demand).
- Measures based on gravity models: these refer to the capacity of a place to reach certain opportunities, generally weighted by distances or journey times. This can also be expressed as a place's capacity to be reached from other places generally weighting the population of the nearby zones by the journey times or distances they have to make.
- Measures based on random utility theory: these have a more disaggregate nature and evaluate the desirability of a set of destinations i.e. the denominator of the multinomial logit model or logsum.

In spite of the many theoretical studies carried out and the wide range of available indicators, in the applied field and especially in Europe, there has been a lack of applications which have evaluated the impact of changing accessibility conditions in urban systems. Because of its greater ease in calculation and interpretation, a gravity indicator was chosen to measure accessibility in each of the zones in the metropolitan area. The accessibility of a zone can be split into two large sub types. They are denominated as active accessibility or the potential reachable opportunities from an area, and passive accessibility or the potential of consumers able to reach the activities taking place in an area (Cascetta, 2009).

The active accessibility of a determined zone to the employment opportunities in the rest of the zones can be calculated using the following expression:

$$Acc(o) = \sum_i \left[\exp\left(\alpha_2 \cdot Cost\left(o, d_i\right)\right) \cdot jobs\left(d_i\right)^{\alpha_1} \right] \tag{2.1}$$

Where:

Cost is a measure of travel cost by transport mode between origin *o* and destination d_i

Jobs (d_i) are the number of jobs in destination zone d_i

α_1 and α_2 are the parameters to be estimated

Passive accessibility can be represented by:

$$Acc(d) = \sum_i \left[\exp\left(\beta_2 \cdot Cost\left(o_i, d\right)\right) \cdot res\left(o_i\right)^{\beta_1} \right] \tag{2.2}$$

Where:

Cost is a measure of travel cost by mode of transport between origin *i* and destination d

res (o_i) are the number of residents present in origin zone o_i

β_1 β_2 are the parameters to be estimated

Various measures of cost can be used, including travel times and expressions of generalised cost. (2.1) and (2.2) can also be calibrated by linearizing the expressions through a logarithmic transformation. As the transport model is multimodal it considers public transport and the relative variables separated by mode.

2.4.3. Residential location model

The main aim of the residential location model is to calculate the number of residents that live in each of the zones in the study area. A similar model to the one presented in this article has been proposed by Hsu and Guo (2006). The model is based on a hypothesis derived from random utility theory that individuals choose locations which maximise their utility. As the modeller cannot know, in an absolute way, how individuals value different locations, a probabilistic discrete choice model is postulated in which the error terms are assumed to be independently and identically distributed Gumbel. The consumers of residential spaces value different zones as a function of their environmental and location attributes relative to their places of work, among other factors. Using these assumptions, the probability that a worker of type *i* (income class) chooses zone *o* as their place of residence conditioned to working in *d* is given by the following well known logit type formulation:

$$P^i_{res-cond}(o|d) = \exp\left[V^i(o|d)\right] \Big/ \sum_o \exp\left[V^i(o|d)\right] \tag{2.3}$$

Where:

$P^i_{res-cond}(o|d)$ is the probability of a type *i* worker choosing to live in zone *o* conditional on working in zone *d*

$V^i(o|d)$ is the systematic utility given to type *i* worker choosing to live in zone *o* conditional on working in zone *d*

Under assumption 1, previously mentioned, the study area is a closed labour market. Therefore, the supply of employment is taken up by the internal demand of employees. From this hypothesis it can be deduced that the number of workers *w* of type *i* who locate in zone *o* is equal to:

$$w^i(o) = \sum_d P^i_{res-cond}(o|d) \cdot Emp^i(d) \tag{2.4}$$

Where $Emp^i(d)$ represents the total number jobs in zone *d* available to class *i* workers.

Finally, knowing the number of workers in each zone of the study area allows us to find the total number of residents in the zone by using the coefficient *k(o)* representing the ratio between residents and workers in each zone:

$$res(o) = k(o) \cdot \sum_i \sum_d P^i_{res-cond}(o|d) \cdot Emp^i(d) \tag{2.5}$$

Expression (2.5) can now be used to calculate the number of residents who locate in each zone within the proposed hypotheses. The number of residents depends on the number of jobs in zones *d* which is consistent according to economic base theory in assuming that any increase in employment has multiplying effects on the population of the urban system.

2.4.4. Economic activities location model

The economic activities location model can be used to determine the distribution of employment in the different zones of the study area disaggregated into sectors. The basic expression of the model is:

$$Emp_a(d) = P_a(d) \cdot EMP_a \tag{2.6}$$

Where:

$Emp_a(d)$ is the number of jobs located in zone *d* belonging to sector *a*

$P_a(d)$ is the probability that a type *a* job is located in zone *d*

EMP_a is the total number of type *a* jobs in the study area

In a similar way to the residential location model, random utility theory is used to simulate the activity location decisions by modelling the decisions as discrete choices. The private agents are assumed to assign a utility to each zone and choose the one that returns the maximum utility. The utility is again assumed to be made up of two parts: systematic and random. If these residuals are independently and identically distributed Gumble, the probability $P_a(d)$ of locating the activity in zone *d* is given by:

$$P_a(d) = \exp[V_a(d)] \Big/ \sum_d \exp[V_a(d)] \tag{2.7}$$

The model is able to differentiate between activities. In urban modelling research these are generally classified into four categories (Nuzzolo and Coppola, 2005):

- Basic sector activities dependent on exporting outside the system and therefore with location not directly tied to the distribution of internal demand, i.e. to the location of households or other economic activities.
- Activities aimed at the internal demand such as retail and non-advanced services which depend on the location of the demand.
- Representative activities such as those whose location depends on particularly attractive zonal characteristics for reasons of prestige or centrality.
- Activities with low spatial efficiency that need large areas of land to function correctly such as car dealerships, industrial complexes, etc.

The economic activity location model used in the integrated modelling system only considers those activities where the location depends on the population distribution, in other words, activities aimed at the internal demand considered by economic base theory to belong to the non-basic sector. The classification of activities as belonging to the basic or non-basic sector has received various criticisms (Camagni, 2005). Even so it may be considered a valid approximation in models which do not seek long-term forecasting of the impact of an expansion of the demand in the base sector in the amount of population and employment of the study area (as is the present case).

Because residential location depends on the location of economic activities and vice versa, an equilibrium problem for both is presented. Mathematically it can be formularised in the following way:

$$\begin{cases} R^i = R\left[\sum_i A^i\right] & \forall i \\ A^i = A\left[\sum_i R^i\right] & \forall i \end{cases} \tag{2.8}$$

Where:

R^i is a vector [n_zones x 1] of type *i* residents

A^i is a vector [n_zones x 1] of total type *i* jobs

The solution to this equilibrium problem can be treated as a fixed point problem with its solution given by the vectors R^{i*} and A^{i*}. The existence of this equilibrium solution comes from the fulfilment of the conditions imposed by the Brouwer's theorem (Cascetta, 2009). The uniqueness of the solution can be checked if the functions R[·] and A[·] are strictly monotonous and the probabilistic location model is additive (as in the case of the logit model).

2.4.5. Real estate price prediction model

The property price simulation model is based on the theory of hedonic regression formalised from the new consumer theory of Lancaster (1966) and, above all, based on the theory of market functioning for heterogeneous goods developed by Rosen (1974). This methodology has been shown to give good results when simulating variations in real estate prices caused by changes in transport or environment conditions (see for example Smith and Gihring (2006) in the case of transport and Nelson et al. (1992) and Zeiss (1990) in the case of environment). In addition, the hedonic model has been specified as endogenous, i.e. zonal housing prices are also affected by the location choices of agents.

Because the LUTI model as a whole simulates the urban system as a discrete space, the hedonic regression needs to be estimated based on the aggregation of the variables in the database. The general structure of the model is as follows:

$$\ln P_j(o) = \beta_0 + \beta_1 X_{j1}(o) + \cdots + \beta_n X_{jn}(o) + \gamma_o \left(\frac{D_j(o)}{S_j(o)} \right) + \varepsilon_j \qquad (2.9)$$

Where:

$P_j(o)$ is the average price of the group of properties type j in zone o

$X_{jn}(o)$ is an attribute of the properties type j in the zone o or of its environmental characteristics

$D_j(o)$ is the real estate demand for properties type j in the zone o

$S_j(o)$ is the real estate supply of properties type j in the zone o

β_1, β_2, β_n, γ_0 are the parameters to be estimated

ε_j is an identically distributed independent error term between observations

The specification of the hedonic model shows a semi-logarithmic functional form i.e. the value of each estimated parameter represents the semi-elasticity. This functional form is one of the more commonly used in applied studies because the value of the coefficient is easier to interpret and it has the added advantage of reducing problems associated with heterocedasticity (Malpezzi, 2008). Nevertheless, the specialised literature does not provide any consensus about which functional form is the most appropriate because theoretically there are no restrictions (Cropper et al., 1988; Stephen, 1999).

2.5. Application of the model to the metropolitan area of Santander

2.5.1. Introduction to the area and the data used

The integrated system of models developed in this work will be applied to the metropolitan area of the city of Santander located in the north of Spain. Santander is a medium sized town and the administrative capital of the region of Cantabria. The city itself currently has a population of 182,700 inhabitants which rises to 280,000 for the larger metropolitan area. The overall area provides around 100,000 jobs, of which 67,000 correspond to the capital. The population and the employment are highly concentrated in the city of Santander and along the axis formed by other important urban centres within the metropolitan area such as Astillero (10,020 hab.), Muriedas (11,279 hab.) and Maliaño (5,272 hab.) (see Fig 2-3). Santander is connected to the other relevant centres by various transport networks and transport services. These networks are mainly made up of urban and interurban roads with their associated public transport services as well as the suburban railway network which connects the most important centres of population within the study area.

The area is administered by 9 different municipal boroughs. The zoning used in this study has divided the metropolitan area into 42 zones (see Fig 2-3). The different submodels will be calibrated with data from 2008 which is taken to be the base year for this research.

The data used in estimating the parameters of the different submodels came from three main sources. The first source was provided by official statistics. The Spanish Institute of Statistics publishes the Population and Household Census and annual municipal registers which provide data on the location and characteristics of the population and households per census district. This data has been used for estimating the residential location models. The Institute of Regional Statistics publishes the Regional Company Directory for 2008 which contains information on the location and characteristics of individual economic activities (number of employees, classification by sector, etc). This data has been used in estimating the location models for economic activities. The data for estimating the real estate pricing model has come from two real estate sources given that the required detail of disaggregated data does not exist in the study area. Finally, the third data source used consisted of a transport survey designed by the authors and others which provided information on the characteristics of the surveyed households and the mobility of each household member.

The data required to run the transport model is imported through a series of text files which can be edited before each simulation cycle. In the present case the zonal trip generation/attraction archive is modified at each iteration depending on the results provided by residential and company location submodels. These submodels were programmed using MATLAB code. An initial procedure imports the results of the transport costs between the zones and calculates the accessibility indicators. A second phase then calculates the real estate price indices as well as the distribution of households and companies through the location models. Finally, this data is once again input into the ESTRAUS software which calculates the trip generation and attraction along with the rest of the stages involved in the transport model.

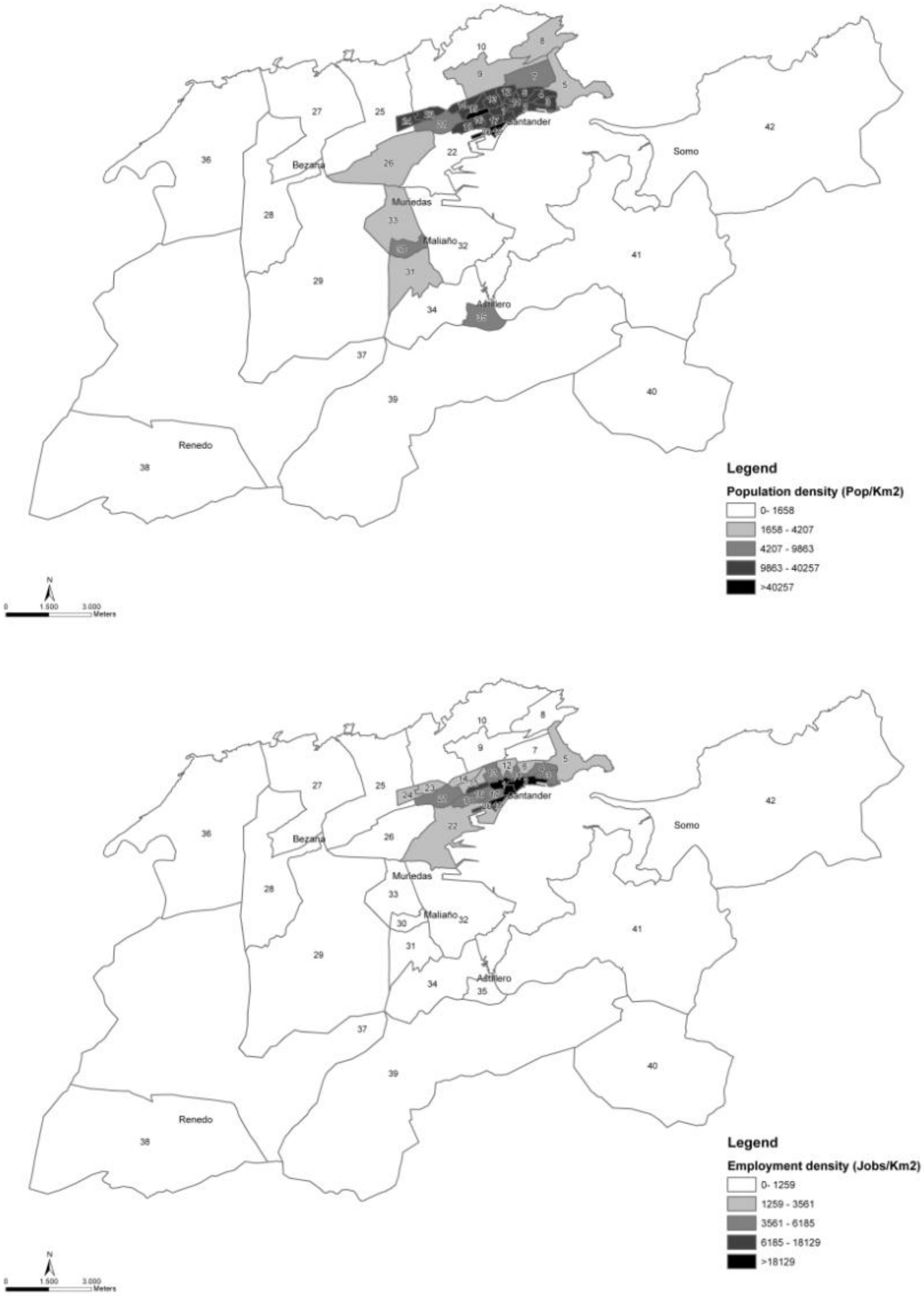

Fig 2-3. Zoning the study area. Population density (top) and Employment density (bottom)

2.5.2. Accessibility indicators

The calibration of the accessibility indicators must be carried out by previously linearising expressions (2.1) and (2.2) taking logarithms to both sides of the equation. The parameters can later be estimated using ordinary least squares (OLS). This type of estimation implies the use of a dependent variable which can be interpreted as a proxy for the accessibility conditions of each zone. This study has taken the proxy variable to be the transport flows generated and attracted by each zone in the base year as the active and passive accessibility indicators, respectively. This interpretation means that the more journeys that are generated or attracted by a zone then the greater will be its active and passive accessibility conditions. The cost variable was obtained using the journey times from the transport model and therefore took congestion into account. The results of the OLS estimation for both indicators are presented in Table 2-1.

Variable	**Active accessibility**		**Passive accessibility**	
	β	p - value	β	p - value
Ln_Jobs/Ln_Res	.265	.000	.269	.000
Cost	-.120	.000	-.195	.000
R^2	0.64		0.72	

Table 2-1. Estimation of the parameters of the accessibility indicators

The journey cost parameter in both models was negative while that of opportunities was positive, which is consistent with theory. Furthermore, the magnitude of the parameter of the passive accessibility model was greater which is also consistent because work related journeys are known to show lower impedance (Ortúzar and Willumsen, 2001). All the parameters were also clearly significant and although a certain degree of colinearity was detected, this was moderate with VIF values lower than 4 in all cases. Similar fits were found with the parameters for both indicators, although slightly better in the case of passive accessibility.

2.5.3. Specification and estimation of the residential and economic activities location models

The location models for both population and economic activities were disaggregated into two sub types of agents. The population was divided into 2 income groups, those with high incomes (>2500 € available household income) and those with medium and low incomes (<2500€). This distinction was made to try and obtain better fitting parameters for the preferences shown by the different agents. The economic activities belonging to the non-basic sector were, in turn, differentiated into those specialising in selling goods and those which provided services. The models were specified using the following variables:

- Journey time by mode of transport from home to work, considering congestion, taken from the transport model and expressed in minutes (CT).
- Active accessibility of each zone to employment (ACCA).
- The natural logarithm of the available housing stock (DW).
- The natural logarithm of the average house price in the zone (PRI).
- The prestige of the zone expressed as a dummy variable with a value of 1 if there are positive environmental externalities such as a beach (PG).
- The passive accessibility of each zone with respect to the population (ACCP).
- The basic employment present in the zone, expressed in 1000s (EMP).
- The presence of each zone as part of the commercial and business centre of the metropolitan area, expressed as a dummy variable with a value of 1 (CBD).
- The characterisation of the zone as a highly developed tourist area expressed as a dummy variable with a value of 1 (TOU).

The metropolitan system has been divided into 4 large sub areas for the economic activity location models which have been introduced as specific constants in the models. The first of these areas, taken as a reference for the identification of the parameters, is practically the entire municipality of Santander city (zones 1 to 20, 22 to 24 and 26). The second constant groups together the zones to the west of Santander which have a high number of retail sector jobs due to the presence of large shopping centres (zones 21 and 25). The third specific constant groups together the zones which are closest to Santander and more strongly integrated into the metropolitan system (zones 27 to 35). Finally, a larger zone has been defined which groups together the areas which are less well integrated into the metropolitan system and mainly located in the eastern part of the study area (zones 36 to 42).

The location model parameters estimated using maximum likelihood can be seen in Table 2-2. The signs of the parameters for the residential location models are in agreement with the theoretical hypothesis. The parameter of the CT variable turned out to be negative in both socio-economic classes which is consistent because the model starts from the hypothesis, derived from urban economic theory, that the residents would tend to prefer locations closer to their places of work, ceteris paribus. The parameter for the higher income residents was of a greater magnitude which reflects a greater disutility of journey time for these users and is coherent with its greater value of time (Glaeser, 2008). The DW variable showed a positive sign associated with the greater attraction for the available residential supply. The PRI variable was only significant in the case of the residents whose incomes were lower than 2500€ while the PG variable was only significant for those people with higher incomes at a 90% confidence level. In both cases the difference in the significance of the variable according to income would appear to be coherent given that the higher income households could be indifferent to the higher housing costs in certain areas, while the prestige factor may not be important for medium and low income households taking into account that the estimation data are based on revealed preferences. For both socio-economic classes the active accessibility variable showed a parameter that was clearly not significant which led to its removal from the model. This lack of significance could be due to part of its effects being captured by CT. Even when the residential location model did not incorporate the active accessibility variable it continued to be sensitive to the costs of the home-work journey, which is an essential characteristic for evaluating policies and projects related to transport. The models had a similar fit for both types of residents using the McFadden's R^2 value as the indicator compared with the constants only model.

Variable	Residents- <2500€		Residents- >2500€		Jobs- Retail Sector		Jobs- Service Sector	
	β	**p -value**	**β**	**p -value**	**β**	**p -value**	**β**	**p -value**
CT	-.106	.000	-.131	.027				
ACCA	-	-	-	-				
DW	1.098	.000	.867	.020				
PRI	-1.541	.000	-	-				
PG	-	-	.328	.094				
ACCP					.643	.000	.206	.004
EMP					.519	.000	.506	.000
CBD					1.666	.000	1.707	.000
TOU					-1.189	.045	1.430	.000
K2					2.404	.000	-.554	.000
K3					-.412	.009	-.083	.425
K4					-1.010	.000	-.356	.002
L(θ)	-776.41		-235.03		-1234.68		-2435.89	
L(0)	-1677.91		-495.23		-4208.60		-7475.33	
R^2	0.53		0.52		0.70		0.67	
N	515		152		1126		2000	

Table 2-2. Parameters estimated for the residential and economic activities location models

The parameters had theoretically correct signs in all cases for the economic activities location models. The specific zonal constants were generally negative with respect to the reference, showing that the municipality of Santander has a high location utility probably derived from economies of agglomeration. Only the K2 constant of the retail location model had a parameter with a positive sign because it captured the concentration of jobs due to the presence of large shopping centres. Furthermore, the K3 constant of the service sector location model was not significant although it also had a negative sign. The parameter of the CBD variable had a positive sign and a high magnitude in both models which highlights the importance of the urban centre as an area with a high concentration of non-basic sector jobs. The dummy variable TOU showed opposing signs in both models probably because highly developed tourist zones have high concentrations of service sector jobs, mainly in catering, while retail development in these areas is much lower. The models had better fits than the residential location models with McFadden R^2 values of 0.70 and 0.67 for retail and services, respectively.

2.5.4. Specification and estimation of the real estate pricing model

The implicit prices model was estimated using OLS. In the present application of the model only residential type properties have been considered due to a lack of reliable price data for other typologies. The resulting parameters (see Table 2-3) correspond to the following variables:

- Average dwelling surface area in square metres of the zone (M2).
- Proportion of dwellings with the use of a lift in the zone (LIFT).
- Proportion of dwellings with balcony in the zone (TER).
- Proportion of dwellings with garage in the zone (GAR).
- Average waiting time for public transport in the zone (TESP).
- Interaction between the presence of a bus stop averaging less than 100 m from the group of properties in the zone and the average number of lines serving the stops in the zone (LIN · BUS).
- Active accessibility of the zone (ACCA).
- Distance in minutes to the urban centre from the zone centroid (TCBD).
- The presence of each zone as part of the commercial and business centre of the metropolitan area, expressed as a dummy variable with a value of 1 (CBD).
- The Presence/Absence of beach in the zone expressed as a dummy variable with a value of 1 (BCH).
- The prestige of the zone expressed as a dummy variable with a value of 1 (PG).
- The ratio between residential demand and supply in the zone (DS).

Some of the variables that were originally introduced into the model had to be discarded because of high colinearity, especially between the square metres variable and the average number of bedrooms and bathrooms in the properties of the zone. These latter variables were, therefore, removed from the model because the M^2 variable is sufficiently explanatory on a zonal scale. The DS variable incorporated into the specification of the models can be calculated using the following equation:

$$DS(o) = \frac{D(o)}{S(o)} \cdot 100 \tag{2.10}$$

Where:

D(o) is the number of residents in zone *o*

S(o) is the number of habitable square metres in zone *o*

Variable	β	P - value
(Constant)	11.713	.000
M^2	.003	.001
LIFT	-	-
TER	.013	.000
GAR	.005	.000
LIN · BUS	.036	.006
ACCA	-	-
TCBD	-.018	.001
TESP	-.006	.037
CBD	.363	.001
BCH	.188	.012
PG	.830	.000
DS	.037	.008
R^2	.935	
R^2 adj	.913	

Table 2-3. Parameters estimated using the hedonic model

This indicator is a measure of the level of occupancy in a zone in the sense of how many residents (demand) are present to each one hundred square metres (supply).

The parameters had theoretically consistent signs and were significant at a 95% confidence level. The LIFT variable corresponding to the proportion of properties equipped with a lift and the ACCA variable corresponding to the active accessibility were removed from the specification because they were clearly not significant. The policy variables, TCBD, TESP and LIN · BUS also proved to be significant and had the correct signs. An increase of one minute in the journey time to the CBD can reduce the average value of properties in a zone by 1.8 % and similarly an increase of one minute in average waiting times at the bus stops in a zone could result in an average price reduction of 0.6%. Changes made to available public transport, for example the location of a bus stop at an average of less than 100 metres from the properties in a zone, were shown to have a positive influence on prices by 3.6% per line serving a stop. The variable DS showed the expected positive sign and was clearly significant. The model's goodness of fit was quite high with a corrected R^2 of 0.91. Two atypical values (outliers) were removed to guarantee the efficiency of the estimators. The chosen selection criteria was that of excluding those cases which showed a residual variable of over 2.5 typical deviations. A Kolmogorov – Smirnov test was applied to check if the model residuals distributed Normal. The test gave a p – value of 0.94 so the null hypothesis of Normal residual distribution could not be rejected.

2.6. Goodness of fit of the model

The goodness of fit of the model was evaluated by simulating the base year and the results were compared with those observed from statistical sources. The transport network and flows were modelled using ESTRAUS software, while the overall group of land-use models were programmed using MATLAB language. Finally, the routines programmed in MATLAB were connected to ESTRAUS to check the interaction between the transport and land-use submodels. The equilibrium solution between residential location, economic activities location, real estate prices and transport can be found using a simple iterative algorithm after having established a stop criterion. After reaching equilibrium in the transport subsystem, the stop criterion can be set as when the variation between one iteration and the next in the location of population, activities and real estate prices is lower than a pre-determined percentage. This research established that the variations in the land-use sub system should be lower than 1% in each one of the 42 zones making up the overall urban system. The model reached the equilibrium solution relatively quickly after 23 iterations.

According to the R2 test the residential location models showed a fit with the observed population of 0.62 and 0.70 for individuals in households with incomes of more and of less than 2500 € respectively (see Fig 2-4). The models managed to capture the location pattern even though they showed an average absolute error of about 20%.

Turning to the fit of the economic activity location model, the retail activity submodel showed a good fit with an R^2 of 0.93, whereas the service sector location model had a slightly inferior R^2 (0.9). The better fit of the retail location model is due to the higher concentration of jobs in this sector mainly found in the city centre and areas where large out of town shopping centres are located, while the location of service sector jobs showed that it depended on a greater number of factors showing high groupings at certain specific locations like tourist areas. Overall, the average absolute error for the location of economic activities was around 18%.

Finally, the goodness of fit of the real estate price model had an R^2 of 0.79. The model had greater errors in the zones located further away from the capital where it tended to overstate average property values. The overall average absolute error was slightly lower than that of the economic activities location model, at around 17%.

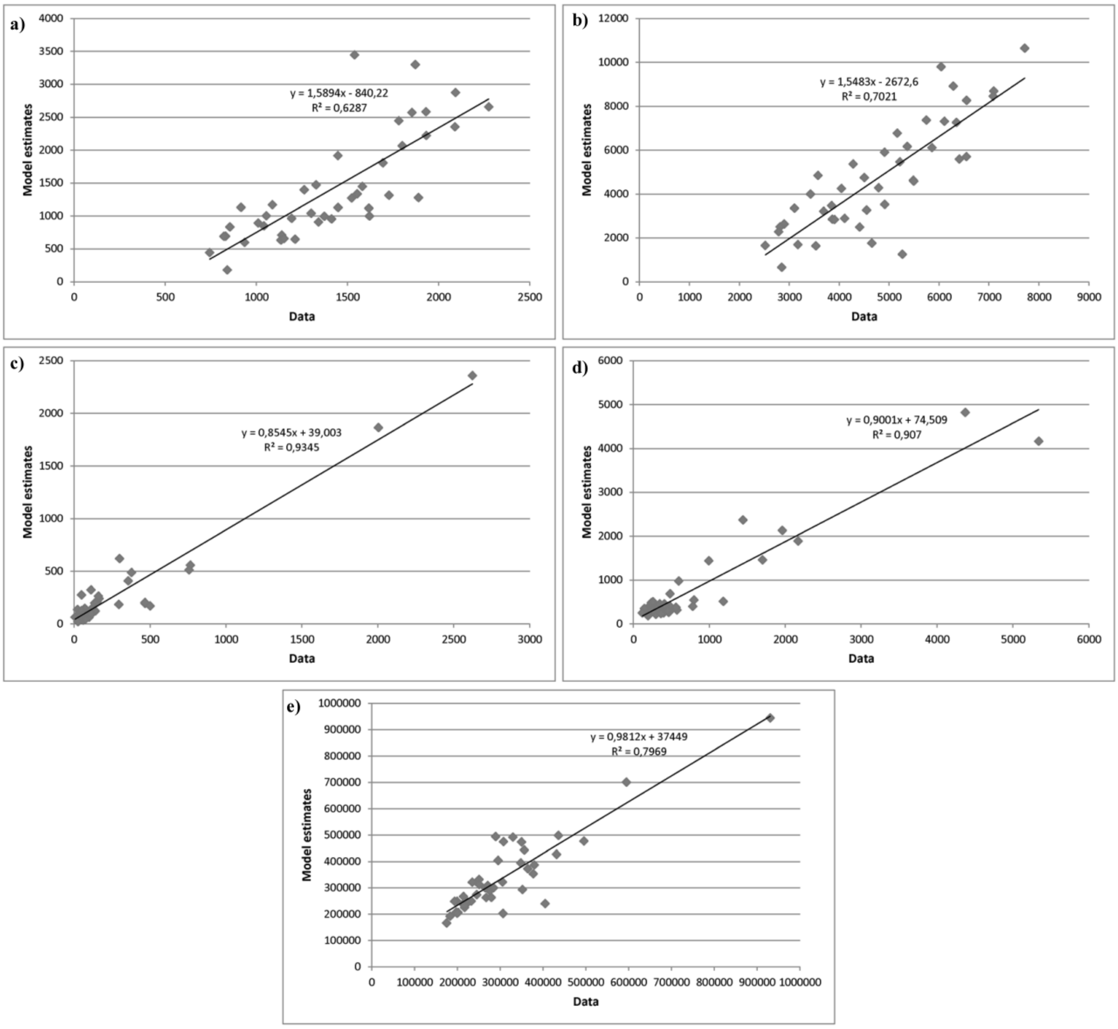

Fig 2-4. Estimations of the model vs. statistical data for: residential location >2500€ (a), residential location <2500€ (b), retail sector activity location (c), service sector activity location (d) and real estate prices (e).

2.7. Conclusions

This article has presented a land-use and transport interaction model which can perform estimations of changes in the location of population, economic activities and real estate prices as a result of the introduction of policies and projects relating mainly to transport. The model was classified as a LUTI simulator between first and second generation as it combines certain aspects of spatial interaction models with random utility theory and hedonic regression techniques for predicting real estate prices. The model was later applied to the metropolitan area of Santander where its parameters were calibrated and the goodness of fit evaluated using observed data from the base year.

The parameters estimated using the different submodels showed theoretically consistent signs. The travel times from home to work were shown to have an explanatory role in household location, causing more disutility for people with incomes of over 2500€. Other aspects such as property prices or the prestige value placed on a zone were only significant according to the income levels of the households which also agreed with theoretical expectations. In the case of Santander, the active accessibility of each zone did not prove to have significant weight in explaining residential location, although part of its effects could have been captured by the transport costs. Nevertheless, passive accessibility along with other factors showed significant weight in determining the location of economic activities. The model was also able to demonstrate the importance of Santander and especially the city centre, as places where high numbers of jobs are concentrated. The aggregated model of real estate prices showed sensitivity to the available levels of transport services such as waiting times, the number of public transport lines serving an area and the journey time to the CBD by private transport.

The model calibrated as a whole reproduced, to an acceptable level, the spatial distribution of the population and economic activities in the study area as well as the real estate prices in the zones. However, the fit of the model could definitely be improved by further disaggregating the types of households and economic sectors as well as by data collection to allow for more complex specification of the utility functions.

Capítulo 3
MODELING THE SPATIAL INTERACTIONS BETWEEN WORKPLACE AND RESIDENTIAL LOCATION

3. MODELING THE SPATIAL INTERACTION BETWEEN WORKPLACE AND RESIDENTIAL LOCATION[2]

3.1. Resumen

El uso del modelo Logit Multinomial (MNL) para la simulación de la localización residencial de los hogares de un área urbana ha recibido críticas debido a que se basa en la hipótesis de la Independencia de las Alternativas Irrelevantes (IIA) la cual no permite que exista correlación espacial entre zonas de residencia. Además no está claro en qué grado afectan la localización del lugar de trabajo y la accesibilidad a empleos en las decisiones de localización realizadas por los hogares, es decir, si la elección del lugar de residencia está condicionada a la elección del lugar de trabajo o si tales elecciones se realizan conjuntamente.

[2] Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P. (2013) Modelling the spatial interactions between workplace and residential location. *Transportation Research Part A: Policy and Practice*, 49, 110-122. http://dx.doi.org/10.1016/j.tra.2013.01.008

En este apartado se especifican modelos de elección residencial Logit Jerárquico (NL) y Logit Jerárquico Cruzado (CNL) para compararlos con el modelo más sencillo MNL e investigar si existe o no correlación espacial entre alternativas de localización. Adicionalmente se ponen a prueba diferentes supuestos considerando la elección de zona de residencia y la elección conjunta de zona de residencia y zona de empleo.

Esta serie de modelos fueron estimados con datos provenientes del área urbana de Santander (España). Los resultados indican que la inclusión en la especificación del modelo de correlación espacial entre zonas mejora el ajuste significativamente. Los tiempos de viaje casa –trabajo fueron un factor estadísticamente significativo en la elección de lugar de residencia mientras que la accesibilidad al empleo presento como variable un signo correcto pero no fue estadísticamente significativa.

3.2. Introduction and objectives

Land use and transport interaction models (LUTI models) are based on the hypothesis of a strong interrelationship between the population location pattern, the economic activities location pattern and the functioning of the transport system (Barra, 1989). Classic LUTI models such as developed by Lowry (1964) postulate that home – work journey costs are one of the basic factors in explaining and simulating household location. Other more recent models (Waddell, 2002) also consider measurements of accessibility to employment in determining how attractive households find specific locations. However, only a limited number of case studies have tried to estimate the weight of these factors (Guo and Bhat, 2001).

Understanding the factors which determine household location decisions is essential if the LUTI models are going to realistically simulate the repercussions of introducing different policies and projects into an urban system. These types of long term choices made by households, condition many of the characteristics of urban systems related both to the demand for transport (trip generation/attraction) and for other distance dependent public services (O'Sullivan, 2007).

The studies based on Random Utility Theory (RUT) stand out among research found in the empirical literature which has tried to determine the most important factors influencing residential location. In fact, RUT based research has become the standard econometric tool both for estimating the determinant factors on residential location as well as in the simulation of these choices within a LUTI system (Pagliara and Wilson, 2010). The application of discrete choice models in the field of residential choice started in the 1970s in the work of Quigley (1976), Lerman (1976) and, from a theoretical point of view, McFadden (1977). These pioneering studies were the forerunners of later studies which have tended to concentrate on specific aspects of residential choice such as its connection with transport choices (Anas, 1981; Eliasson and Mattsson, 2000; Pinjari et al., 2011), the interdependence between the choice of where to reside and the place of work (Anas, 1981; Eliasson and Mattsson, 2000; Pinjari et al., 2011) or the role of accessibility to certain opportunities in the location decisions (Waddell, 1993; Waddell et al., 2007a) . This article is framed mainly within the latter line of work and, more specifically, the relationship between journey time to work, accessibility to employment and residential location.

Residential choice models can, therefore, help in estimating both the more important factors influencing location decision making, as well as the trade – offs households are faced with. Nevertheless, their use has been criticised (Chen et al., 2008) in that most of the models proposed follow a Multinomial Logit (MNL) specification which is based on the Independence of Irrelevant Alternatives hypothesis (IIA). Although this assumption may be considered correct in other choice contexts it is thought to be inadequate in the field of residential choice. Various authors (Hunt et al., 2004) have argued for the need to estimate using models which consider the existence of spatial correlation between alternatives and, therefore, more complex substitution patterns. A simple substitution pattern as assumed in the MNL model ignores the fact that some places, especially those which are closest, could be better alternatives to the rest in the residential choice context of a certain household. Further criticism concentrated on questioning the validity of the assumption that residential choice is exogenous to the choice of work place (Waddell, 1993). Given that there could be a case where the choice of work place is conditioned by a previous residential choice or that both decisions are made simultaneously, some authors have proposed the use of models which consider both types of choice at the same time.

This article presents the specification and estimation of a series of choice models for residential zone and residential zone together with zone of employment. These models will be estimated in order to check the hypothesis, in accordance with basic economic theory, that the locations with better transport conditions and accessibility to employment will have, ceteris paribus, greater utility for the agents and, therefore, will have a greater probability of being chosen as residential areas. The transport conditions in a determined area will be characterised by using home-work journey times considering congestion whereas the accessibility to employment will be based in a gravity type indicator (Handy and Niemeier, 1997). Furthermore, in order to consider the possible existence of spatial correlation, residential choice models will be compared with and without considering the existence of spatial dependence between alternatives. An MNL model will be estimated first, using all the available variables, to estimate the weight of the factors relating to transport to the work place and accessibility to employment. Given that the MNL model is based on the IIA hypothesis, a second step was used to estimate a Nested Logit (NL) model with a structure of correlation between alternatives which was consistent with the maximisation of utility for all possible values of the explicative variables. Finally, a Cross Nested Logit (CNL) model was estimated considering a free structure of correlation between the alternatives. The specification and estimation of the three models presented here: MNL, NL and CNL leads to posing a question of methodological interest in the field of residential choice: do the models considering spatial dependence between alternatives have a significantly better fit to the data?

The estimated models were applied to a case study which was the urban area of Santander (Spain). The results confirmed that journey times to work place had the expected signs and were a statistically significant factor in explaining household locations. Nevertheless, the accessibility indicator was not significant even though it had the positive sign assumed in the hypothesis. The models considering the possible existence of correlation between alternatives, showed, according to the likelihood-ratio test (LR) a goodness of fit significantly better than that of the MNL, both in the case of the single choice models dealing with residential zone and, at a higher level of confidence, in the choice models dealing with residential zone together with work place location.

The article also provides empirical evidence, in the context of a medium sized city, showing that the cost of the journey from home to work continues to be a statistically important factor when explaining household residential location choices. Previous research also supports this conclusion although the data came from larger metropolitan areas (Dallas – Fort Woth, Puget Sound, etc.) which have many notably differences to medium sized cities (Vernon, 1997). This conclusion is supported with the use of simpler discrete choice models (MNL) as well as by using those that consider spatial correlation between alternatives (NL and CNL).

This work is structured in the following way. Section 3.3 presents a review of the state of the art in the field of residential location modelling. Section 3.4 describes the methodology followed and the formulation of the MNL, NL and CNL models. Section 3.5 presents the application of the methodology to the urban area of Santander and the specification and estimation of the models in order to check the proposed hypotheses. Finally, the conclusions coming from this work are presented in section 3.6.

3.3. Bibliographic review

Work concentrating on the modelling of residential location already has a long history as an independent area of study. There are currently various traditional fields of research that can be identified. Starting from the classification proposed by Pagliara and Wilson (2010), the present article differentiates three broad approaches: based on urban economic theory, based on spatial interaction theory and the approach based on econometric modelling.

The origins of location theory derived from urban economic theory can be traced back to the classic work of Von Thünen (1826) on agricultural land use and rents. This pioneering work later became the basis for the creation of the distribution theory of land use and rents in urban areas mainly proposed in the works of Alonso (1964), Muth (1969) and Mills (1972a). In the case of residential land use, the nucleus of the theory rests in the modelling of certain trade – offs facing households when choosing where to live. According to the classic theory, the most important trade – off is formed by the dilemma between choosing lower transport costs to the Central Business District (CBD) (where the hypothesis assumes all the jobs are concentrated) and the space occupied by the household. This dilemma can be mathematically modelled by using a bid rent function which decreases as distance to the CBD increases. Following the Alonso – Muth – Mills model, urban location theory has continued to adapt to the ever increasing complexity of urban systems by expanding the classic theory. An excellent systemisation of research can be found in Fujita (1989). It should be pointed out that various LUTI models have also been derived from applied economic – urban theory such as the Penn – Jersey model developed by Herbert and Stevens (1960) from the work of Alonso.

The second of the traditional areas in which the problem of urban residential location has been addressed has been based on spatial interaction theory. This theory appeared at the end of the 19th century to try and explain the regularities found in spatial population flows. The theory establishes an analogy between the movements of people and the attraction of physical objects using universal gravity law. Relevant work in this line can be found in Reilly (1931), Hoyt (1939), Zipf (1949), Isard (1956) and Hansen (1959) whose work is considered to be the beginning of modern spatial interaction theories. However, it was Wilson (1970) who provided a new interpretive framework to the theory of spatial interaction by overcoming the gravitational analogy and substituting it with a probabilistic paradigm based on the maximisation of the entropy of the studied system. Currently, spatial interaction theory under Wilson's paradigm is considered to be a realistic basis for making residential location predictions using optimization techniques with known destination constraints.

The third line of investigation in the field of residential choice is derived from econometric modelling and, more specifically, discrete choice models. This type of model has proved to be very useful in the field of applied simulation, for making predictions about residential choices made by households about dwelling or residential area. Work developed in the 1970s by Quigley (1976) and Lerman (1976) can be considered as pioneering in this type of modelling, while research carried out by McFadden (1977) signified a considerable advance in this type of study by addressing both the problem of applying the MNL model to choice groups containing many alternatives and to the consideration of the possible correlation between alternatives by introducing the Generalized Extreme Value (GEV) family of models.

As indicated in the introduction, three major lines of investigation can be differentiated in the application of discrete choice models in the field of residential location. A lot of research has taken place relating household location decisions with transport choices. Lerman (1976) specified a choice model which connected residential location, mode of transport to work place and car ownership based on a household survey carried out in the urban area of Washington D.C. The model estimated by Lerman was one of the first to use logit techniques in simulating location choices as a real alternative to traditional techniques based on spatial interaction theory. Anas (1981) applied a nested logit choice model connecting location with mode of transport using aggregated data to the metropolitan area of Chicago. The estimations made by Anas showed that the use of aggregated data based on small areas was practically analogous to the use of disaggregated data. In more recent research Eliasson and Mattsson (2000) specified a choice model which connected residential location, journey frequencies, destination and mode of transport. The model developed by Eliasson and Mattsson demonstrated the existence of heterogeneity between destinations as well as the agents within a coherent microeconomic structure similar to that derived from a standard NL model.

Work involving the relationship between accessibility conditions and residential choice has formed another of the main streams of research. Noteworthy research by Guo and Bhat (2001) considered measuring accessibility to green zones, schools, employment opportunities and other public services in a household survey performed in the urban area of Dallas – Fort Worth (U.S.A.). The authors showed that accessibility to employment did not appear to be an explanatory factor for residential choice except in the case of highly educated workers. The other measures of accessibility to opportunities which were either leisure or commercial were significant as explanatory variables for making residential choices. Srour et al (2002) calculated accessibility indexes to be used as explanatory variables in a residential choice model also applied in the Dallas – Fort Worth area. The indicators of accessibility to employment, retail premises and public park space were shown to be significant in the resulting residential choice model and accessibility to employment was shown to be more important than the measures of accessibility to other opportunities. Chen et al. (2008) estimated a MNL model with indicators of accessibility to employment, open spaces, retail areas and leisure opportunities in a panel of households in the metropolitan area of Puget (U.S.A.). The authors found that the journey distance to work was a significant factor in residential location. They also underlined the existence of a clear trade – off between journey distance to work and accessibility to open spaces which led them to conclude that households only chose more densely populated areas when they were considerably closer to their places of work.

Although less numerous, another field of research has paid attention to the interrelationship between residential and work place location choices made by households. As shown by the Alonso – Muth – Mills type of models based on the assumption that the urban system being studied has a mono-centric nature with all employment being located around the CBD. This assumption has been one of the more criticised within classic economic theory mainly because of the growing tendency of urban systems in presenting various sub-centres and a growing dispersion of places providing employment (Glaeser, 2008). Within the framework of discrete choice studies, Waddell (1993) specified a choice model addressing place of work, place of residence and type of tenancy based on a household survey in the Dallas – Fort Worth area. Given the large number of possible alternatives in the household choice group, the specified model was based on a survey of alternatives which followed the methodology established by McFadden (1977). The MNL and NL models estimated by Waddell demonstrated that the grouped choice specification of work place and residential location fit the data better. Waddell et al. (2007a) later examined the interdependence between the residential choices and the place of work in the urban area of Puget. To address the problem of the high number of alternatives the authors used a methodology which analysed the sequence of residential and work place location choices by using latent market segmentation. The results showed that this methodology improved the fit of the residential and employment location choice models estimated separately.

One of the most frequently repeated criticisms made against residential choice models based on the use of multinomial logit type specifications has been their inability to address spatial correlation between alternatives based on the supposition of IIA. Haynes and Fotheringham (1990) pointed out that aspects such as spatial aggregation or spatial contiguity could produce correlation between alternatives and, therefore, substitution patterns which went against the supposed IIA. The solution to this problem has appeared with the specification of models which allow for more complex substitution patterns. The GEV family of models developed from the work of McFadden (1977), notably the NL model, has provided an initial solution to this problem. However, the NL model still requires the analyst to specify a priori, a correlation structure (Hunt et al., 2004; Pellegrini and Fotheringham, 2002). Bhat and Guo (2004) proposed a model derived from the GEV family which they called Mixed Spatially Correlated Logit (MSCL) for the case of spatially correlated alternatives. This model is based on a special case of the Generalized Nested Logit (GNL) model as originally formulated by Wen and Koppelman (2001) with the incorporation of a distribution of mixes to capture the heterogeneity of tastes between individuals. Bhat and Guo applied the model to a database with the residential choices of a series of households in the metropolitan area of Dallas – Fort Worth. The results of the estimation revealed the importance of the home-work journey costs along and the employment and retail accessibility indicators in choosing where to live. The consideration of the spatial correlation between alternatives improved the fit of the MSCL model over a MNL model. Sener et al. (2011) presented a Generalized Spatially Correlated Logit (GSCL) model to analyse residential choice with correlation between alternatives. This type of model was derived from the GEV family and was based on the MSCL model developed by Bhat and Guo with a more flexible correlation structure between alternatives in which the assignment parameters were considered as a function of a series of zonal characteristics. The GSCL model was applied to a sample of households in the metropolitan area of San Francisco and showed a better fit than the MNL and MSCL models.

3.4. Application of discrete choice techniques to modelling residential choice

This section summarises the modelling structure used based on the different models within the logit family. The main features of the more basic MNL model will be described first. This will be followed by introducing the NL and CNL models which allow modelling the existence of correlation between alternatives.

3.4.1. The multinomial logit model

Within the family of discrete choice models, the MNL model is the one that has been applied on most occasions in the field of residential choice because of its greater flexibility and simplicity of estimation. The overall utility U of a residential alternative i for a household n can be separated into two components: a systematic utility V_{in} and a random utility ϵ_{in}:

$$U_{in} = V_{in} + \epsilon_{in} \quad (3.1)$$

Where V_{in} takes the form:

$$V_{in} = \sum_{k} \beta_k x_{ikn} \quad (3.2)$$

Where parameters β_k are to be estimated and where variables X_{ikn} can refer to zonal characteristics in the case of residential choice models, or even in terms of the interaction between the socio-demographic characteristics of the households n and the characteristics of zones i. In household choice models and, generally, in any discrete choice model with alternatives having particular characteristics at an individual level, a series of specific constants can be specified for each alternative. These constants capture the effect of the mean of all the factors which are not observed by the explanatory variables (Ben-Akiva and Bierlaire, 1999). However, in residential choice models, given the aggregated character of the alternatives and the fact that they present characteristics which are identical for all the individuals, specific constants cannot be introduced in the i-1 utility functions. This is because the generic parameters of the different zonal variables may get confused with the specific constants in the estimation. It is still possible though to include dummy variables within the utility functions that represent the influence of large residential districts. However, this practice is not overly recommended because it makes the model less sensitive to simulating the introduction of new policies (Waddell, 2010). Finally, the probability that a household n chooses to live in zone i can be written as:

$$P_{in} = \frac{e^{\lambda V_{in}}}{\sum_{j \in C} e^{\lambda V_{in}}} \tag{3.3}$$

Where λ is a scaling parameter and C is a choice group.

3.4.2. Models of the GEV family: NL and CNL

In the context of residential choice models, although the MNL model is easily estimated even in the presence of large choice groups, the IIA property is present. As mentioned earlier, even though this property may be acceptable in multiple choice concepts, it is doubtful that it can be applied to choices that have a strong spatial element which is the case of making a residential choice. So the application of a MNL model to a group of residential choices may lead to the estimation of biased or inconsistent parameters.

The models of the GEV family proposed by McFadden constitute a series of specifications which allow a variety of substitution patterns between alternatives (Train, 2009). All the models in this family share the property that the portions of unobserved utility of all the alternatives distribute together as a generalized extreme value. In the case of the more widely used model from the GEV family, the NL model, the alternatives can be grouped into nests. Within each nest the IIA property is maintained, which isn't true for the alternatives of different nests. In the present study the probability of household n choosing a residential zone i is given by:

$$P_{in} = \frac{e^{\lambda_k V_{in}} (\sum_{j \in C_k} e^{\lambda_k V_{jn}})^{\lambda/\lambda_k - 1}}{\sum_{\ell=1}^{K} (\sum_{j \in C_\ell} e^{\lambda_\ell V_{jn}})^{\lambda/\lambda_\ell}} \tag{3.4}$$

Where the parameter λ/λ_k is a measure of the degree of independence of the unobserved utility among the alternatives in nest k. The value of λ/λ_k should be between 0 and 1 in order for the model to be consistent with the maximisation of the utility for all the values of the explanatory variables. However, the alternatives belonging to each nest should be specified a priori and require, in the case of the residential location simulation, that the grouping of the residential areas be based on the possible common characteristics not captured by the independent variables.

Ben – Akiva and Bierlaire (1999) proposed a new model in the GEV family which they called Cross Nested Logit (CNL). The CNL model is an extension of the NL model where an alternative is allowed to belong to more than one nest at the same time. Given this characteristic, the CNL model can be applied in the case of choosing a residential zone without the prior need to impose a spatial structure of correlation between alternatives.

According to Bierlaire (2006) there have been various formulations of the CNL model presented in the literature such as proposed by Small (1987) or Vovsha (1997). However the most general were those presented by Ben – Akiva and Bierlaire (1999) and Wen and Koppelman (2001). The formulation of Ben – Akiva and Bierlaire is based on the following function:

$$G(x_1,....,x_j)=\sum_k(\sum_{j\in C}\alpha_{jk}x_j^{\lambda_k})^{\lambda/\lambda_k} \qquad (3.5)$$

Where k is the index of the nest, λ_k is the scaling parameter associated to nest k and α_{jk} represents the degree alternative j belongs to nest k. The formulations proposed by Wen and Koppelman and Ben – Akiva and Bierlaire are equivalent although Wen and Koppelman proposed the condition λ=1, a common norm in the GEV family of models. Bierlaire (2006) has formally demonstrated that this function fulfils the conditions proposed by McFadden (1977) which allow the CNL model to belong to the GEV family. In addition, the α_{jk} parameters should fulfil the following constraint in order to make the model identifiable:

$$\sum_k \alpha_{jk}=1 \quad \forall j \qquad (3.6)$$

Finally, in a CNL model, the probability that a residential area i will be chosen by a household n is given by:

$$P_{in} = \sum_k \frac{\alpha_{ik} e^{\lambda_k V_{in}} (\sum_{j \in C} \alpha_{jk} e^{\lambda_k V_{jn}})^{\lambda/\lambda_k - 1}}{\sum_{\ell=1}^{K} (\sum_{j \in C} \alpha_{j\ell} e^{\lambda_\ell V_{jn}})^{\lambda/\lambda_\ell}} \tag{3.7}$$

3.5. Application of residential choice models to the metropolitan area of Santander

3.5.1. Demographic and Socioeconomic characteristics of the Study Area

Santander is a medium sized city located on the north coast of Spain. The city is also the capital of the region of Cantabria. The current population of the city is about 182,700 inhabitants in the urban nucleus; however the overall metropolitan area contains more than 260,000 residents.

The distribution of the population in the study area has dramatically changed recently with a striking increase in the number of people living in the peripheral municipalities during the period 2001-2009 coinciding with a very low population growth of hardly 1 % in Santander itself. This supports the hypothesis that the area is becoming more metropolitan and the population is spreading out over a larger area rather than having high concentrations of people in central urban nuclei. This process is normally described as urban sprawl when it is accompanied by the development of low density residential areas (García Palomares, 2007).

More recently, the economy of the metropolitan area has gone through a transition towards the service sector which parallels both the regional and national economies. This can be traced back to the changes occurring in the late 1970s starting a process of continual service sector growth which has led to this sector holding the dominant position within the economic structure of the study area. Grouping all the municipalities in the study area together, in 1986 the service sector represented 60.3% of all employment (Nogués, 1990), whereas in 2008 this figure had grown to 71.8%.

Removing the municipality of Santander, the service sector continues to be the predominant provider of employment in the rest of the municipalities, with almost 50% of overall employment. This percentage is slightly above the employment figures provided by secondary industries (manufacturing and construction), which supports the notion of the spread of deindustrialisation throughout the municipalities located around the bay. Although in the 1980s most employment was still provided by secondary industries (47.1% of all jobs) and the primary sector still had an important economic role to play (more than 10% of all jobs), both sectors are currently less significant, especially the primary sector which in 2007 represented less than 1% of overall employment.

The economic structure of the study area can be seen to have undergone important changes over a relatively short recent time scale. These changes can be summarised as:

- Economic activity has definitely moved towards the service sector.
- Although manufacturing still exists (with an important contribution from construction in many municipalities) its presence has been reduced.
- Primary sector industries have lost so much activity that they can be regarded as practically marginal. This is more noteworthy in the eastern municipalities where there has always been a tradition of dairy farming.
- From a spatial point of view, the traditional zonal division into the capital as the central provider of services, the north western zone containing the industrial municipalities and the eastern zone occupied by dairy farming and leisure activities has to a certain extent been redrawn. There is currently a dominance of tertiary sector industries in almost all the municipalities.
- Activity continues to be highly concentrated in the municipality of Santander although a certain trend towards some diffusion of these services can be seen in favour of other metropolitan municipalities.
- The creation of large out-of-town shopping malls has strengthened this diffusion in terms of retail trade by drawing a large number of shoppers outside the commercial centre of Santander towards the suburbs whilst at the same time attracting people from other neighbouring areas.

3.5.2. Available data

A sample of 534 households was used. This sample came from a simple random sample mobility survey carried out in 2008 on a section of the population from households in the urban area. Given that the study concentrates on the influence of factors relating to accessibility and transport to work, the chosen sample is made up of households that have only one member with a salary. This choice removes the problem associated with the number of workers present in each household when specifying the model (Waddell, 1996).

In fact, of the 534 surveyed households, 396 had at least one worker and of those 275 had only one worker. The geographical distribution of the properties according to the number of workers they hold does not follow any particular pattern and can be considered to be quite disperse, it has been assumed that the hypothesis of including one-worker households does not affect the end result of the study.

The survey collected information on household location, their basic characteristics (number of members, income level…) and their revealed mobility preferences detailed in a journey diary. The postal address was used to code each observation into a geographical information system (GIS). The introduction of the data into the GIS meant that certain variables could be obtained about the surroundings and locations of each one of the households by cross-referencing the point location data of each home with the variables present in the land use zoning used. This zoning divided the municipal area into a total of 26 zones. Given the limited size of the overall study area it was thought that 26 zones with an median area of 0.33 Km^2 fulfilled the condition imposed by Anas (1981) who worked with zoning in squares of 0.65 Km^2.

The variables contained in the database are presented in Table 3-1 and can be classified into three types: variables relating to the environmental conditions in the zone, variables relating to the accessibility and transport to work conditions, and interaction variables. The summary of the descriptive statistic of the variables is given in Table 3-2.

Variable	Description	Type
Accessibility	Accessibility to employment zone)	Transport
Journey time	Journey time between residential zone – employment zone	Transport
Residence/work	Dummy =1 where the zone of residence and zone of work coincide)	Transport
Waiting time	Average waiting time at public transport stops in the zone	Transport
Employment	Nº of jobs in the zone	Environmental
Foreigners	Nº of none EU foreign residents present in the zone	Environmental
Dwellings	Natural log of the nº of dwellings in the zone	Environmental
Prestige	Dummy for a zone with special prestige	Environmental
House price	Average house price for the zone	Environmental
Learning	Learning centres at least 1000 m from zone centroid	Environmental
High Income	Dummy for income >2500	Interaction
Age	Age	Interaction
Work	Nº Worker per household	Interaction

Table 3-1. Description of the explanatory variables used

Variable	Mean	Standard Deviation	Minimum	Maximum	Measurement unit
Accessibility	29.34	7.42	10.02	48.05	-
Journey time	13.51	1.88	10.10	18.83	Minutes
Residence/work	0.07	0.25	0.00	1.00	-
Waiting time	10.51	0.78	8.50	12.11	Minutes
Employment	2602.26	2737.53	520	11359	No. Jobs
Foreigners	461.12	224.36	108.00	843.00	No. Foreigners
Dwellings	2650.95	567.20	1072.00	3804.00	No. dwellings
Prestige	0.09	0.29	0.00	1.00	-
House price	287609.6	126695.66	174920.83	930899.18	€
Learning	2.22	1.70	0.00	7.00	No. Schools
Income	2.90	0.79	1.00	4.00	-
Age	43.85	11.91	16.00	82.00	Years
Work	1.8	0.68	1.00	4.00	No. Workers per household

Table 3-2. Descriptive statistic

The choice of explanatory variables was strongly conditioned by the available sources. An attempt was made to make the independent variables take into account the following the three main aspects which are generally considered to be a key part of residential choice:

- The environmental conditions in the residential area: this group contains “employment”, “foreigners”, “prestige”, “house price”, “learning” and “dwellings”.

- The accessibility and transport conditions of each residential area: these are the most interesting for the objectives of this research. This group may include: “accessibility”, “journey time”, “Residence/work”and “Waiting time”.
- The demographic characteristics of the population. Household income has been taken as particularly relevant because budgetary constraints are usually an important conditioning factor on location choice.

Furthermore, other variables referring to the socio-demographic characteristics of households were considered, even though they were not thought to be theoretically relevant or did not turn out to be significant in the process of specifying the residential choice models: these were the age of household members (“Age”), and the number of workers in the household (“Work”).

Interactions consider how the environmental and transport conditions differentially affect households according to their income levels. A dummy variable was used to consider these differential effects in households with income levels above the third quartile (above 2500 €) on the variables “accessibility”, “journey time”, “foreigners”, “dwellings”, “prestige”, “house price”, “learning” and “waiting time”.

The variables “employment”, “foreigners”, “dwellings” and “learning” were established from different statistical operations made by the Spanish Office of National Statistics (INE) and by the Cantabrian Institute of Statistics (ICANE). The “learning” variable can be considered as an accumulated opportunities type of accessibility indicator to centres of primary and secondary education in each of the zones. The average price per zone (“house price”) is collected from a series of real estate sources which represent asking prices on a sample of 845 properties in the study area. The prestige of the zones (“prestige”) depended on their belonging to the city centre or the better known neighbourhoods on a regional or even national scale (e.g. the El Sardinero neighbourhood).

The variables considered to be more relevant to the objectives of this study are those which refer to the accessibility and transport to work conditions: “accessibility”, “journey time”, “Residence/work” and to a lesser extent, “waiting time”. The average waiting times in the zones were calculated from the data provided by the municipal public transport service. The initial hypothesis was that this variable would have a negative effect on a zone's utility, even though it did not account for all the journey to work costs using public transport it was not thought to be fundamental for the goals of this study. The "journey time" variable represents access time in minutes from the head of the household's residential zone to their employment zone. This variable has been calculated with a transport model using the real morning rush hour traffic flows. The times that were included in the database therefore take congestion into account and this variable is expected to present a parameter with a negative sign in the models. "Residence/work" represents a dummy variable taking a value of 1 if the residential zone and employment zone coincide. This variable is expected to have a positive sign which signifies that the households tend to assign greater utility to those zones where their work place is located ceteris paribus. Finally "accessibility" is a Hansen type gravity indicator of accessibility to employment which considers the possible multicenter nature of the urban area. The indicator was calculated from the employment data present in the zones and the following expression was used (Nuzzolo and Coppola, 2007):

$$Acc(o) = \sum_i [\exp(\alpha_2 \cdot Cost(o, d_i)) \cdot jobs(d_i)^{\alpha_1}] \quad (3.8)$$

where *Cost* is a measure of the journey cost between origin *o* and destination d_i by car calculated using a transport model which considered congestion at morning rush hour. Jobs (d_i) are the number of employments present in the destination zone d_i and α_1 and α_2 are parameters to be estimated. The parameters may be calculated by linearization (3.8) taking logarithms to both sides of the expression. The parameters are then estimated by ordinary least squares taking the transport flows generated by each zone as the dependent variable for the accessibility conditions. The parameter α_1 presented an estimated value of 0.26 while α_2 had a value of -0.12. The estimation had a R^2 goodness of fit of 0.7. The parameter of this variable was expected to have a positive value meaning that households tended to give greater utility to the zones with more employment close by.

3.5.3. Multinomial Logit model for choosing residential zone and residential zone considering employment zone

This subsection presents the estimation of two residential choice models using MNL. The first of the estimated models (MNL – 1) considers the location of the place of employment as an exogenous factor. Under this supposition the group of choices available to households is limited to a total of 26 alternatives. The second model to be estimated is a joint logit model which considers the joint choice of place of work and place of residence (see MNL-2 in Table 3-3 with the p-values of the significance of the parameters in brackets). In this case the expression of utility is formulated as:

$$U_{rwn} = V_{rn} + V_{wn} + V_{rwn} + \epsilon_{rwn} \quad (3.9)$$

Where U_{rwn} is the overall utility of choosing zone of residence r and place of work w for household n. V_{rn} is the systematic utility of choosing zone of residence r for household n, V_{wn} is the systematic utility of choosing zone of work place w, V_{rwn} is the systematic utility of the specific combination of zone of residence and work rw and ϵ_{rwn} is the random component of the alternative's utility.

As stated before, some authors have pointed out the need to consider that the choice of residential zone is not independent of the choice of work place zone, because if it isn't considered it may reduce the goodness of fit of the model and bias the estimated parameters (Waddell, 1993). Therefore, in this model the households present a total of 26*26 = 676 alternatives, which is an ample group of choices (see Fig 3-1 to compare the structure of the models). Previous research with large choice groups has used the random sampling of alternatives technique proposed by McFadden, even though this can only be used in the case of MNL models unless additional corrections are used (Lee and Waddell, 2010). However, the current software available for estimations (e.g. Biogeme (Bierlaire, 2003)) enable working with large choice groups but with rather high estimating times.

Type	Variables	Interaction	MNL-1. Choice of residential zone	MNL-2. Joint choice of residential and work location zone
Transport	"Accessibility"	-	.005 (.56)	.006 (.53)
	"Accessibility"	"High Income"	-.016 (.38)	-.016 (.38)
	"Journey time"	-	-.102 (.01)	-.103 (.01)
	"Journey time"	"High Income"	-.000 (.92)	-.007 (.92)
	"Residence/work"	-	.235 (.29)	.233 (.29)
	"Waiting time"	-	-.132 (.10)	-.137 (.09)
	"Waiting time"	"High Income"	-.013 (.94)	-.021 (.91)
Environmental	"Employment"	-	-	.000 (.00)
	"Foreigners"	-	-.000 (.04)	-.000 (.04)
	"Foreigners"	"High Income"	-.001 (.12)	-.001 (.13)
	"Dwellings"	-	1.39 (.00)	1.39 (.00)
	"Dwellings"	"High Income"	1.49 (.07)	1.42 (.08)
	"Prestige"	-	-.897 (.00)	-.934 (.00)
	"Prestige"	"High Income"	1.90 (.00)	1.93 (.00)
	"House price"	-	-.000 (.04)	-.000 (.05)
	"House price"	"High Income"	-.000 (.57)	-.000 (.55)
	"Learning"	-	-.087 (.05)	-.090 (.04)
	"Learning"	"High Income"	.242 (.00)	.246 (.00)
	Null Log-Likelihood		-1739.82	-3479.64
	Log-Likelihood		-1658.58	-3274.40
	LR test Null		162.48	410.48
	Nº Alternatives		26	676
	N		534	534

Table 3-3. Parameters estimated for the MNL residential location models

The residential choice model MNL – 1 estimated with a choice group of 26 alternatives presented a log –likelihood in convergence of -1658.58. From the transport related variables, the "accessibility" parameter was not significant neither for the overall group of households nor for the households with incomes of over 2500 €, although it did present a theoretically believable sign. Nevertheless, the coefficient of "journey time" did turn out to be significant in the sample group of households and it had a negative sign which agreed with the initial expectations.

The parameter of the variable referring to "waiting times", was significant at a 90% confidence level and had a negative sign meaning that households considered that increased waiting times for public transport reduced utility. Finally although the "Residence/work" variable produced a positive parameter it was only significant at a 75% confidence level.

The coefficients relating to the environmental variables such as "foreigners", "dwellings", and "house price" had the expected signs and, in general, were significant. In the cases where they presented theoretically incoherent signs in the population group such as in the cases of "prestige" or "learning", they were as expected for the higher income group of households. So, given that the data are based on a revealed preferences survey, only the higher income households tend to locate in areas with a high number of schools or in prestigious zones because their parameters, once added to the parameters for the overall population, continue to be positive.

The MNL – 2 joint choice model for residential zone and work place zone generally presented similar parameters for all the variables to those estimated using MNL – 1. An additional variable ("employment") was introduced for modelling the choice of work place. "Employment" variable was clearly significant adding a utility of 0.0001 to the alternatives for each additional job.

Both models presented a significant fit with respect to the equiprobable model according to the LR test even though the result was higher when places of residence and work were modelled together (410.48 versus 162.48).

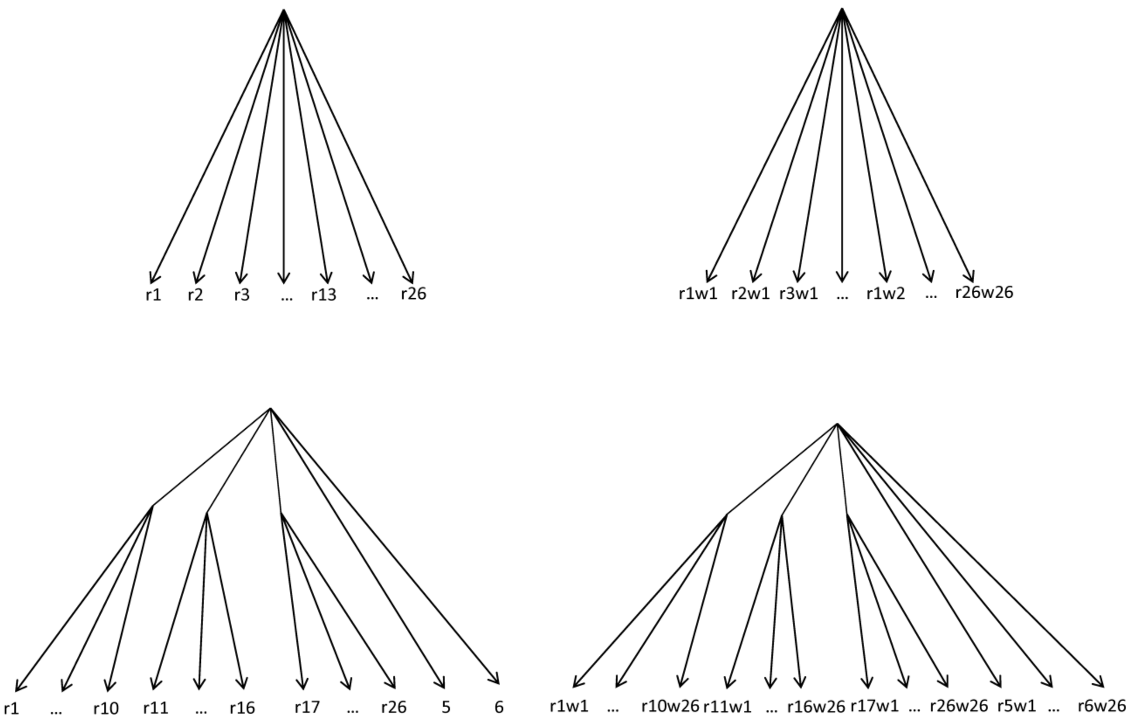

Fig 3-1. Nested structures of the MNL -1 models (top left), MNL -2 (top right), NL -1 (bottom left and NL-2 (bottom right)

3.5.4. Models considering spatial correlation between alternatives: NL and CNL

This section introduces the estimations made for the models considering the existence of spatial correlation between alternatives. Two NL models were initially estimated, the first of these (NL – 1) considered, as in the case of the MNL models, the choice of employment zone as being exogenous to the choice of residential zone, whereas the NL – 2 model again considers the choices of residential zone and employment zone as a join process.

Various specifications were estimated to consider different structures of prior correlation between zones by grouping them into different nests. The specification that was finally selected for NL – 1 (see Fig 3-1 and Table 3-4) had three nests plus two alternatives hanging directly from the root nest. The overall model should fulfil certain conditions in order to be consistent with random utility theory (Train, 2009). Firstly, one of the scaling factors has to be established for it to be estimated. In this case an upper level normalisation of $\lambda = 1$ was chosen. Secondly, the estimated nest parameters should fulfil the condition that λ_k >1 so the coefficient λ / λ_k is between 0 and 1 thereby assuring that an increase in the utility of any of the alternatives from the nest increases the probability of choosing it as a whole.

Finally, one of the proposed nesting structures was chosen with the alternatives grouped into three main nests. These nests divided the study area into three large zones (see Fig 3-2). Firstly, nest A grouped alternatives 1 to 4 and 7 to 10 corresponding to the central and north-eastern areas of the city, a zone that generally had a higher residential status. Nest B grouped together the alternatives 11 to 16 belonging to the neighbourhoods around the city centre which were considered as residential areas with a generally lower status. Finally, nest C grouped together alternatives 17 to 26 corresponding to the western part of the studied area, with a generally varied nature consisting of urban and peri-urban residential areas, as well as industrial zones and transport infrastructure installations. Two alternatives, 5 and 6 were grouped into nest D with a fixed parameter of 1 and therefore directly connected with the root nest. This was because different specifications showed that they had little correlation with the other areas or between themselves. In the NL – 2 model, the same correlation structure was replicated between the alternatives grouping the joint choices of residential zone and work place zone rw according to the residential zone.

The parameters estimated using NL – 1 and NL – 2 were very similar to those obtained from the MNL specifications. Among the environmental variables only the “learning” coefficient became non-significant in the NL models, even though for the higher income households (“learning” interacting with “High income”) the greater number of centres of learning was a positive and significant factor in choosing location. The “waiting time” parameter was the least significant from the variables relating to accessibility and transport to work place, although it kept its negative sign in the overall household survey. Furthermore, the “waiting time” interacting with “high income” parameter presented an unexpected positive sign, even though it was clearly not significantly different from zero. The number of jobs present in the zone variable, “employment”, once again returned a clearly significant parameter. The remaining interesting variables “accessibility”, “journey time” and IN didn’t show any changes in their parameters nor in their signs and practically not even in magnitude and significance. The coefficients of the A, B and C nests were greater than 1 although they were not significant in all cases. In particular, nest C was only significant at a 72% confidence level and at 65% for the residential location and joint residential and work place location models, respectively.

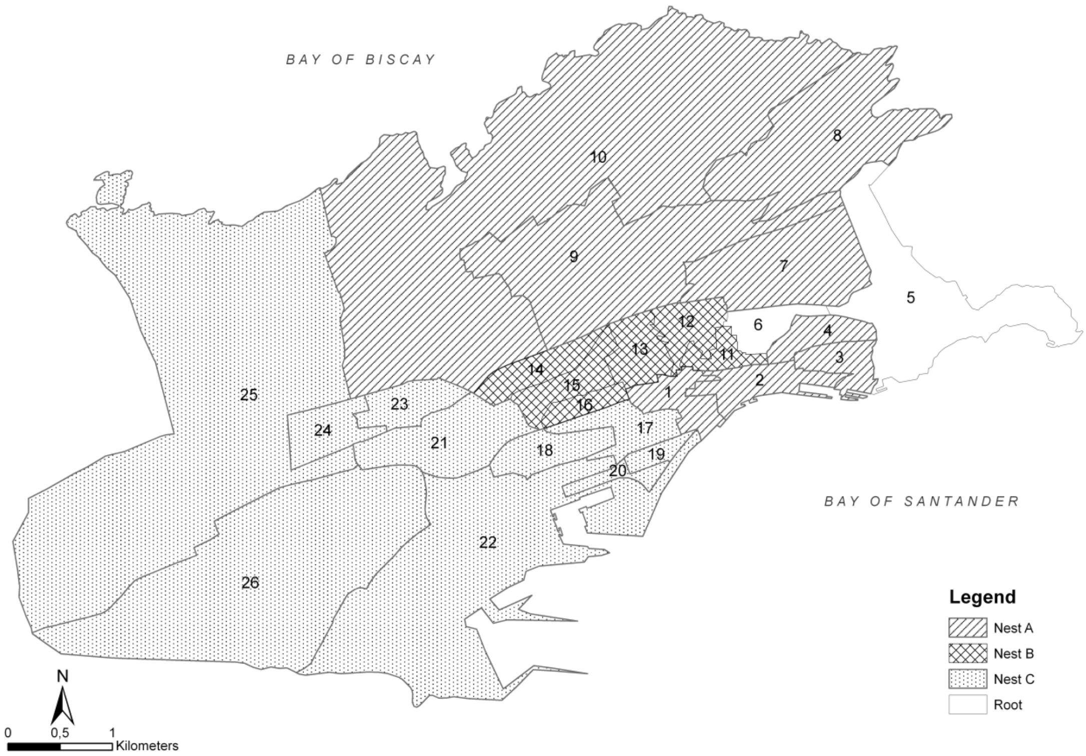

Fig 3-2. Land use zoning used in the study area

Both the residential location model and the joint location model had clearly better and more significant fits than the respective equiprobable models. Furthermore, if their fit is compared to the one obtained for the MNL models using the LR test, the residential location model offers a significantly better fit at a 90% confidence level, while the joint location model presented a better value in the test (8.46) at a confidence level greater than 95%.

Type	Variables	Interaction	NL-1. Choice of residential zone	NL-2.Joint choice of residential and work place zone	CNL-1. Choice of residential zone	CNL-2. Joint choice of residential and work place zone
Transport	"Accessibility"	-	.005 (.51)	.007 (.42)	.006 (.33)	.002 (.80)
	"Accessibility"	"High Income"	-.013 (.39)	-.015 (.38)	-.003 (.79)	-.008 (.62)
	"Journey Time"	-	-.089 (.01)	-.093 (.02)	-.068 (.02)	-.064 (.08)
	"Journey time"	"High Income"	-.000 (.99)	-.009 (.89)	-.000 (.99)	-.014 (.82)
	"Residence/work"	-	.214 (.26)	.236 (.26)	.203 (.14)	.317 (.13)
	"Waiting time"	-	-.097 (.18)	-.083 (.30)	-.069 (.32)	-.127 (.14)
	"Waiting time"	"High Income"	.048 (.76)	.027 (.88)	.061 (.65)	-.047 (.79)
Environmental	"Employment"	-	-	.000 (.00)	-	.000 (.00)
	"Foreigners"	-	-.000 (.04)	-.000 (.04)	-.000 (.13)	-.000 (.15)
	"Foreigners"	"High Income"	-.001 (.12)	-.001 (.13)	-.000 (.07)	-.001 (.06)
	"Dwellings"	-	1.28 (.00)	1.46 (.00)	.531 (.07)	1.26 (.00)
	"Dwellings"	"High Income"	1.15 (.08)	1.34 (.09)	.685 (.13)	1.52 (.06)
	"Prestige"	-	-.681 (.01)	-.685 (.02)	-.339 (.12)	-.921 (.00)
	"Prestige"	"High Income"	1.76 (.00)	1.90 (.00)	1.19 (.00)	1.59 (.00)
	"House Price"	-	-.000 (.01)	-.000 (.01)	-.000 (.00)	-.000 (.16)
	"House price"	"High Income"	-.000 (.50)	-.000 (.40)	-.000 (.53)	-.000 (.83)
	"Learning"	-	-.053 (.17)	-.053 (.23)	.000 (.98)	-.109 (.02)
	"Learning"	"High Income"	.215 (.00)	.236 (.00)	.152 (.00)	.191 (.01)
	NESTA		1.24 (.11)	1.07 (.09)	3.97 (.01)	1.82 (.00)
	NESTB		1.28 (.05)	1.10 (.02)	1.85 (.05)	1.21 (.03)
	NESTC		1.11 (.28)	1.03 (.35)	1.27 (.00)	1.00 -
	NESTD		1.00	1.00	1.00	1.00

Null Log-Likelihood	-1739.82	-3479.64	-1739.82	-3479.64
Log-Likelihood	-1654.73	-3270.17	-1626.64	-3070.01
LR test Null	170.17	418.94	226.35	819.26
LR test MNL/NL	7.69	8.46	56.18	400.32
Nº Alternatives	26	676	26	676
N	534	534	534	534

Table 3-4. Estimated parameters for the NL and CNL residential location models considering the existence of correlation between alternatives

Finally, CNL models were estimated with identical specification to the NL models but without the need for prior detailing a spatial structure of correlation between alternatives. Nevertheless, in order to limit the number of parameters to be estimated in the models, a technique that had been previously used in other studies (Bhat and Guo, 2004) was applied to establish the condition that a zone could only belong to those nests which presented common borders. Only alternatives 5 and 6 continued to hold direct connected with the root nest.

The models had long calculation times because of the many parameters to be estimated, especially in the case of the CNL joint choice model of residential and work place zone. In the case of the joint residential-work place model the software used required several days to calculate the variances – covariances matrix for obtaining the statistical significance (see Table 3-4).

The CNL – 1 residential choice model, showed quite similar parameters to those obtained from the MNL and NL models. Among the environmental variables, the “dwellings” parameter had the expected sign but with a lower magnitude and was only significant at a 93% confidence level. The coefficient of “learning” was clearly not significant even though, the number of schools present in the zones of higher income households (“learning” interacting with "High Income”), did turn out to be significantly different from zero with a positive sign. Among the variables related to accessibility and transport to work not even the parameters of “accessibility” or “waiting time” were significant even though they had a p – value of around 0.3. The parameter of “journey time” however, continued to present the expected negative sign and was clearly significant. Finally, the parameter of the “Residence/work” variable and its significance were similar to those present in the MNL and NL models. In the case of the

CNL – 2 model the estimated parameters were once again quite similar to those already mentioned above. The “accessibility” variable presented a parameter with a positive sign and a magnitude a little greater than estimated in the MNL and NL models as in the case of the CNL – 1 model. The parameters of “journey time”, IN and “waiting time” continued to show signs and magnitudes which were very similar to those estimated using the other models.

In the case of the CNL – 1 all the λ_k scaling parameters were significantly different from one, indicating that the CNL specification helped increase correlation inside the nests. In the joint choice model CNL – 2, nests A and B presented parameters significant, while the nest C took a value equal to 1 which is an indication of the non-existence of correlation between alternatives, a fact similar to the present in the NL models.

The goodness of fit of the CNL models was greater than found using the MNL and NL models. The LR test used in CNL – 1 against the fit of NL – 1 and in CNL – 2 against the fit of NL – 2, was significantly better at 90% and 95% confidence levels, respectively, in spite of the great number of additional parameters that had to be estimated for the CNL type of models.

3.6. Conclusions

This article has presented the specification of three types of discrete choice models MNL, NL and CNL considering the choice of residential zone and the joint choice of residential zone and work place zone. The models were specified to estimate the influence of accessibility and transport to work conditions on the choice of residential zone made by a sample of households. The estimation of these models using data from the urban area of Santander also allowed them to be compared to determine if the models that considered the existence of spatial correlation between alternatives presented a significantly better fit to the data.

In considering the effects of the factors relating to accessibility and transport to work, the Hansen type accessibility to work indicator introduced in the specification of the models was generally found not to be significant, even though it always presented a positive sign for the group of households in the survey. On the other hand, the parameter of the variable referring to home-work journey times was clearly significant in all the models at least a level of confidence of 90% with no significant differences between the overall group of households and the group with higher incomes. This result can therefore be considered to be compatible with the results obtained in previous research such as Guo and Bhat (2001) and Bhat and Guo (2004). The "Residence/work" variable referring to the location of work place and residence in the same zone also had a positive sign using all the models even if it was only significant at a confidence level of between 71% and 87%. Finally, although waiting times for public transport had the expected negative sign in most of the models, they were only significant at a 90% confidence level in the MNL models which offered the worst fit. These facts show that, in the sample analysed, journey times to work were definitely an important factor for households in choosing where to live. Furthermore, the lack of significance of the accessibility indicator could be due to part of its effects being captured by the "journey time" and "Residence/work" variables. It can therefore be concluded that, in agreement with the hypothesis proposed at the beginning of this research, lower journey times to work continue to be an important factor when households are deciding where to live.

The environmental variables, considered to be secondary for the purposes of this study, generally presented the expected negative signs in the cases of "foreigners" (an indicator of the existence of a certain degree of spatial segregation) and "house price" and positive in the case of "dwellings". The parameters of the variables "prestige" and "learning" also behaved differentially according to household income levels: while the prestige of a zone was a strong positive factor for the utility of the alternatives in the case of higher income households, in the case of the overall group of households this variable presented a negative parameter using most of the models. Similar results were found with the parameters of the variable referring to the number of schools.

The joint choice models for residential and work place location zones showed similar parameters to those obtained in the purely residential choice models, including the additional ("employment") variable, always significant and with the correct sign. Furthermore, the LR test showed that the NL and CNL joint choice models offered a better fit than the MNL at a 95% confidence level. This level of confidence was certainly better than the 90% obtained by the purely residential choice models.

The NL models considering the existence of spatial correlation between alternatives had a better fit with the data than the MNL models, even though in some cases the parameters of the nests were not significant to a high level of confidence. In comparison, the CNL – 1 model had a good fit with all the parameters of the nests, significantly different from one and at a confidence level of at least 95%. This fact shows the presence of a certain degree of spatial correlation between the alternatives which was captured more precisely by the CNL model than by the previously specified and closed structure of the NL models. The parameters of the CNL – 2 models were similar to those of other models. Moreover the correlation structure allowed getting a better fit to the data even if this was at the expense of rather high estimating times caused by the greater number of parameters. Therefore this type of model looks promising for the future in the field of residential choice as the capacity of software increases to enable estimations to be made with large choice sets and specifications with the presence of a large number of parameters.

Capítulo 4
MODELING THE SPATIAL INTERACTIONS BETWEEN WORKPLACE AND RESIDENTIAL LOCATION

4. MODELING TRANSPORT AND REAL-ESTATE VALUES INTERACTIONS IN URBAN SYSTEMS[3]

4.1. Resumen

En este apartado se presentan modelos de Regresión Lineal Múltiple (MLR), modelos de regresión hedónica autoregresiva (SAR), modelos de regresión hedónica en el término de error (SEM) y modelos hedónicos Durbin (SDM) para estimar las variaciones de precios inmobiliarios como resultados de cambios medioambientales y en las condiciones de accesibilidad y transporte del área urbana. La bondad de ajuste de los diferentes modelos se ha comparado conjuntamente con una serie de hipótesis sobre la utilidad de especificar modelos que consideren relación espacial entre observaciones. El estudio de caso de este análisis ha sido nuevamente el área urbana de Santander. Los modelos considerando dependencia espacial entre observaciones

[3] Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P., Dominguez, A. (2012) Modelling transport and real-estate values interactions in urban systems. *Journal of Transport Geography* 24, 370-382. http://dx.doi.org/10.1016/j.jtrangeo.2012.04.012

ofrecieron un mejor grado de ajuste en un escenario de fuerte autocorrelación espacial entre los residuos al mismo tiempo que presentaban parámetros estadísticamente significativos con signos coherentes según la teoría de la economía urbana y del transporte. El modelo seleccionado como el mejor mostró incrementos del 1,8% por cada línea de transporte adicional presente en el área cercada a cada vivienda así como una reducción del 1,1% en los precios por cada minuto adicional de tiempo de viaje al centro comercial y de negocios de la ciudad. La cercanía a las estaciones de tren sin embargo implicó, en los modelos considerando relaciones espaciales entre viviendas, reducciones en los precios inmobiliarios.

4.2. Introduction and objectives

Classic urban economic theory proposed by Alonso (1964), Muth (1969) and others is based on the trade-off between accessibility and space. Locations with better access to the Central Business District (CBD) have higher land values per unit area, because certain agents are more willing to pay higher prices for them. This fact implies that investment in transport can improve accessibility to certain locations and have repercussions on property values.

Hedonic studies stand out in the empirical literature which tries to verify hypotheses on urban economic theory. This technique has become the standard econometric tool for estimating the determinant factors on the prices of heterogeneous goods such as property values (Malpezzi, 2008). The development of hedonic studies had its roots in both early empirical studies (Court, 1939) and the reformulation of consumer theory carried out by Lancaster (1966). However, it was Rosen (1974) who finally formalized the theory of how markets worked for heterogeneous goods. According to this theory, real estate can be seen as goods priced as a function of the group of their characteristics. These characteristics may not only refer to the structural aspects of the properties but also to the characteristics of the surrounding area and their access to different land uses.

Hedonic models can help in estimating the increase in real estate prices derived from environmental and other local improvements, making them a potentially useful tool to support investment in transport projects through value capture policies. Furthermore, when integrated in Land Use/Transport Interaction Models (LUTI models) hedonic models can help simulate the complex interactions caused in an urban system where location choices for housing or companies depend very strongly on the real estate market (Löchl and Axhausen, 2010; Waddell et al., 2007b).

This article presents four hedonic regression estimators to verify, in accordance with the accepted economic theory, the hypothesis that the dwellings with better accessibility do capitalize, to a certain extent, these benefits; in other words, to verify to what extent a relationship between the accessibility conditions and the dwelling market values does exist. Accessibility is here measured by three types of indicators: an accumulated opportunities indicator, gravity-based indicators and the journey time to CBD taking into account congestion.

The application of hedonic regression techniques in research carried out within the real estate market has had various methodological problems which will be tackled throughout this study. One of the more basic problems is the existence of strong spatial relationships between observations. These relationships can violate the basic hypothesis of multiple linear regression model residual independence (LeSage and Pace, 2010). Anselin (1988) differentiates two basic types of spatial relationships. Spatial dependence or autocorrelation which is defined as the existence of a functional relationship between what occurs at a point in space and what occurs at nearby or neighbouring points, and spatial heterogeneity (or spatial non-stationarity), that is the lack of structural stability in the parameters or of spatial errors in a model. In the context of real estate markets both effects could be present due to various factors: lack of equilibrium between housing supply and demand in different sectors of an urban area (Bitter et al., 2007), diffusion effects of market prices for housing in nearby areas or, simply, the omission of relevant variables that were not included in the model because of the lack of or poor quality available data. Therefore, it would be necessary to use spatial econometric models in order to avoid biased or inefficient parameters in case studies in which these effects play a significant role (LeSage and Pace, 2009).

Throughout this study, after section 4.3 presents the state of the art of the models developed to estimate the impacts of changing accessibility on real estate, hedonic regression models will be estimated with and without considering the existence of spatial dependence between observations. In section 4.4 a Multiple Linear Regression (MLR) model is estimated using the traditional hedonic regression technique applied to a dataset from the urban area of Santander. The residuals of this model show a strong degree of spatial autocorrelation between observations. To overcome such autocorrelation, spatial autoregressive models (SAR), spatial autoregressive model in the error term (SEM) and spatial Durbin models (SDM) have been estimated along with the following questions of methodological interest:

1. do the models considering spatial dependency between observations have a significantly better fit to the data?
2. which spatial regression model combines a better fit with higher parsimony in the hedonic function?
3. which type of spatial relationship between observations provides a better fit?

The specification of the SAR, SEM and SDM models are presented in section 4.5 and the model estimates are analyzed to answer the questions posed. Finally, in section 4.6, some conclusions are drawn on the opportunity to take into account the existence of spatial effects when modelling the real estate market as well as on the opportunity to use a given modelling specification to capitalize the accessibility conditions onto the housing market prices.

4.3. Bibliographic review

Research about how transport conditions influence real estate prices can be classified into two main streams of thought. On the one hand there are the more theoretical studies initiated by Von Thünen (1826) in his work on agricultural land rents. This pioneering work later served as a basis for creating a theory about the distribution of land use and rents in urban areas proposed largely by Alonso (1964), Muth (1969) and Mills (1972a). The nucleus of the theory lies in the modelling of certain trade-offs in the choice of location, mainly between the transport costs of getting to the CBD and the cost of the space, which can be modelled using bid-rent functions. This tradition, which makes up the theoretical nucleus of the urban economy, has continued to grow through the use of ever more complex models. An excellent systematic review of the various research work carried out can be found in Fujita (1989).

On the other hand, the second main body of research, based around the relationship between transport and real estate values, is of a more empirical nature and has provided a growing number of case studies. These have been generally supported by the well-known hedonic regression technique formulised by Rosen (1974) to describe how markets function with heterogeneous goods. The hedonic studies relating real estate prices with transport conditions have therefore complemented the theories on urban economy and tested their hypotheses through multiple case studies. Most of these studies have concentrated on the relationship between real estate prices and access to rail transport with very varied results (Pagliara and Papa, 2011). Debrezion et al. (2007) carried out a meta-analysis with more than 50 hedonic studies to explain the variability in the results of the research. The authors controlled the influence of variables like the type of property being modelled, the type of station or the functional form chosen for the hedonic model. The results detected a significant influence of variables such as the type of property or station being studied on the variability of the relationship between accessibility to railway stations and real estate prices. Commercial properties generally showed higher price rises than residential properties, while suburban train stations also had a greater influence on local real estate prices than light rail or metro stations did. Nevertheless, the authors found overvaluations of positive impacts if the specification of the models omitted variables which considered the influence of other modes of transport on accessibility.

Senior (2009) found that Metrolink had no effect on house prices in Greater Manchester (Forrest et al., 1996), however Overnell (2007) was later able to identify a positive effect on the prices of properties located between 0.5 and 1 kilometre from the Metrolink stations. Andersson et al. (2010) studied the effects of accessibility to high speed rail in Taiwan and discovered that overall it had a very minor effect on property prices. This clearly contrasts with the results of Debrezion et al. (2006) in

Holland, where the effects of proximity to a railways station were more than double those found in Taiwan. Banister and Thurstain-Goodwin (2011) found that, at the micro level, non-transport benefits provided by investment in railways can be seen in the land and property markets. These effects are a reduction in land prices immediately around the railways stations due to increased noise levels and greater crime rates (Bowes and Ihlanfeldt, 2001), and these effects can radiate out up to a range of 1 kilometre (Overnell, 2007). It is important to point out that, although small investment may have local effects, it is the large investments that have really significant effects on the housing market (Buchanan and Partners, 2003).

Another, less numerous, series of studies has concentrated on the impact caused by Bus Rapid Transit (BRT) systems on real estate prices. Rodriguez and Mojica (2009) studied the impact on property values caused by introducing a BRT system in the city of Bogotá in Bolivia and found price increases of between 13% and 14%. The influence of the same BRT system was again examined by Munoz-Raskin (2010) who found that the properties nearest the bus stops had a value 4.5% lower than the rest of the properties in Bogotá. However, they also found that properties located less than 5 minutes walk away from the stops were valued 8.7% higher than those located between 5 and 10 minutes walk away concluding that households were prepared to pay more to be located close to the BRT system. Cervero and Kang (2011) used multilevel hedonic regression to estimate the capitalisation of introducing a new BRT system in Seoul, South Korea and found increases in property values of up to 10% for residents less than 300m from a stop on the network.

Nevertheless, in spite of the usefulness of hedonic studies they are not exempt from technical problems. Armstrong and Rodríguez (2006) point out three of them: the problem of omitting variables, the problem of choosing the functional form and the problem of spatial autocorrelation in sample observations. The omission of theoretically relevant variables may, as is well known, bias the estimated parameters (Gujarati and Porter, 2009) and the problems detected by Debrezion et al. (2007) in a great many models of omitting the accessibility provided by other modes of transport is a clear example of this. The problem of specifying the functional form is common to all hedonic studies. There is currently no theoretical basis which recommends using one particular functional specification rather than another, even if Cropper et al. (1988) showed that the linear form produced lower errors in cases where the model presented omitted variables. Malpezzi (2008) recommends the use of the log-linear form because it allows estimated parameters to be interpreted as semi-elasticities and has the capacity to reduce the problems derived from heterocedasticity. Finally, the problem of spatial autocorrelation in the sample may lead to the estimation of parameters which are inefficient or even biased, requiring the use of spatial econometric models (LeSage and Pace, 2009).

Within the range of available models in the field of spatial econometrics, the SAR and SEM models have enjoyed the greatest number of practical applications. These types of models were systematized for the first time in the early contribution made by Paelinck and Klaassen (1979) and received later contributions by various authors (Anselin, 2010). These types of models have received increasing attention in the field of hedonic studies applied to the relationship between transport and real estate prices. Armstrong and Rodríguez (2006) used a spatial autoregressive model to examine rising real estate values following the opening of a suburban railway in Eastern, Massachusetts, USA. The estimated model could capture the existence of spatial dependence between observations and price rises of up to 10% in the properties close to the stations. The properties located close to the lines also showed significant, but negative, changes in value. Martínez and Viegas (2009) examined the relationship between the availability of transport infrastructure and property values in Lisbon, Portugal in order to establish value capture schemes to introduce new public transport services. The authors used a MLR model and a SAR model to demonstrate the existence of spatial autocorrelation between observations. However, the MLR model showed similar parameters to those estimated by the SAR model and a lower Akaike information criteria (AIC) leading the authors to conclude that the MLR model was preferable because it offered sufficiently well-fitting predictions with greater parsimony. Löchl and Axhausen (2010) compared hedonic type MLR, SEM, SDM and based on geographically weighted regression (GWR) models. This comparison was made to establish the best specification of a hedonic regression to be introduced into a land use-transport interaction model (UrbanSim). The authors chose the SEM model as the most appropriate because the SDM model showed a large number of variables not significant and because the GWR showed a strong correlation between the estimated parameters, a phenomenon also found in other research (Ibeas et al., 2011; Wheeler and Tiefelsdorf, 2005).

4.4. Multiple Linear Regression (MLR) models

The data set used to estimate the hedonic models will be presented in this section. A second step will introduce the various specifications proposed for the MLR models along with a discussion about the parameters obtained. Finally, the presence of spatial autocorrelation in the residuals of the models will be evaluated, something that would violate one of the fundamental assumptions of the MLR models.

4.4.1. The data set

The hedonic regression models used in the present application will be estimated with data from the metropolitan area of Santander. Santander is a medium sized city, capital of the region of Cantabria in the North of Spain. The city currently has 182,700 inhabitants in its urban nucleus but the population rises to around 280,000 if the surrounding metropolitan area is taken into account. Apart from the capital, other important urban centres within the metropolitan area are Astillero (10,020 pop), Muriedas (11,279 pop) and Maliaño (5,272 pop). The city of Santander is located to the North of the study area (see Fig 4-1 and Fig 4-2) and is connected to the other urban centres by transport networks and services. These networks are mainly made up of urban and interurban road systems, the urban and interurban public transport services and the interurban railway network connecting the most important nuclei in the study area.

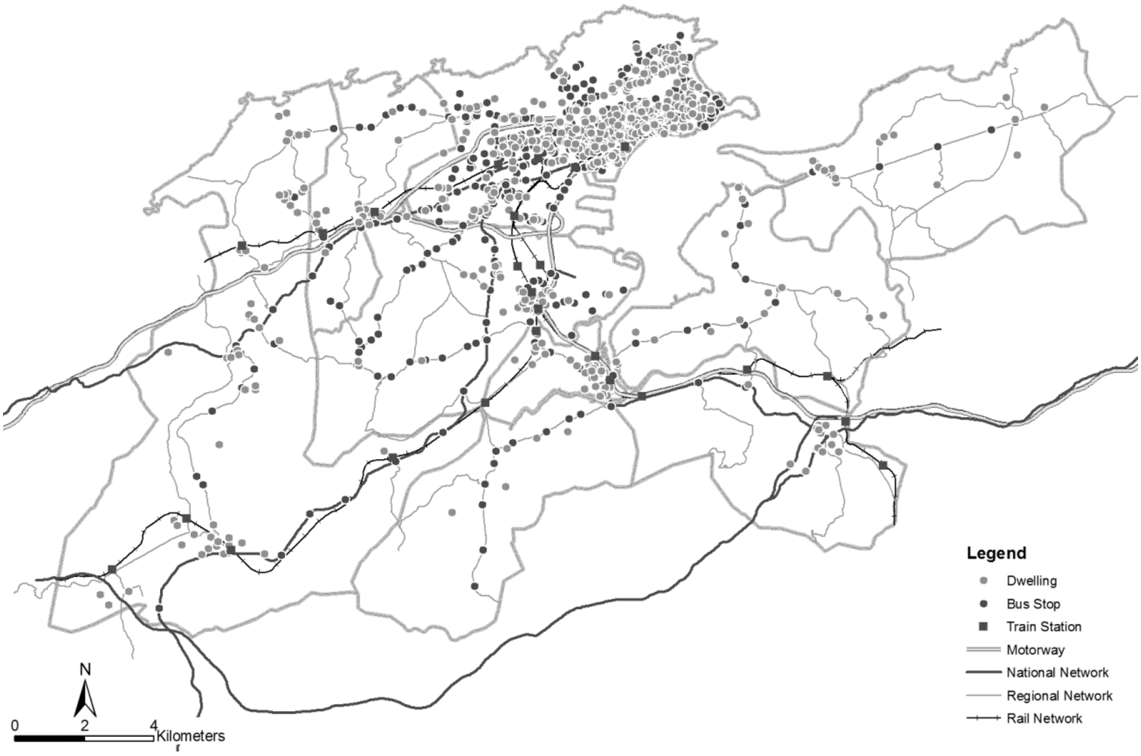

Fig 4-1. Location of sampled dwellings, bus stops and transport networks in the study area

The household sample comes from a cross sectional data base obtained from various on-line real estate platforms. The data was collected in June 2009 and contains information about asking prices and other structural characteristics for 1562 properties located in the metropolitan area. The availability of the address of each of the sample observations meant that they could be coded with a geographical information system (GIS).

The spatial distribution of the aggregated asking prices over large administrative areas shows how the highest average prices (around 500,000 euros) are concentrated in the city of Santander and more specifically in the residential area El Sardinero, located to the east of the city. This neighbourhood is characterised by a range of environmental attractions such as its status as a garden city, the prestige associated with an address there, its closeness to beaches and parks, etc. Other areas with high average prices are located along the central axis of the city of Santander (central zone) as well as in more recently developed neighbourhoods close to the el Sardinero area, which largely share the same environmental attractions (e.g. the Valdenoja neighbourhood). Two high price areas were detected away from the urban nucleus. The first of these was found to the west of Santander with a suburban style of development made up mainly of individual family houses and also close to the coast and beach areas (San Román, Liencres and Bezana). The second area is found to the east of Santander where several coastal settlements are made up mainly of second homes (Somo and Galizano, among others). The areas with the lowest average asking prices are located in a range of residential neighbourhoods found around the urban centre where the households with the lowest incomes live, as well as along the south eastern link around the Bay of Santander, where the area has been strongly influenced by negative environmental spillovers from industrial development and port activity. An overall north-south spatial gradient can be detected in the housing prices which is a function of proximity to the coast and the beaches. However, the existence of a pricing pattern depending on distance to the town centre of Santander is not so easy to identify using purely cartographic representation.

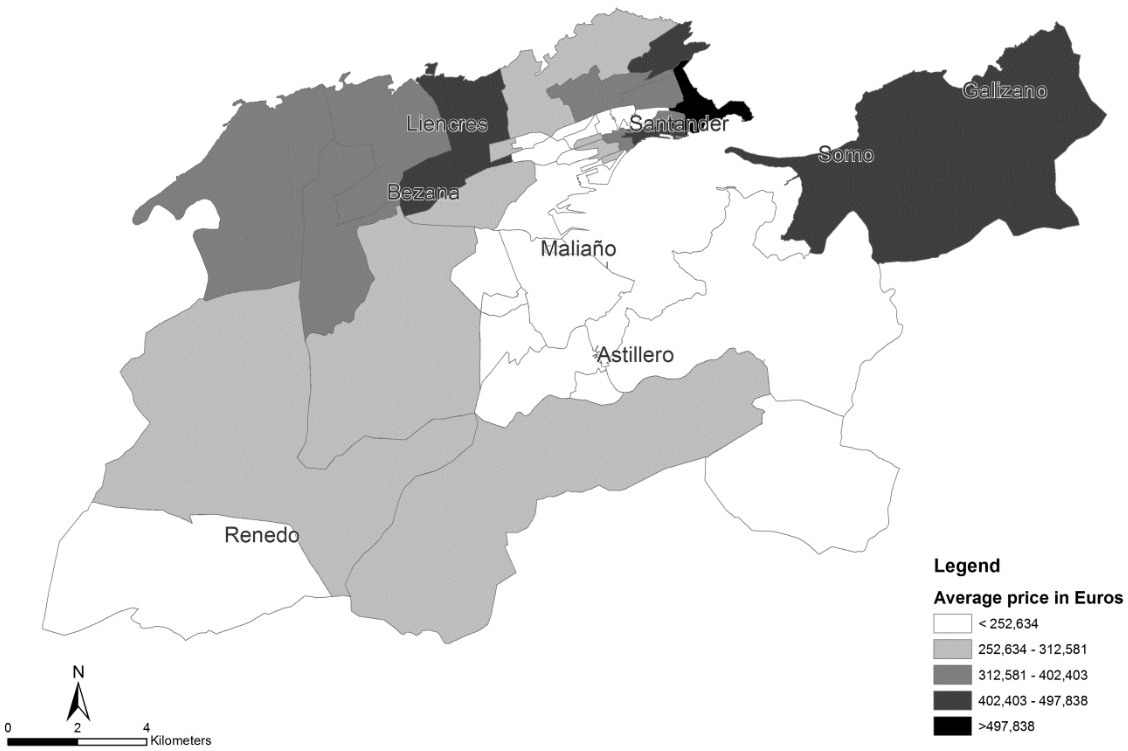

Fig 4-2. Spatial distribution of average asking prices aggregated by administrative zones in the study area

Putting all the data into the GIS meant a diverse range of variables relating to the environment and the location of the properties could be obtained when the data was crossed with the exact location of each household with various other socio-demographic characteristics present in the population and household census zoning data. The following variables are contained in the data base (see Table 4-1):

- LN(P) is the natural logarithm of the property asking price.
- IMPROV is a dummy variable taking a value of 1 if the property requires major improvement.
- DETAC is a dummy variable taking a value of 1 if the property is detached.
- ROOMS is the number of bedrooms at the property.
- BATH is the number of bathrooms at the property.
- FLOOR is the floor where the property is located in the building.
- LIFT is a dummy variable taking a value of 1 if the building where the property is located has a lift (elevator).
- TER is a dummy variable taking a value of 1 if the property has a terrace.
- GAR is a dummy variable taking a value of 1 if the property has a garage.
- SQM is the surface area of the property in square metres.
- LINES is a dummy variable taking a value of 1 if the property has a bus stop less than 400 metres away interacting with the number of lines servicing that bus stop.
- CBD is the time in minutes which it takes at morning rush hour to reach the city's CBD from the property using the road network, considering congestion.
- TRAIN is a dummy variable taking a value of 1 if the property is less than 500 metres from a suburban train station.
- ACC is a Hansen type measure (gravitational) of employment accessibility.
- CEN is a dummy variable taking a value of 1 if the property is located in the city centre.
- BCH is a dummy variable taking a value of 1 if the property is located at a beach.
- DEN is a measure of the zone's population density. Calculated as inhabitants per area.
- JOBS is the number of employments present in the area where the property is located.

- EXT is the proportion of the population from overseas living in the area where the property is located.

Variable	Minimum	Maximum	Mean	Std. Deviation	Measurement unit
LN(P)	11	14.91	12.49	.55	Ln (Price €)
IMPROV	0	1	.07	.26	-
DETAC	0	1	.23	.41	-
ROOMS	0	12	2.97	1.15	No. of bedrooms
BATH	0	4	1.86	.83	No. of bathrooms
FLOOR	0	12	2.38	1.99	Floor number
LIFT	0	1	.52	.5	-
TER	0	1	.26	.43	-
GAR	0	1	.56	.49	-
SQM	20	850	124.41	79.73	m^2
LINES	0	15	4.26	4.62	No. of lines
CBD	.1	28.73	8.22	6.42	Minutes
TRAIN	0	1	.26	.43	-
ACC	5.01	48.05	18.56	12.12	-
CEN	0	1	.07	.26	-
BCH	0	1	.17	.37	-
DEN	.006	9.41	1.31	1.94	Pop./Area
JOBS	.006	4.60	.54	.63	No. Jobs in 1000s
EXT	.008	.322	.08	.05	Proportion of foreigners

- Table 4-1. Descriptive statistics of the variables contained in the residential property data base (N=1562)

There are some problems associated with the characteristics of the data source used. The main restriction is that the property prices are not market values, they are asking prices. Nevertheless, previous research has shown that asking prices have a high correlation to selling prices and generally represent 90% of the equilibrium price (Hometrack, 2005). Therefore, changes in the dependent variable are not expected to cause significantly different parameter estimations. One of the aspects that could condition the use of this is the fact that it corresponds to a period which included the start of the housing crisis in Spain, narrowly connected to the international financial crisis. However, the price variations have been limited, especially in the urban centre of Santander, with a fall of between 10% and 15% compared with the current situation. Furthermore, changes in real estate prices are not thought to have altered the trade-offs between the characteristics of the properties being considered, meaning that the estimated parameters should not change. Additionally, the number of variables in the characteristics of the diverse properties included in the data base is limited. Unfortunately, there is no official data base currently available for public viewing in Spain showing the characteristics and final selling prices of real estate.

The dependent variable, the property asking price, has been specified in logarithmic form following the recommendations of Malpezzi (2008), meaning that the estimated parameters can be interpreted as semi-elasticities. The most relevant variables, in accordance with the aim of this study, are those which refer to transport conditions: LINES, CBD, TRAIN and ACC. The LINES variable, as described earlier, represents the interaction between the presence of a bus stop at least 400 m away and the number of lines which service that bus stop. This variable is therefore an indicator of the supply of bus services available to each of the properties. The possibility of measuring access to bus services by only using the 400 m zone was initially tested but it was discarded because of the lack of variability in the resulting dummy variable as 85% of properties had access to urban or interurban bus stops. The CBD variable represents access time in minutes to the urban centre of Santander calculated with a transport model which uses real morning rush hour traffic flows. The times input into the data base assume the existence of congestion and represent a road accessibility indicator in accordance with the theoretical mono-centric model proposed by Alonso (1964). The TRAIN variable represents an accumulated opportunities measure of accessibility to metropolitan railway services (Handy and Niemeier, 1997). Therefore, it has been assumed that all the properties located at less than 500m from one of the 25 railway stations in the study area (see Fig 4-1) have access to the train mode. Finally, ACC is a Hansen (1959) type indicator of gravitational accessibility to employment which therefore, considers the possible multi-centric nature of the urban area. This indicator has been calculated from the employment figures available in census data. The expression used was as follows (Coppola and Nuzzolo, 2011):

$$Acc(o) = \sum_i [\exp(\alpha_2 \cdot Cost(o,d_i)) \cdot jobs(d_i)^{\alpha_1}] \qquad (4.1)$$

where *Cost* is a measure of the journey cost between origin *o* and destination d_i calculated using a transport model considering congestion at morning rush hour. Jobs *(d_i)* are the current employments in the destination zone d_i and α_1, and α_2 are the parameters to be estimated. The parameters can be estimated by ordinary least squares, linearizing (4.1) using logarithms at both sides of the expression. The transport flows generated by each zone have been selected as a proxy of the accessibility conditions (dependent variable). The parameter α_1 presented an estimated value of 0.26 while α_2 had a value of -0.12, showing a R^2 goodness of fit of 0.7.

4.4.2. MLR estimates

The first hedonic model (MLR1) was estimated with the 18 independent variables contained in the data base (see Table 4-2 showing the p-values of the parameters in brackets). All the parameters had theoretically correct signs, although four of them: DETAC, ACC, DEN and JOBS were not significant according to the t test. The parameter of the DETAC variable indicated that being a detached property raises average house prices, although this effect was not high enough to be statistically significant, so it was eliminated in subsequent models. Furthermore, two of the variables which measure the influence of property accessibility, CBD and ACC had a high degree of colinearity with a correlation coefficient of -0.78 and a VIF value of 3.6 and 6.7 respectively. As the DETAC, DEN and JOBS variables were not significant they were removed from the MLR2 and MLR3 models and only one indicator of road accessibility was kept in each of them, CBD in MLR2 and ACC in MLR3. In both models, CBD and ACC were significant at a 99% confidence level. They also had the correct signs, although the MLR2 model had a slightly higher goodness of fit taking into account the R^2_{adj} as the Akaike information criteria (AIC). Both models showed moderate colinearity between independent variables and in no cases were Variance Inflation Factor (VIF) values over 3 nor condition indexes over 20 found. Finally, for the MLR4 and MLR5 models, respectively similar in their specification to MLR2 and MLR3, eight observations were removed as outliers because they showed studentized residuals[4] higher than three

[4] Given the residuals depend on the unit of measurement of the dependent variable, it would be better to standardise them by dividing them by the standard error of the regressions. This means the residuals can be compared with those of other regressions and facilitate the detection of outliers by distributing the residuals with an average of zero and a variance close to the unit. In this case it was chosen to use the studentized residuals, which are identical to the standardised residuals with the exception that the

typical deviations. In five cases, this was due to negative residuals while in the remaining three cases the residuals were positive. Among the negative-residual observations, four of them concerned detached properties with residential surfaces and number of rooms and bathrooms high above the average. This causes the model to overestimate asking prices (although they were in themselves already high). The remaining observation with a negative residual corresponds to an apartment with an asking price of 75,000 Euros which is much lower than the average even when taking into account its characteristics; this appears to be due to an error in the digitisation of the data base. Finally, the three observations with positive residuals have very high prices (in all cases above 500,000 Euros) which the model underestimates. These high asking prices are because the properties are located in tourist zones, a factor which is partly captured by the BCH variable; however in these three cases it has a much higher weight than normal.

The removal of the outliers slightly improved the fit of both models. The MLR4 and MLR5 models were therefore specified as (4.2) and (4.3) respectively:

$$\begin{aligned} Ln(\hat{P}_i) = \beta_0 + \beta_1 IMPROV_i + \beta_2 ROOMS + \beta_3 BATH_i + \beta_4 FLOOR_i + \beta_5 LIFT + \beta_6 TER_i + \beta_7 GAR + \\ \beta_8 SQM_i + \beta_9 LINES_i + \beta_{10} CDB_i + \beta_{11} TRAIN_i + + \beta_{12} CEN_i + \beta_{13} BCH_i + \beta_{14} EXT_i + \varepsilon_i \end{aligned} \tag{4.2}$$

$$\begin{aligned} Ln(\hat{P}_i) = \beta_0 + \beta_1 IMPROV_i + \beta_2 ROOMS + \beta_3 BATH_i + \beta_4 FLOOR_i + \beta_5 LIFT + \beta_6 TER_i + \beta_7 GAR + \\ \beta_8 SQM_i + \beta_9 LINES_i + \beta_{10} ACC_i + \beta_{11} TRAIN_i + + \beta_{12} CEN_i + \beta_{13} BCH_i + \beta_{14} EXT_i + \varepsilon_i \end{aligned} \tag{4.3}$$

standard error of the i*th* residual is calculated by eliminating this observation. This guarantees that the variance of the distribution of the residuals is truly unitary (Gujarati, 2009).

Variables	MLR1	MLR2	MLR3	MLR4	MLR5
(Constant)	11.367	11.393	11.227	11.402	11.240
	(.000)	(.000)	(.000)	(.000)	(.000)
IMPROV	-.085	-.080	-.087	-.079	-.086
	(.002)	(.004)	(.002)	(.004)	(.002)
DETAC	.038				
	(.167)				
ROOMS	.069	.068	.068	.058	.058
	(.000)	(.000)	(.000)	(.000)	(.000)
BATH	.131	.134	.136	.123	.125
	(.000)	(.000)	(.000)	(.000)	(.000)
FLOOR	.011	.010	.012	.009	.011
	(.009)	(.012)	(.003)	(.019)	(.005)
LIFT	.244	.232	.253	.235	.256
	(.000)	(.000)	(.000)	(.000)	(.000)
TER	.046	.045	.039	.042	.036
	(.004)	(.005)	(.016)	(.007)	(.021)
GAR	.107	.112	.108	.109	.105
	(.000)	(.000)	(.000)	(.000)	(.000)
SQM	.003	.004	.004	.004	.004
	(.000)	(.000)	(.000)	(.000)	(.000)
LINES	.020	.021	.021	.021	.022
	(.000)	(.000)	(.000)	(.000)	(.000)
CBD	-.010	-.011		-.011	
	(.000)	(.000)		(.000)	
TRAIN	-.060	-.062	-.076	-.065	-.075
	(.001)	(.000)	(.000)	(.000)	(.000)
ACC	.002		.005		.005
	(.269)		(.000)		(.000)
CEN	.123	.145	.088	.140	.084
	(.001)	(.000)	(.009)	(.000)	(.011)
BCH	.338	.336	.315	.326	.305
	(.000)	(.000)	(.000)	(.000)	(.000)
DEN	-.008				
	(.130)				
JOBS	.002				
	(.903)				
EXT	-.570	-.638	-.875	-.631	-.863
	(.001)	(.000)	(.000)	(.000)	(.000)
R^2	.767	.767	.763	.773	.770
R^2_{adj}	.765	.764	.761	.771	.768
F	282.67	362.92	355.54	375.13	367.30
p - value F	.000	.000	.000	.000	.000
p - value Moran's I	.000	.000	.000	.000	.000
AIC	334.75	331.56	356.10	252.48	277.75
Log-Likelihood	-148.37	-150.78	-163.51	-111.24	-123.87
N	1562	1562	1562	1554	1554

Table 4-2. Estimated parameters of the MLR models

Given the semi-logarithmic specification of the models, the parameters can be directly interpreted as semi-elasticities which allows the dependent variable's relative change caused by an absolute change in the value of an independent variable to be estimated ceteris paribus. This contrasts with log-log regressions where each parameter measures the independent variable elasticity, i.e. the dependent variable's relative change related to a relative change in the value of an independent variable (Gujarati and Porter, 2009). However, in the case of the dummy variables, the parameters cannot be directly interpreted in this way. These can be correctly calculated by applying the following expression (Halvorsen and Palmquist, 1980):

$$\left[\exp(\beta_n)-1\right]*100 \qquad (4.4)$$

In the MLR4 and MLR5 models all the variables that were introduced were statistically significant and had theoretically believable signs (see Table 4-2). Among the variables related to the property's structural conditions, IMPROV was the only one that gave a negative sign which is interpreted as a reduction in property value of between 7.5 and 8.2 % if the building had to be renovated. Noteworthy from among the variables which had a positive effect on property values, the presence of an additional bathroom, which increased prices by 12%, having a garage which increased prices by around 10% and, most strikingly, if the building was equipped with a lift it implied an increase in value of 20%. The important parameter related to this last variable may not only be due to the availability of a lift but also the age of the building because more modern and better quality construction normally includes a lift.

The parameters related to the environmental conditions of the properties showed positive signs if the buildings were located in the city centre or, even more so, in the beach areas, a variable which could also include the effects of better landscaping. A property's location in a beach area could imply an average increase of 38% in its value according to the parameter of the MLR4 model. The parameter of the EXT variable was negative and high which could also be due to the capture of other environmental effects by this variable such as population density (even though DEN did not turn out significant), worse urban services and a lower presence of public installations.

Finally, most of the parameters related to the transport conditions had signs which agreed with the hypothesis that improved transport conditions resulted in increased real estate values. The exception was the TRAIN variable which had a negative sign implying that properties with a train station less than 500 metres away had a value which was between 6% and 7% lower. This result which goes against established theory has also been found in other previous studies (Forrest et al., 1996). But, unlike in the case of Metrolink, where the stations and most of the nearby housing were built in the 19^{th} and early 20^{th} centuries (making them less desirable and with lower prices), the reason is unclear and may be because railway installations and infrastructure carry with them a series of negative spillovers associated with noise and a generally lower landscaping and environmental quality (Armstrong and Rodríguez, 2006). We must also consider the fact that the railway constitutes a minimum percentage of the modal split (around 1%) which helps to explain why the presence of a railway station does not have a positive impact on property prices.

The parameter of the bus accessibility indicator LINES indicates that each additional nearby public transport line can imply an additional increase of up to 2.2% in property values. However, this result should be interpreted with care because the city centre areas of Santander are the ones that are best served by public transport. In this sense, the hypothesis that a better supply of public transport services implies increased property values could be reversed and be due simply to the fact that the more central areas have more public transport services due to the demand for travelling to that area from all the other parts of the city. This phenomenon is very typical of European cities where long standing historic urban centres have always had a high residential and commercial presence and high real estate values (Felsenstein et al., 2010).

The road network accessibility indicators produced believable signs, negative in the case of CBD in the MLR4 model and positive in the case of ACC in the MLR5 model. The CBD parameter indicates that an additional minute in travelling time to the urban centre of Santander could imply slightly over a 1% reduction in property values which confirms the existence of the price gradient assumed by urban economic theory. The parameter of the ACC variable indicates that an increase of one unit in the employment accessibility index implies an increase of 0.5% in real estate values.

For both the MLR4 and the MLR5 models the parameters of the CBD and ACC variables had the correct signs and were significant, meaning that both these indicators of accessibility to employment opportunities and commercial activities can be considered valid. However, a variable like CBD can be adapted to most traditional mono-centric urban areas while the ACC variable can be considered as a more valid indicator for polycentric zones (Ottensmann et al., 2008). In the case of this research, the nuclei of the alternative urban centres to Santander within the metropolitan area do not as yet show enough development in the number of job opportunities available to be able to compete with the city centre of Santander which has more than 20% of the total jobs and around 65% if the city as a whole is considered. It can therefore be stated that the metropolitan area still presents a strong mono-centric character meaning that the CBD variable can be considered as a strong indicator for measuring accessibility. The MLR4 model also had a slightly better fit in both its R^2_{adj} and AIC and will therefore be chosen as the reference model for specifying the models which consider the existence of spatial dependence.

4.4.3. Autocorrelation analysis

The well-known Moran´s I index (Griffith, 1987) was used to test for the presence of residual autocorrelation. Before the index was calculated, the geographical point information on property location was transformed into zonal information using Thiessen polygons (Maguire et al., 2005). The index was initially applied with Queen type spatial contiguity (Griffith, 1987) and later at different fixed distances and clearly significant values appeared in all cases (see Table 4-2).

A Getis-Ord Gi* statistic was also calculated (see Fig 4-3 for the residuals of the MLR4 model). This index is useful for detecting clusters of autocorrelated residuals which may not become evident using only a global autocorrelation index (Ord and Getis, 1995). The Gi index showed significant values at a 95% confidence level for the presence in various zones of both positive and negative residual clusters. The zones showing significant correlation in the positive residuals are located in the northern part of the study area, specifically in the zones close to the Bay of Santander where there is a strong demand for housing due to the attractive characteristics of the area. Of particular relevance was the case of the El Sardinero sector where properties had very high prices which were almost certainly due to the lack of supply and the high demand, apart from the factors mentioned above. The negative residuals were spatially auto-correlated in certain nuclei to the south of the Bay and to the west of the study area, zones which currently have a high number of commuting trips to the city centre.

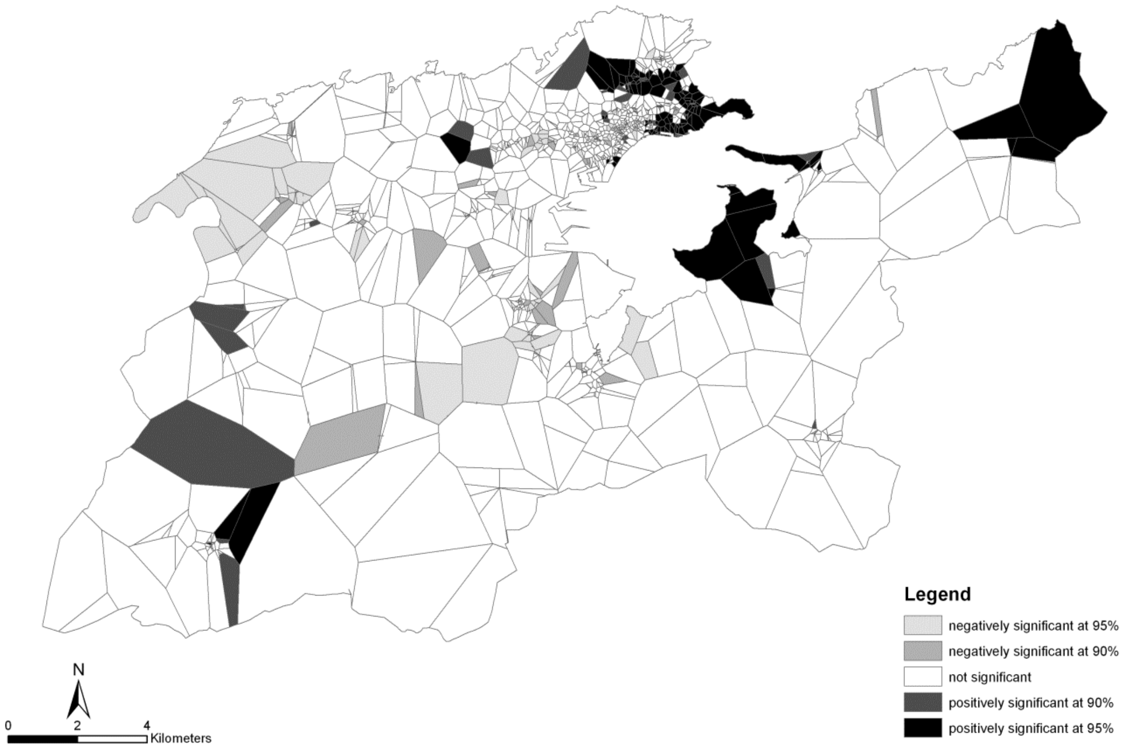

Fig 4-3. Significance of the Getis-Ord Gi* statistic values on the residuals of the MLR4 model

Anselin (1988) recommends using the Lagrange multiplier test (LM) to detect specification errors due to not considering spatial dependence in MLR models. This test can detect specification errors caused by not including the autoregressive parameter in the dependent variable (LM-Lag) or in the error term (LM-Error). These tests can also provide robust versions if both specification errors are significant. In this case study they prove significant values both in their LM-Lag and LM-Error versions. The robust tests were also significant except in the case of the spatial dependence of the dependent variable using a nearest neighbour distance matrix (see the following section).

4.5. Spatial econometric models: SAR, SEM and SDM

To capture the effect of the strong spatial autocorrelation present in the residuals of the MLR models, a further series of models were estimated considering the existence of spatial dependence between observations. Both of the functional forms more commonly found in the literature, SAR and SEM, and the spatial Durbin model (SDM) were applied. The models considering spatial relationships were estimated using the specification of the MLR4 model as this gave a better fit and the measure of road accessibility from the journey time to the CBD was considered to be a correct hypothesis.

4.5.1. SAR, SEM and SDM specifications

LeSage and Pace (2009) provided an extensive introduction to the spatial econometric models developed in the literature. The most well-known spatial model is the simultaneous autoregressive (SAR) model which assumes the existence of a diffusion process in the dependent variable and can be specified as follows:

$$y = \rho Wy + X\beta + \varepsilon \quad (4.5)$$

where y is a vector of observations for the dependent variable, ρ is the parameter of spatial autocorrelation, W is a matrix $N \times N$ of spatial weightings where N *is* the number of observations, β is a vector of estimated parameters, X is a matrix with observations for the independent variables and ε is a vector of independent and identically distributed (IID) error terms.

If we only wish to specify the existence of spatial dependence in the error term of the observations, then an autoregressive simultaneous spatial error model (SEM) can be used. This type of model can be written as:

$$y = X\beta + u \quad (4.6)$$

$$u = \lambda Wu + \varepsilon \quad (4.7)$$

where W is a $N \times N$ matrix of spatial weightings, λ is a spatial autocorrelation parameter of errors u and ε is a vector of independent and identically distributed errors. So in this model, the dependent variable of any location is a function not only of the independent variables but also of the errors *u* of the neighbouring locations.

Finally, although applied in fewer research works, a third type of model found in the literature is the spatial Durbin model (SDM):

$$y = \rho Wy + X\beta + \gamma WX + \varepsilon \qquad (4.8)$$

where W is a *N x N* matrix of spatial weightings, ρ and γ are spatial autocorrelation parameters of the dependent and independent variables, respectively and ε is a vector of independent and identically distributed errors. The application of this model is especially recommended by LeSage and Pace (2009) because of two of its properties. In the first place its greater robustness against the omission of relevant independent variables. Secondly, given its more general model characteristics than those of SAR and SEM which may be considered as nested versions of SDM. So the addition of the constraint $\gamma = 0$ to the SDM model leads to the SAR model, while the constraint $\gamma = -\rho\beta$ leads to a SEM model and, finally, the constraints $\rho = 0$ and $\gamma = 0$ produce a standard MLR model. Furthermore, the SDM model is the only one that produces unbiased parameters even when the true data generation processes are provided by SAR, SEM, SDM models or even by a general spatial model which includes dependence in both the dependent variable and the error terms.

The *W* spatial weighting matrices present in these models define the connectivity between the units of analysis. It is important to correctly specify the matrix elements, w_{ij}, to ensure that spatial econometric models are correctly applied. The four most common ways of defining connectivity are: Queen type contiguity, Rook type contiguity, fixed number of closest neighbours and neighbours located at a maximum determined distance. Queen type and Rook type are two different contiguity definitions coming from the game of chess. Thus, Queen type contiguity considers all adjacent locations sharing an edge or a vertex with a given location as neighbours, while Rook type contiguity considers neighbours to be only those adjacent locations sharing an edge with a given location (Anselin, 1988).

LeSage and Pace (2009) argued that the estimated parameters should not be excessively variable under changes in the definition of neighbourhood matrices. This aspect will be verified in the following application. Finally, spatially dependent model parameters should be estimated using maximum likelihood because estimation using ordinary least squares could lead both to inconsistent parameters and standard errors.

Although in spatial econometric models the estimated parameters are directly interpreted using SEM type models, the same does not occur in the cases of SAR and SDM models which consider lags in the dependent variable. Simultaneous feedback exists with these specifications because a change in the dependent variable of an observation causes changes in the neighbouring observations which in turn have repercussions on the first observation. Therefore, in the cases of the SAR and SDM specifications, the estimated parameters should be seen as the representation of a state of equilibrium in the modelling process which includes the effects of spatial diffusion (Ward and Gleditsch, 2008). Given that the effects provided by each variable take the form of a matrix in this situation, LeSage and Pace (2009) recommend the use of a series of scaling indicators to correctly interpret the functional relationships:

a. Average direct effect: calculated as the mean of the elements of the main diagonal of the parameter matrix. It can be interpreted as the effects caused by the group of observations of an independent variable on the dependent variable.
b. Average indirect effect: calculated as the mean of the elements outside the main diagonal of the parameter matrix. It can be interpreted as the diffusion effects between observations caused by changes in an independent variable.
c. Average total effect: calculated as the mean of the elements of the parameter matrix of observations. It can be interpreted as the total effect, direct and indirect, received by the dependent variable.
d. These measures can be calculated separately and require the use of simulation techniques if inferences are to be made of their significance.

4.5.2. SAR, SEM and SDM estimates

The models were estimated with three types of neighbourhood matrices once again starting from the transformation of the point information to zonal information using

Thiessen polygons. An initial matrix assumes first order Queen type spatial relationship between observations, that is considered that there is contiguity between observations only when they have common boundaries. The second of the proposed matrices (referred as “K10”, in Table 4-2 and Table 4-3) considers the 10 closest observations to each property. The third and final matrix (referred as “D1750”, in Table 4-2 and Table 4-3) starts from a distance of 1750 metres and considers all the observations found within this radius to be neighbours of the property. This distance was chosen because it was the minimum in which all the observations had at least one neighbouring observation. The three matrices gave a progressively increasing number of average links: 5.7 for the Queen type contiguity matrix, 10 for the matrix with k close neighbours and 219 using the distance matrix. The three matrices progressively consider wider neighbourhood relationships, which should prove useful when checking the hypothesis that the different neighbourhood matrices do not affect the estimated parameters to any large degree.

A total of nine models were estimated considering spatial dependency between observations, combining the different functional specifications SAR, SEM and SDM with three types of neighbourhood matrices. Although the interpretation of the estimated parameters is direct in the case of the SEM type models, the same does not occur with the SAR and SDM models because these consider lags in the dependent variable. In this latter case the scaling measures proposed by Le Sage and Pace (2009) should be used to obtain the authentic magnitudes of the direct and indirect effects of the independent variables.

The parameters estimated using the SAR and SEM models in all cases showed identical signs to those obtained using ordinary least squares in the MLR4 model. Only the variables TRAIN and FLOOR, the latter only in the SEM with the D1750 neighbourhood specification, were not significant at a 95% confidence level.

Conversely, in the SDM models there was a greater presence of not significant parameters, especially for the theoretically interesting variables CBD and TRAIN in the SDM-QUEEN and SDM-D1750 models. With the SDM-K10 model all the regressors were significant and showed identical signs to those present in the SAR, SEM and MLR4 models. The spatial lag of the independent variables turned out to be not significant in many cases and in others had signs which were theoretically difficult to interpret such as the positive signs for IMPROV or the negative signs for GAR and SQM using SDM-QUEEN.

The parameters estimated for the different neighbourhood matrices (see Table 4-3) did not show any great changes with respect to the MRL models and in no case did they show sign changes except in the CBD and TRAIN variables between, on the one hand, the SDM-D1750 model and the SDM-QUEEN and SDM-K10 models. Nevertheless, these latter parameters showed little significance in the SDM models, meaning that their change of sign did not imply that the parameter estimations were clearly different. Moreover, in the three types of models the fit provided by the specifications considering Queen-Contiguity were superior to those provided by the 10 closest neighbours (K10) and maximum distance matrices (D1750). This statement is valid using both the log-likelihood and the AIC indexes.

The Moran´s I index was used to test for the presence of residual auto-correlation in the spatial models. Only four models did not show any auto-correlation that was not significant: SEM-QUEEN, SEM-D1750, SDM-QUEEN and SDM-D1750. In all the other cases the spatial residual auto-correlation was significant even though the index showed values that were clearly inferior to those presented by the MLR models.

Taking into account these factors, the models considered as being the best for each type were chosen. Three selection criteria were used: Firstly, the goodness of fit of the models with the sample data; secondly, considering the significance and coherence of the parameters present in the models with the initial hypotheses derived from geographic and economic theory, if and when the unexpected signs were not clearly significant; finally, it was considered important that the residuals derived from the models should not show any significant degree of spatial autocorrelation in order to maintain parameter unbiassedness and efficiency.

	SAR			SEM			SDM		
	QUEEN	K10	D1750	QUEEN	K10	D1750	QUEEN	K10	D1750
(Constant)	8.507	8.520	6.313	11.443	11.430	11.469	6.607	7.643	4.812
IMPROV	-.098	-.095	-.082	-.093	-.090	-.085	-.092	-.096	-.089
ROOMS	.061	.054	.059	.060	.054	.057	.059	.051	.058
BATH	.109	.120	.117	.118	.126	.123	.114	.126	.118
FLOOR	.011	.009	.008	.009	.010	.006*	.010	.009	.006*
LIFT	.212	.228	.242	.196	.225	.232	.193	.221	.229
TER	.035	.038	.034	.034	.040	.036	.033	.037	.032
GAR	.104	.110	.113	.114	.108	.127	.112	.108	.135
SQM	.003	.003	.003	.003	.003	.003	.003	.003	.003
LINES	.017	.020	.022	.018	.020	.015	.004*	.018	.014
CBD	-.009	-.007	-.007	-.011	-.009	-.010	-.008*	-.005	.001*
TRAIN	-.052	-.052	-.017*	-.051	-.047	-.006*	-.000*	-.028	.013*
CEN	.094	.140	.104	.129	.133	.076	.065*	.131	.086
BCH	.223	.245	.205	.295	.264	.227	.084*	.195	.269
EXT	-.288	-.609	-.547	-.505	-.710	-1.00	-.213*	-.683	-1.12
Lag.IMPROV							.118	.144*	.134*
Lag.ROOMS							-.047	.004*	.061*
Lag.BATH							-.031*	-.072*	.102*
Lag.FLOOR							-.006*	-.000*	-.031*
Lag.LIFT							.037*	-.015*	.049*
Lag.TER							.034*	-.020*	-.051*
Lag.GAR							-.077	.054*	-.031*
Lag.SQM							-.001	-.000*	-.001*
Lag.LINES							.009*	-.000*	.015*
Lag.CBD							.005*	-.083*	-.001*
Lag.TRAIN							-.057*	-.083	-.065*
Lag.CEN							.005*	.014*	.890
Lag.BCH							.099*	.090*	-.163
Lag.EXT							-.224*	.935	.963
ρ	.235	.232	.406				.415	.311	.568
λ				.447	.429	.756			
p - value ρ/λ	.000	.000	.000	.000	.000	.000	.000	.000	.000
p - value Moran´s I	.000	.000	.000	.060	.000	.791	.068	.000	.631
Log-Likelihood	-49.66	-63.91	-75.01	-22.51	-74.33	-34.80	7.99	-42.42	-2.06
AIC	133.33	161.82	184.04	79.02	182.67	103.62	46.02	146.85	66.13

*Not significant at 0.05

Table 4-3. Estimated parameters for the SAR, SEM and SDM models

From among the SAR and SEM models, the best were thought to be those estimated using QUEEN type neighbourhood matrices because of their better fit. However, from among the SDM models, the SDM-K10 model was thought to be better than the SDM-QUEEN and SDM-D1750 in spite of its inferior fit, because it was the only model in which the most important variables in this study: LINES, CBD and TRAIN were significant in all cases.

Table 4-4 shows the average and total direct and indirect impacts of the auto regressive models on the dependent variable. Simulation based on 100 samples was used to estimate the inference measures. The number of samples obtained had to be reduced to 100 because of the amount of calculation involved in simulating models with a great number of parameters. A comparison between Table 4-2 and Table 4-3 shows that the estimated parameters experienced slight variations. Furthermore, in cases like that of the parameter of the BCH variable, situations could occur where the direct or indirect effects of a variable can change its significance. The SAR-QUEEN model also showed indirect impacts in all the significant cases with theoretically believable signs.

Variables	SAR-QUEEN			SDM-K10		
	Average direct Impact	Average indirect Impact	Total average impact	Average direct Impact	Average indirect Impact	Total average impact
IMPROV	-.099	-.029	-.128	-.088	.158*	.070*
ROOMS	.062	.018	.080	.053	.029*	.082*
BATH	.110	.032	.143	.123	-.046*	.077*
FLOOR	.011	.003	.015	.009	.003*	.012*
LIFT	.214	.063	.277	.225	.074*	.300
TER	.036	.010	.046	.036	-.012*	.024*
GAR	.105	.030	.136	.113	.122*	.236
SQM	.003	.001	.004	.003	.000*	.004
LINES	.017	.005	.022	.018	.003*	.022
CBD	-.009	-.002	-.012	-.005	-.001*	-.007*
TRAIN	-.052	-.015	-.068	-.033	-.127	-.161
CEN	.095	.028	.124	.135	.076*	.212
BCH	.225	.066	.292	.204	.210	.414
EXT	-.291	-.085	-.376	-.637	1.00	.365*

*Not significant at 0.05

Table 4-4. Average direct, indirect and total impacts estimated for the SAR-QUEEN and SDM-K10 models

Finally, the best model was chosen from the three models selected above. The SEM-QUEEN was chosen because it provided a better fit, theoretically coherent signs (except in the case of the TRAIN variable, mentioned above) and no significant spatial auto correlation in the residuals at a 95% confidence level. The Getis-Ord Gi* statistic was calculated for the residuals of this model (see Fig 4-4). A visual comparison between Fig 4-3 and Fig 4-4 shows that the SEM-QUEEN model notably reduced residual auto correlation in various zones even though it continued to be high in the South East zone of Santander.

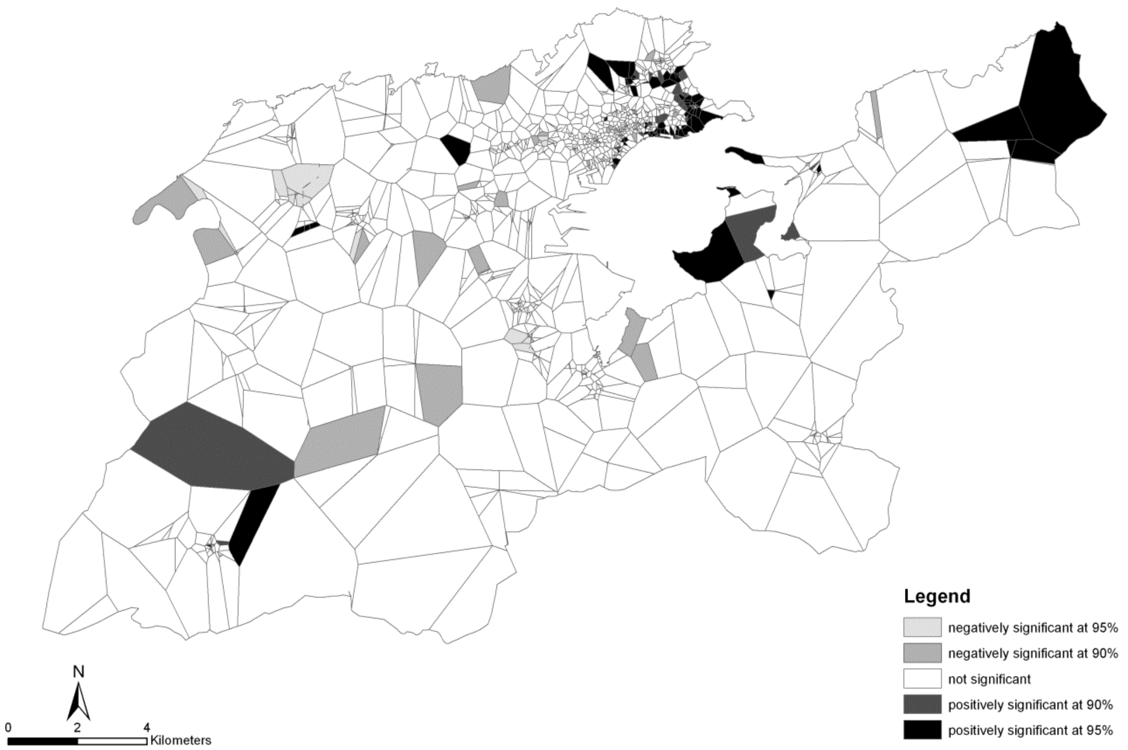

Fig 4-4. Significance of the Getis-Ord Gi*statistic values on the residuals of the SEM-QUEEN model

4.6. Conclusions

The work described in this article specified four types of hedonic models: multiple linear regression, spatial autoregressive, spatial autoregressive model in the error term and spatial autoregressive Durbin in order to determine the influence of transport conditions on the prices of a series of real estate properties. These models were estimated using data collected in the metropolitan area of the city of Santander and were compared to determine which of them presented the best fit and how the estimated parameters were affected by setting different neighbourhood relationships between the observations.

The estimated models were useful in calculating how different transport characteristics affected the prices of real estate properties. The LINES variable, as a measure of accessibility to bus transport, had a positive sign in all cases with estimations of its influence on property prices of between 1.4% and 2.2% depending on the model and 1.8% according to SEM-QUEEN. However, this result should be interpreted with care because the causality direction, increase in supply of bus services —higher property prices is arguable for the study area used in this research. The city centre of Santander has a large supply of available public transport which is mainly geared to providing access to the city centre from the surrounding areas. The CBD variable was used as an indicator of journey time by car to the city centre and in all the models the parameter had a negative sign with estimations of property price reductions of between -0.5 and -1.1% per minute of journey time. The specifications that measured accessibility through the CBD variable were chosen rather than through a gravity type indicator like ACC because the models gave a better fit and the study area could be characterised as very mono-centric. The TRAIN variable indicated accessibility using suburban railway transport and presented a negative parameter, which, in agreement with previous research (Armstrong and Rodríguez, 2006), was almost certainly due to noise and other spillovers associated with this type of infrastructure. The impact of the TRAIN parameter on real estate values was -4.9% in the SAR-QUEEN model and oscillated between -2.7% and -6% depending on the specification. These results were not far from those obtained by Forrest et al. (1996), who got an estimation of -4.5% in the average property prices placed less than 1 kilometre from a railway station in Greater Manchester. In addition, the models confirmed the existence of a price gradient as a function of accessibility to the CBD. However, the effect that accessibility to public transport has on property prices was more debatable in spite of the fact that the zones with the most bus lines, especially the centre of Santander, had higher average house prices.

The comparison between the MLR model and the SAR, SEM and SDM models which considered spatial dependency between observations allowed to verify the hypothesis put forward at the start of the study. The spatial econometric models were specified using different neighbourhood matrices, making it possible to check if they had an important influence on the estimated parameters. In general, the estimated parameters were found not to show any important variations or sign changes. Only in the case of the SDM model some parameters, bordering on 0 and not significant, change their sign. These results show that in the case of this research, the different neighbourhood matrices did not have any notable influence on the results obtained.

The SEM-QUEEN specification was chosen as the best model because it gave a better fit and also had parameters that were clearly significant with theoretically consistent signs. The good performance of the SEM models indicates that the MRL models had specification problems by omitting variables related to the local environment resulting in their overvaluation of properties in certain zones and undervaluation in others. The models which considered spatial dependence between observations, particularly the SEM models, helped to reduce this effect with better fits than the MLR models using both the log-likelihood and the AIC. Furthermore, the spatial models reduced the presence of residual spatial auto-correlation according to the Moran's I index and in some cases this did not turn out to be significant. Nevertheless, in spite of their better fit, the SDM models which consider the existence of spatial lags in the independent variables, in some cases had theoretically inconsistent signs and non-significant parameters mainly in the lag variables but also in some key variables like CBD or TRAIN. For this reason they were considered to be less valid for the purposes of this research.

Capítulo 5
CONCLUSIONES FINALES Y LÍNEAS DE INVESTIGACIÓN FUTURAS

5. CONCLUSIONES FINALES Y LÍNEAS DE INVESTIGACIÓN FUTURAS

5.1. Conclusiones

En esta tesis se han presentado una serie de modelos útiles para el estudio de los impactos provocados por los cambios del sistema de transporte en los usos del suelo. En estos modelos se ha considerado específicamente el rol de la influencia de las condiciones de accesibilidad en los patrones de localización de los hogares y empresas así como en los precios de los bienes inmobiliarios. La finalidad de estos modelos ha sido la de generar conocimiento sobre el funcionamiento de los sistemas urbanos para permitir una mejor evaluación ex – ante, de acuerdo a los criterios de sostenibilidad, de los proyectos y políticas de transporte que puedan querer implantarse. Las presentes conclusiones trataran cuatro temas principales:

- El rol de la influencia del transporte en los patrones de localización de hogares y empresas y en los precios de los bienes inmobiliarios.
- El grado de influencia de otras variables diferentes al transporte a la hora de condicionar los patrones de localización y los precios inmobiliarios.

- La importancia de considerar la componente espacial y las relaciones de dependencia entre observaciones en los modelos utilizados.
- Los aspectos más relevantes del empleo de modelos econométrico – espaciales en el ámbito práctico de la evaluación de políticas y proyectos de inversión en transporte.

5.1.1. Conclusiones sobre los impactos en los patrones de localización y los precios de bienes inmobiliarios

Los resultados aportados por los modelos permiten contrastar varias hipótesis derivadas de la economía urbana y del transporte. En primer lugar el hecho de que la accesibilidad sigue siendo un factor relevante en las decisiones de localización de hogares y empresas, algo cuestionado por algunos autores (Giuliano, 1995). Los modelos de localización residencial estimados en los apartados 2 y 3 mostraron como la variable tiempo de viaje casa – trabajo, una medida de accesibilidad relativa desagregada (Geurs y van Wee, 2004), fue claramente significativa en todos los casos en línea con otros estudios anteriores (Bhat y Guo, 2004; Guo y Bhat, 2001). El parámetro de desutilidad estimado osciló entre -0.064 y -0.106 según la especificación y la muestra empleada. La estimación separada del parámetro para los hogares de ingresos superiores a los 2500€ fue algo mayor en el caso del modelo de localización del apartado 2 mostrando una superior desutilidad (-0.131). Sin embargo en el modelo de localización residencial del apartado 3, la estimación del parámetro separado para los hogares de altos ingresos mediante la interacción con una variable dummy no fue significativa. Este hecho indica que en el área de estudio no hay una diferencia estadísticamente significativa entre los valores del tiempo de los hogares con menores y mayores ingresos.

Los indicadores de accesibilidad agregados basados en el potencial de oportunidades de empleo no fueron significativos ni en el modelo de localización del apartado 2 ni en el modelo del apartado 3. Esto pudo deberse por un lado a la captura de parte de su efecto por el indicador de accesibilidad tiempo casa – trabajo. Otra hipótesis ad – hoc que se podría plantear es la existencia de distintos patrones de localización entre los hogares de distintos ingresos, una tesis que podría sustentarse en los diferentes signos de los parámetros para los hogares de ingresos altos y medios/bajos obtenidos en los modelos del apartado 3. De esta forma los hogares de ingresos altos podrían localizarse de forma segregada a las grandes áreas de empleo pero cercanas a zonas de empleo más selectivas, mientras que los hogares de ingresos medios y bajos se situarían más cerca de las grandes zonas de empleo. Ambos factores explicarían el hecho de que los hogares del área de estudio siendo sensibles a un indicador de accesibilidad relativo y desagregado del tipo tiempo casa – trabajo no lo sean a un tipo de indicador del potencial de oportunidades de empleo. Sin embargo este hecho no modifica la afirmación general de que, efectivamente, la accesibilidad a empleos sigue siendo un factor relevante a la hora de explicar la localización de los hogares. Esta afirmación se ve reforzada además por la variable dummy cuyo valor toma el valor uno en caso de que las zonas de residencia y trabajo coincidan la cual presentó en todos los modelos del apartado 3 un parámetro positivo si bien con significatividad variable según la especificación.

El modelo de localización de actividades económicas dependientes de la demanda interna al área de estudio, mostró ser sensible a la accesibilidad pasiva a la población. La estimación desarrollada en el apartado 2 presento un parámetro en la función de utilidad de 0.643 claramente significativo por lo que puede decirse que, en coherencia con la teoría de la economía urbana, las actividades de servicios y comercio detallista tienen a preferir aquellas localizaciones cercanas a la demanda.

En cuanto a la influencia de las condiciones transporte en los precios inmobiliarios, los resultados de los modelos sustentan la hipótesis de que los beneficios obtenidos de las mejoras en las condiciones de accesibilidad son capitalizados en los bienes inmobiliarios. En el modelo hedónico estimado en el apartado 2, cada minuto adicional de acceso al centro urbano supuso un descenso del 1.8% en el valor de las viviendas de la muestra. Los indicadores de accesibilidad al transporte público también presentaron los signos esperados y fueron significativos especialmente en lo referente al indicador de oportunidades acumuladas según el número de líneas de transporte público presente en las cercanías de los inmuebles de una zona (3.6% de incremento en el valores inmobiliarios de las viviendas de una zona por cada línea de transporte público adicional en las proximidades). Los resultados de los modelos hedónicos estimados en el apartado 4 fueron similares a los anteriores. El indicador de accesibilidad al centro urbano presentó estimaciones de reducciones en los precios inmobiliarios por minuto de viaje adicional de entre el -0.5% y el -1.1%. La accesibilidad a las oportunidades acumuladas de transporte público presentó una influencia en los precios inmobiliarios de entre el 1.4% y el 2.2% dependiendo de la especificación utilizada.

Dos resultados aportados por los modelos hedónicos requieren comentarios adicionales. En primer lugar el indicador de potencial de accesibilidad a empleos resulto ser muy colineal con el indicador de accesibilidad al centro urbano por lo que no pudieron estimarse ambos en los modelos del apartado 4. Esto pudo deberse al todavía notable carácter monocéntrico del área de estudio (el centro urbano de Santander concentra en torno al 20% del empleo y la totalidad del municipio en torno al 65%). En segundo lugar el indicador de accesibilidad a las estaciones de tren suburbano presentó un parámetro con signo negativo con estimaciones entre -2.7% y -6%. Otros autores (Armstrong y Rodríguez, 2006; Forrest et al., 1996) han obtenido resultados similares en diferentes áreas de estudio y han aportado evidencia de que esta disminución en los valores inmobiliarios se debe al ruido y otras externalidades negativas asociadas a la infraestructura ferroviaria por lo que este hecho no debe considerarse excepcional.

Los resultados aportados apoyan por lo tanto las hipótesis teóricas que afirman la relevancia de las condiciones de accesibilidad a la hora de condicionar las elecciones de localización de los agentes urbanos así como la capitalización de los beneficios derivados del transporte en los bienes inmobiliarios.

5.1.2. Conclusiones sobre la influencia de otros factores en las elecciones de localización y en los precios de los bienes inmobiliarios

El factor accesibilidad es únicamente uno de los determinantes de las elecciones de localización de los agentes urbanos. Otros factores como el precio del suelo o las características ambientales de cada zona, son también importantes en la localización urbana (Fujita, 1989). En los modelos econométrico – espaciales estimados a lo largo de esta tesis, se han obtenido los parámetros de toda una serie de variables no relacionadas con el transporte y, aunque son factores secundarios para los objetivos de esta investigación, se resumirán brevemente algunos de los resultados más notables ya que también pueden ser importantes de cara a la medición de los impactos provocados por la implantación de distintas políticas y proyectos de movilidad.

En los modelos de localización residencial el número de viviendas presentes en cada zona o el precio de éstas fueron factores significativos. La interacción con los ingresos fue importante para determinar el que otras variables como el prestigio o el número de centros de enseñanza de una zona son factores también muy relevantes en la localización de los hogares de mayor capacidad adquisitiva.

El modelo de localización de actividades económicas mostró cómo los factores número de empleos básicos de cada zona y otras características ambientales son elementos centrales en la localización de las actividades orientadas a la demanda interna. Considerando el resto de factores como constantes, las actividades de servicio mostraron una mayor utilidad al localizarse en áreas turísticas, mientras que el factor centro urbano fue significativo tanto para las actividades de comercio minorista como de servicios.

Por último en los modelos hedónicos estimados en el apartado 4, factores estructurales de los inmuebles y las edificaciones como el número de cuartos de baño, la presencia de ascensor o la propiedad de un garaje asociado aumentaron el valor de los inmuebles de forma notable en más de un 10% por unidad adicional. Entre las características ambientales fue especialmente importante la presencia de playas en las cercanías de los inmuebles con incrementos en los precios inmobiliarios de en torno al 30%.

Por lo tanto factores relacionados con el mercado inmobiliario (oferta y demanda de vivienda) y con las características ambientales de cada área tienen una elevada importancia a la hora de explicar tanto la localización de los agentes urbanos como los precios inmobiliarios. Es importante tener en cuenta este hecho ya que, como se comentará más adelante, las políticas de transporte sólo tendrán los resultados deseados en materia del impacto sobre los usos del suelo si se coordinan adecuadamente con otras políticas públicas.

5.1.3. Conclusiones sobre el uso de modelos considerando relaciones espaciales

Los modelos especificados en esta tesis tienen en todos los casos una fuerte componente espacial. En este tipo de modelos factores como los costes de desplazamiento entre lugares, la localización de las actividades y la contigüidad o la difusión de los fenómenos se consideran clave a la hora de simular la interacción entre el transporte y los usos del suelo.

El modelo conjunto formulado en el apartado 2 si bien presentaba una clara componente espacial (zonificación y red de transporte), no incluía explícitamente en la estimación de los parámetros de las ecuaciones la posible dependencia entre observaciones. La importancia de la existencia de relaciones de dependencia espacial entre observaciones es subrayada por la denominada primera ley de la geografía (Tobler, 2004; Tobler, 1970): "Todos los fenómenos están relacionados entre sí pero los fenómenos más próximos están más relacionados que los más lejanos". Los modelos de localización residencial y de estimación de precios inmobiliarios desarrollados en los apartados 3 y 4, mejoraron los modelos desarrollados en el apartado 2 al simular explícitamente estas relaciones de proximidad. La simulación de la dependencia entre observaciones tiene el objetivo de evitar problemas relacionados con la estimación de parámetros ineficientes e incluso sesgados (LeSage y Pace, 2009) así como patrones de sustitución entre alternativas poco realistas en el caso de los modelos de elección zonal (Hunt et al., 2004).

Los modelos de elección residencial Nested Logit (NL) y Cross – Nested Logit (CNL) estimados en el apartado 3 presentaron ajustes significativamente mejores a los datos que los modelos más simples logit multinomial (MNL) según el test de razón de verosimilitud (LR). En el caso de los modelos NL la estructura de correlación entre alternativas tiene que especificarse a priori por lo que sus resultados deben interpretarse con cierta cautela al depender la zonificación del criterio aplicado por el investigador. En cambio los modelos CNL incluyen la estimación de la estructura de correlación dentro del proceso general de optimización. El modelo denominado como CNL – 1 presentó valores de los parámetros significativamente distintos a uno en todos los nidos lo que reveló la presencia de cierto grado de correlación espacial entre las distintas áreas de la zona de estudio. Por lo tanto a nivel general este tipo de modelos presenta un mayor grado de realismo y bondad estadística a la hora de realizar previsiones. Sin embargo los modelos con correlación entre alternativas también tienen sus desventajas ya que su estimación, especialmente en el caso de los CNL con conjuntos de elección muy grandes (como suele ser el caso en el ámbito de la elección de localización) puede conllevar tiempos de estimación elevados. Sin embargo es probable que con el rápido avance del hardware y el software especializado los tiempos de estimación disminuyan y este tipo de modelos puedan utilizarse de forma estándar en el contexto de la elección espacial.

Los modelos hedónicos autoregresivos (SAR), autoregresivos en el término de error (SEM) y espacial Durbin (SDM) estimados en el apartado 4 ofrecieron también una mejor bondad de ajuste en una muestra que presentó una fuerte correlación espacial en los residuos de los modelos de regresión convencionales (MLR). El modelo SEM seleccionado como el mejor del conjunto de estimaciones realizadas, presentó un ajuste claramente superior a la del mejor modelo MLR (log – verosimilitud de -22.51 versus -111.24). Esto se debió seguramente a que el modelo MLR presentó un error de especificación por variable omitida relacionada con características del ambiente lo que puede causar residuos fuertemente correlacionados. En cambio el modelo espacial Durbin con términos autoregresivos tanto en las variables independientes como en la variable dependiente mostró resultados contraintuitivos en el signo de ciertas variables en algunas de sus especificaciones. Dado que este modelo aún tiene un carácter experimental tanto a nivel teórico como a nivel de implementación de software, por el momento parece recomendable priorizar el uso de los modelos más conocidos SAR y SEM en el ámbito de la planificación. Estos dos modelos únicamente presentan la dificultad añadida respecto a los modelos hedónicos convencionales de tener que especificar una estructura de correlación espacial a priori. En este estudio sin embargo, los parámetros estimados fueron bastante similares en todas las estructuras de correlación empleadas con la excepción ya comentada del modelo Durbin. Por lo tanto puede concluirse que este factor no modifica en gran medida los resultados obtenidos. Una conclusión que además está respaldada por otros estudios anteriores (Mejía-Dorantes, 2011).

5.1.4. Conclusiones sobre la aplicación práctica de los modelos

El objetivo final de los modelos especificados y allí donde pueden tener una mayor utilidad social es en ayudar a predecir ex – ante los efectos de la implantación de políticas y proyectos de transporte sobre los usos del suelo a medio y largo plazo.

Los resultados del modelo conjunto estimado en el apartado 2 fueron comparados con los datos observados mostrando resultados satisfactorios en su bondad de ajuste. Puede considerarse por lo tanto como una herramienta útil para predecir cambios ante la implantación de medidas relacionadas con la planificación del transporte. Badoe y Miller (2000) recomiendan usar este tipo de modelos integrados ya que son los únicos que pueden replicar las complejas interacciones causales que se dan en los sistemas urbanos.

Los modelos estimados en los apartados 3 y 4 mostraron que tanto la localización de los agentes urbanos como los precios inmobiliarios son sensibles a las condiciones de accesibilidad al transporte. Sin embargo esto no implica que la inversión en proyectos como un nuevo modo de transporte o una nueva infraestructura signifique automáticamente que se consigan los impactos esperados en materia de aumento del valor de los bienes inmobiliarios, de generación de una mayor densidad de población o de revitalización del centro urbano. En primer lugar porque estos dos últimos modelos sólo trataron un aspecto de la interrelación entre los usos del suelo y el transporte con lo que deberían integrarse en un modelo conjunto para realizar análisis y predicciones más adecuadas. En segundo lugar porque los resultados que aporta la literatura, derivados de múltiples estudios de caso ex –post, muestran resultados bastante diversos en función del proyecto y las características del área de implantación.

En el caso de la inversión en nuevas infraestructuras de transporte privado como las autopistas, hay cierta evidencia que avala la hipótesis de que éstas podrían aumentar la dispersión urbana dados los aumentos de accesibilidad lineal que provocan a lo largo de todo su recorrido. Es por lo tanto un tipo de medida que no incrementa a medio y largo plazo la sostenibilidad urbana aunque dado que no es el único factor causal del sprawl es dudoso que no construir nuevas autopistas reduzca la tasa de incremento de la dispersión (Handy, 2005). En cualquier caso la magnitud de este tipo de efectos podrían ser evaluados ex – ante por modelos similares a los especificados a la largo de esta tesis.

En el caso de la inversión en nuevas infraestructuras y servicios de transporte público, se han constatado resultados variados. Algunas experiencias han sido claramente exitosas tanto en el uso del nuevo modo de transporte público como en sus impactos sobre los usos del suelo, mientras que otras han presentado el carácter de casos de relativo fracaso o, al menos, con resultados por debajo de las expectativas creadas. Si se considera como objetivo la revitalización del centro y la generación de una mayor densidad urbana se pueden contar entre los casos más representativos de éxito sistemas como el del metro de Washington DC, el metro ligero de Calgary o el sistema Bus Rapid Transit de Curitiba. Entre los casos con resultados por debajo de los esperados estarían sin embargo el sistema de metro del área de la bahía de San Francisco (BART) o el sistema de metro ligero de Búfalo, Nueva York (Berechman y Paaswell, 1983; Cervero y Landis, 1997).

En todos los casos, los investigadores que han evaluado los resultados de estos proyectos han resaltado el carácter de factor necesario pero no suficiente de la inversión en transporte público a la hora de modificar el desarrollo urbano. Cervero (1998) ha resumido las conclusiones obtenidas en la experiencia internacional en los siguientes puntos:

- Los sistemas de transporte público redistribuyen más que crean crecimiento.
- Un prerrequisito para que un sistema de transporte público tenga un impacto elevado es que la economía regional esté en crecimiento. Además los impactos serán mayores si el nuevo modo de transporte se pone en marcha justo antes de que se produzca un ciclo de crecimiento.
- Los sistemas de transporte público con una configuración radial pueden reforzar los centros urbanos. Además también pueden fomentar el desarrollo inmobiliario y comercial de los centros urbanos con problemas si se acompañan con incentivos a la iniciativa privada.
- Los sistemas de transporte público generalmente refuerzan los patrones de descentralización. Por lo tanto es necesario un planeamiento proactivo si se quiere que el crecimiento descentralizado se concentre en subcentros.

- Las inversiones en sistemas de transporte público deben acompañarse con medidas pro – desarrollo: incentivos al desarrollo inmobiliario denso, disponibilidad de suelo, diseño orientado al movimiento peatonal etc.
- La limitación de la oferta de aparcamiento y la eliminación del aparcamiento gratuito ayuda en gran medida al desarrollo en el entorno de las estaciones de transporte público.

Por lo tanto puede decirse que para revertir o al menos minimizar el fuerte equilibrio impulsado por el círculo vicioso del transporte y los usos del suelo, se necesitan realizar toda una serie de políticas coordinadas que afecten tanto al ámbito de la planificación del transporte como a la de los usos del suelo. Entre las iniciativas por el lado de los usos del suelo se requieren políticas de diseño urbano que faciliten el acceso peatonal a las estaciones, la densificación de los desarrollos inmobiliarios, la mezcla de usos del suelo y la reducción de los espacios de parking para vehículos privados. Los diseños urbanos del tipo Transit Oriented Development (TOD) reúnen todas estas características y han demostrado ser útiles en varios casos de aplicación, tanto en el terreno del fomento del uso del transporte público como a la hora de implementar un desarrollo urbano más sostenible (Cervero et al., 2004). Por el lado de las políticas de transporte, la creación de nuevas infraestructuras y servicios se pueden complementar con políticas de gestión de la demanda como la introducción de tarifas de congestión y de aparcamiento en las áreas centrales que reduzcan el uso del vehículo privado (Shoup, 2005). Sin el complemento de estas políticas, es probable que no se generen incentivos suficientes para cumplir el objetivo de una movilidad y un desarrollo urbano más sostenible. A esto hay que añadir que sin crecimiento económico es improbable que la inversión en transporte público genere desarrollo urbano por si sola (véase Fig 5-1).

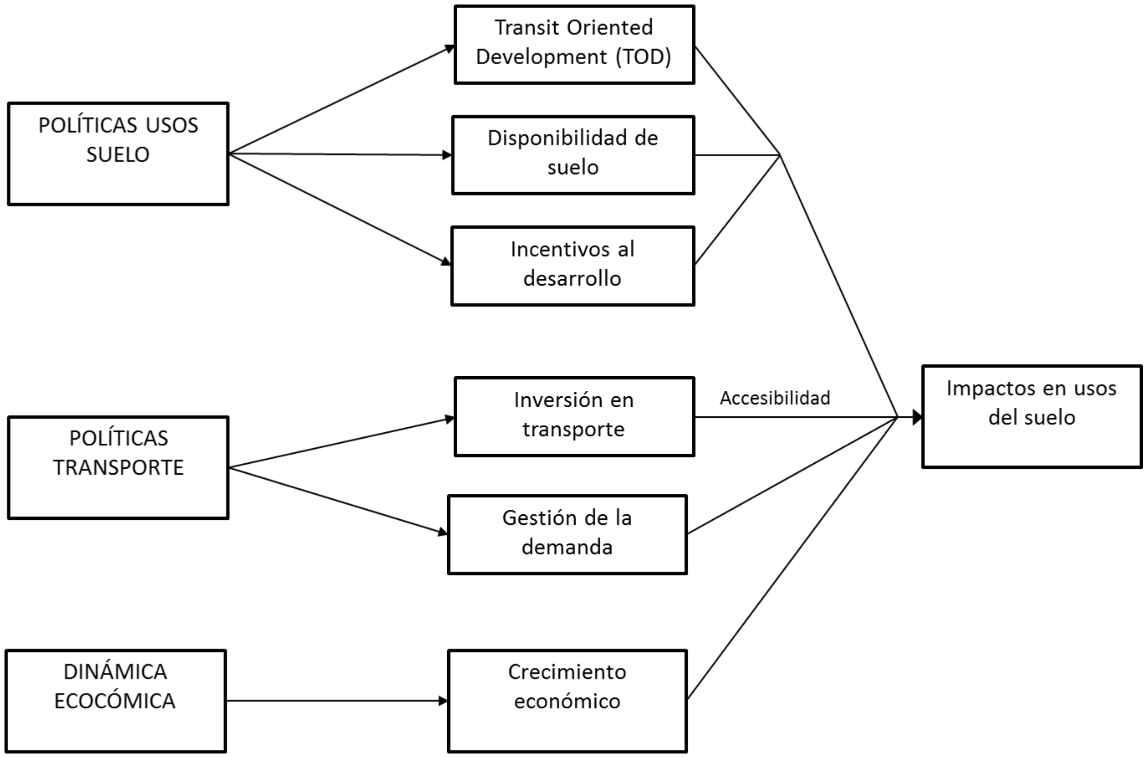

Fig 5-1. Factores que influyen en los impactos sobre los usos del suelo. Basado parcialmente en Knight y Trygg (1977)

En cuanto a las políticas destinadas al incremento de los valores inmobiliarios, es interesante subrayar que los modelos hedónicos pueden emplearse para realizar políticas de captura del valor que ayuden a financiar los proyectos de transporte (Smith y Gihring, 2006). Este tipo de políticas pueden incrementar notablemente los recursos de financiación para realizar inversiones en transporte público y ya se han implantado con notable éxito en ciudades como Hong Kong (Cervero y Murakami, 2009) o Singapur (Chi-Man Hui et al., 2004). En el contexto español algunos autores han defendido la posibilidad de aplicar políticas de captura del valor a través de impuestos como el de bienes inmuebles (Mejia-Dorantes y Vasallo Magro, 2010). En cualquier caso estas políticas de captura del valor deberían aplicarse como estrategias de búsqueda de una mayor eficiencia y equidad en la inversión en los proyectos de transporte donde los agentes más beneficiados contribuyan en mayor medida a su financiación (Zhao et al., 2012). Además en el caso de las infraestructuras ferroviarias debe tenerse en cuenta también el diseño urbano de tal forma que se minimicen sus posibles externalidades negativas (ruido, vibraciones etc.) que, como se ha visto anteriormente, pueden tener un impacto negativo sobre el precio de los inmuebles presentes en su área de influencia.

5.2. Líneas Futuras de Investigación

A partir de los trabajos realizados en esta tesis quedan abiertas diferentes líneas de investigación que se espera continuar en el futuro:

i. Recolectar datos más exhaustivos y precisos sobre la movilidad de la población y sobre la localización y características de hogares, empresas y bienes inmobiliarios.

ii. Realizar un análisis de sensibilidad que determine la influencia de diversos patrones de localización de población y actividades a la solución final proporcionada por los modelos.

iii. Especificar y estimar un modelo de localización de actividades económicas que tenga en cuenta la dependencia espacial entre observaciones.

iv. Aplicar el modelo conjunto de cara a la simulación de los efectos de la realización de un proyecto concreto: la implantación de un metro ligero en la ciudad de Santander. Puede verse un ejemplo de análisis preliminar en el Anexo B.

v. En caso de realizarse un proyecto del tipo anterior y pasado un tiempo suficiente podrían compararse los resultados de la evaluación ex – ante realizada a partir de los modelos presentados con una evaluación ex – post.

REFERENCIAS

REFERENCIAS

Alonso, W. (1964) *Location and land use: toward a general theory of land rent*. Harvard University Press, Cambridge.

Anas, A. (1981) The estimation of multinomial logit models of joint location and travel mode choice from aggregated data. *Journal of Regional Science* 21, 223-242.

Anas, A. (1982) *Residential location markets and urban transportation: economic theory, econometrics, and policy analysis with discrete choice models*. Academic Press, New York.

Anas, A. (1983) Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research Part B: Methodological* 17, 13-23.

Andersson, D.E., Shyr, O.F., Fu, J. (2010) Does high-speed rail accessibility influence residential property prices? Hedonic estimates from southern Taiwan. *Journal of Transport Geography* 18, 166-174.

Andrews, R.B. (1953) Mechanics of the Urban Economic Base: Historical Development of the Base Concept. *Land Economics* 29, 161-167.

Anselin, L. (1988) *Spatial econometrics : methods and models*. Kluwer Academic Publishers, Dordrecht; Boston.

Anselin, L. (2010) Thirty years of spatial econometrics. *Papers in Regional Science* 89, 3-25.

Armstrong, R., Rodríguez, D. (2006) An Evaluation of the Accessibility Benefits of Commuter Rail in Eastern Massachusetts using Spatial Hedonic Price Functions. *Transportation* 33, 21-43.

Badoe, D.A., Miller, E.J. (2000) Transportation-land-use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research Part D: Transport and Environment* 5, 235-263.

Banister, D., Berechman, J. (2000) *Transport investment and economic development*. UCL Press, London.

Banister, D., Thurstain-Goodwin, M. (2011) Quantification of the non-transport benefits resulting from rail investment. *Journal of Transport Geography* 19, 212-223.

Barra, T.d.l. (1989) *Integrated land use and transport modelling : decision chains and hierarchies*. Cambridge University Press, Cambridge ; New York.

Ben-Akiva, M., Bierlaire, M. (1999) Discrete Choice Methods and their Applications to Short Term Travel Decisions. *Handbook of Transportation Science,* .

Berechman, J., Paaswell, R.E. (1983) Rail rapid transit investment and CBD revitalisation: methodology and results (Buffalo, New York State). *Urban Studies* 20, 471-486.

Bhat, C.R., Guo, J. (2004) A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B: Methodological* 38, 147-168.

Bierlaire, M. (2003) Biogeme: A free package for the estimation of discrete choice models. *Proceedings of Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland*.

Bierlaire, M. (2006) A theoretical analysis of the cross-nested logit model. *Annals of Operations Research* 144, 287-300.

Bitter, C., Mulligan, G., Dall'erba, S. (2007) Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems* 9, 7-27.

Bowes, D.R., Ihlanfeldt, K.R. (2001) Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics* 50, 1-25.

Brotchie, J.F., Dickey, J.W., Sharpe, R. (1980) *TOPAZ : general planning technique and its applications at the regional, urban, and facility planning levels*. Springer-Verlag, Berlin ; New York.

Buchanan, C., Partners (2003) Economic and Regeneration Impact of Tramlink. *Report for the South London Partnership. Funded by the London Development Agency*. Transport for London and the London Boroughs of Croydon and Sutton,July.

Camagni, R. (2005) Economía Urbana. ed Bosch, A., Barcelona.

Cascetta, E. (2009) *Transportation systems analysis : models and applications*, 2nd ed. Springer, New York.

Cervero, R. (1998) *The transit metropolis : a global inquiry*. Island Press, Washington, D.C.

Cervero, R., Kang, C.D. (2011) Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy* 18, 102-116.

Cervero, R., Landis, J. (1997) Twenty years of the Bay Area Rapid Transit system: Land use and development impacts. *Transportation Research Part A: Policy and Practice* 31, 309-333.

Cervero, R., Murakami, J. (2009) Rail and Property Development in Hong Kong: Experiences and Extensions. *Urban Studies* 46, 2019-2043.

Cervero, R., National Research Council (U.S.). Transportation Research Board., Transit Cooperative Research Program., United States. Federal Transit Administration., Transit Development Corporation. (2004) *Transit-oriented development in the United States : experiences, challenges, and prospects*. Transportation Research Board, Washington, D.C.

Coppola, P., Nuzzolo, A. (2011) Changing accessibility, dwelling price and the spatial distribution of socio-economic activities. *Research in Transportation Economics* 31, 63-71.

Coppola, P., Ibeas, Á., dell'Olio, L., Cordera, R. (2013) A LUTI Model for the Metropolitan Area of Santander. *Journal of Urban Planning and Development*, 139, 3, 153-165.

Court, A.T. (1939) Hedonic Price Indexes With Automobile Examples. *The dynamics of automobile demand*. General Motors, New York.

Cropper, M.L., Leland, B.D., McConnell, K.E. (1988) On the Choice of Funtional Form for Hedonic Price Functions. *The Review of Economics and Statistics* 70, 668-675.

Chen, J., Chen, C., Timmermans, H. (2008) Accessibility Trade-Offs in Household Residential Location Decisions. *Transportation Research Record: Journal of the Transportation Research Board* 2077, 71-79.

Chi-Man Hui, E., Sze-Mun Ho, V., Kim-Hin Ho, D. (2004) Land value capture mechanisms in Hong Kong and Singapore: A comparative analysis. *Journal of Property Investment & Finance* 22, 76-100.

De Cea, J., Fernandez, E. (1993) Transit Assignment for Congested Public Transport Systems: An Equilibrium Model. *Transportation science* 27, 133-147.

De Cea, J., Fernandez, J.E., Dekock, V., Soto, A., Friesz, T.L. (2003) ESTRAUS: A computer package for solving supply-demand equilibrium problems on multimodal urban transportation networks with multiple user classes. *Annual Meeting of the Transportation Research Board*, Washington, D.C.

Debrezion, G., Pels, E., Piet, R. (2006) The Impact of Rail Transport on Real Estate Prices: An Empirical Analysis of the Dutch Housing Markets. ed Vrije Universiteit Amsterdam, a.T.I., Amsterdam.

Debrezion, G., Pels, E., Rietveld, P. (2007) The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics* 35, 161-180.

Echenique, M.H. (1994) Urban and regional studies at the Martin Centre: its origins, its present, its future. *Environment and Planning B: Planning and Design* 21, 517-533.

Echenique, M.H. (2011) Land Use/Transport models and economic assessment. *Research in Transportation Economics*, 45-54.

Eliasson, J., Mattsson, L.-G. (2000) A model for integrated analysis of household location and travel choices. *Transportation Research Part A: Policy and Practice* 34, 375-394.

Felsenstein, D., Axhausen, K.W., Waddell, P. (2010) Land Use-Transportation Modeling with UrbanSim: Experiences and Progress. *The Journal of Transport and Land Use* 3, 1-3.

Foot, D.H.S. (1981) *Operational urban models : an introduction*. Methuen, London ; New York.

Forrest, D., Glen, J., Ward, R. (1996) The Impact of a Light Rail System on the Structure of House Prices: A Hedonic Longitudinal Study. *Journal of Transport Economics and Policy* 30, 15-29.

Fujita, M. (1989) *Urban economic theory : land use and city size*. Cambridge University Press, Cambridge Cambridgeshire; New York.

García Palomares, J.C. (2007) Movilidad laboral en la Comunidad de Madrid. Universidad Complutense.

García Palomares, J.C., Gutiérrez Puebla, J. (2007) Pautas de la movilidad en el área metropolitana de Madrid. Cuadernos de Geografía 81/82, 7-29.

Geurs, K.T., van Wee, B. (2004) Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography* 12, 127-140.

Giuliano, G. (1995) The weakening Transportation-Land Use Connection. *Access* 6, 3-11.

Glaeser, E.L. (2008) *Cities, agglomeration, and spatial equilibrium*. Oxford University Press, Oxford.

Griffith, D. (1987) *Spatial Autocorrelation: A Primer*. Association of American Geographers Resource Publication, Washington, DC.

Gujarati, D.N., Porter, D.C. (2009) *Basic econometrics*, 5th ed. McGraw-Hill Irwin, Boston.

Guo, J., Bhat, C. (2001) Residential Location Choice Modeling: Accommodating Sociodemographic, School Quality and Accessibility Effects. University of Texas, Austin.

Halvorsen, R., Palmquist, R. (1980) The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70, 474-475.

Handy, S. (2005) Smart Growth and the Transportation-Land Use Connection: What Does the Research Tell Us? *International Regional Science Review* 28, 146-167.

Handy, S.L., Niemeier, D.A. (1997) Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A* 29, 1175-1194.

Hansen, W.G. (1959) How Accessibility Shapes Land Use. *Journal of the American Institute of Planners* 25, 73-76.

Harris, B. (1985) Urban Simulation Models in Regional Science. *Journal of Regional Science* 25, 545–567.

Haynes, K.E., Fotheringham, A.S. (1990) The Impact of Space on the Application of Discrete Choice Models. *Review of Regional Studies* 20, 39-49.

Herbert, J.D., Stevens, B.H. (1960) A model for the distribution of residential activity in urban areas. *Journal of Regional Science* 2, 21-36.

Hometrack (2005) Key Housing Market Stastistics for England and Wales, The hometrack national May 2004 survey of the housing market. UK House Price Statistics.

Hoyt, H. (1939) *The Structure and Growth of Residential Neighbourhoods in American Cities*. Federal Housing Administration, Washington D.C.

Hsu, C., Guo, S. (2006) CBD Oriented Commuters' Mode and Residential Location Choices in an Urban Area with Surface Streets and Rail Transit Lines. *Journal of Urban Planning and Development* 132, 235-246.

Hunt, L.M., Boots, B., Kanaroglou, P.S. (2004) Spatial choice modelling: new opportunities to incorporate space into substitution patterns. *Progress in Human Geography* 28, 746-766.

Iacono, M., Levinson, D., El-Geneidy, A. (2008) Models of transportation and land use change: A guide to the territory. *Journal of Planning Literature* 22, 323-340.

Ibeas, A., Cordera, R., dell'Olio, L., Moura, J.L. (2011) Modelling demand in restricted parking zones. *Transportation Research Part A: Policy and Practice* 45, 485-498.

Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P., Dominguez, A. (2012) Modelling transport and real-estate values interactions in urban systems. *Journal of Transport Geography* 24, 370-382.

Ibeas, Á., Cordera, R., dell'Olio, L., Coppola, P. (2013) Modelling the spatial interactions between workplace and residential location. *Transportation Research Part A: Policy and Practice* 49, 110-122.

Isard, W. (1956) *Location and Space-Economy*. MA: MIT Press, Cambrige.
Joyce, D., Dermot, G., Martin, S. (2007) Vehicle Ownership and Income Growth, Worldwide: 1960-2030. *The Energy Journal* 28, 143-170.
Knight, R.L., Trygg, L.L. (1977) Land-use impacts of rapid transit systems: implications of recent experience. US Department of Transportation.
Lancaster, K.J. (1966) A New Approach to Consumer Theory. *Journal of Political Economy* 74, 132-157.
Lee, B., Waddell, P. (2010) Residential mobility and location choice: a nested logit model with sampling of alternatives. *Transportation* 37, 587-601.
Leontief, W. (1966) *Essays in economics; theories and theorizing*. Oxford University Press, New York.
Lerman, S.R. (1976) Location, housing, automobile owership and mode to work: a joint choice model. *Transportation Research Record* 610, 6-11.
LeSage, J.P., Pace, R.K. (2009) *Introduction to spatial econometrics*. CRC Press, Boca Raton.
LeSage, J.P., Pace, R.K. (2010) Spatial Econometric Models. *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Springer, Berlin, 355-376.
Litman, T. (2010) Evaluating Transportation Land Use Impacts. Victoria Transport Policy Institute, Victoria.
Löchl, M., Axhausen, K.W. (2010) Modelling hedonic residential rents for land use and transport simulation while considering spatial effects. *The Journal of Transport and Land Use* 3, 39-63.
Lowry, I.S. (1964) *A model of metropolis*. Rand Corporation, Santa Monica, Calif.
Maguire, D.J., Batty, M., Goodchild, M.F. (2005) *GIS, spatial analysis, and modeling*, 1st ed. ESRI Press, Redlands, Calif.
Malpezzi, S. (2008) Hedonic Pricing Models: A Selective and Applied Review. *Housing Economics and Public Policy* ed Tony O'Sullivan, K.G., 67-89.
Martínez, F. (1997) MUSSA: Land use model for Santiago City. *Transportation Research Record* 1552, 126-134.
Martinez, F.J. (1992) The bid-choice land-use model: an integrated economic framework. *Environment & Planning A* 24, 871-885.
Martínez, F.J. (2000) Towards a Land-use and transport interaction framework. *Handbook of Transport Modelling* eds Hensher, D.A., Button, K.J. Elsevier Science, 145-164.
Martínez, L., Viegas, J. (2009) Effects of Transportation Accessibility on Residential Property Values. *Transportation Research Record* 2115, 127-137.
May, A.D., Minken, H., Jonsson, D., Shepherd, S.P., Järvi, T., Page, M., Pearman, A., Pfaffenbichler, P.C., Timms, P., Vold, A. (2003) *Developing Sustainable Land Use and Transport Strategies - A Methodological Guidebook*, Oslo.
McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. ed Zarembka, P. Academic Press, New York.
McFadden, D.L. (1977) Modelling the Choice of Residential Location. Cowles Foundation for Research in Economics, Yale University.
Mejía-Dorantes, L. (2011) Transportation infraestructure impacts on house prices and firms location:The effect of a new metro line in the Suburbs of Madrid. Universidad Politécnica de Madrid.
Mejia-Dorantes, L., Vasallo Magro, J.-M. (2010) Financing Urban Transport Through Value Capture. *Highway and Urban Environment* eds Rauch, S., Morrison, G.M., Monzón, A. Springer Netherlands, 15-21.
Mills, E.S. (1972a) *Studies in the structure of the urban economy*. Published for Resources for the Future by Johns Hopkins Press, Baltimore.
Mills, E.S. (1972b) *Urban economics*. Scott, Glenview, Ill.
Munoz-Raskin, R. (2010) Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy* 17, 72-84.

Muth, R.F. (1969) *Cities and housing; the spatial pattern of urban residential land use*. University of Chicago Press, Chicago.

Nelson, A., Genereux, J., Genereux, M. (1992) Price Effects of Landfills on Residential Land Values. *Journal of Urban Planning and Development* 118, 128-137.

Nogués, S. (1990) Caracterización socioeconómica y territorial del área de la Bahía de Santander. *La bahia de Santander : actas de las Jornadas de estudios territoriales de Cantabria*. Asociación Cántabra de Ciencia Regional, Santander, 30-55.

Nuzzolo, A., Coppola, P. (2005) S.T.I.T.: a system of mathematical models for the simulation of land-use and transport interactions. *Proceedings of European Transportation Conference*, Strasbourg,France.

Nuzzolo, A., Coppola, P. (2007) Accessibility and socioeconomic activities location. *Proceedings of European Transportation Conference*, Noordwijkerhout, The Netherlands.

O'Sullivan, A. (2007) *Urban economics*, 6th ed. McGraw-Hill/Irwin, Boston ; London.

Ord, J.K., Getis, A. (1995) Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* 27, 286-306.

Ortúzar, J.d.D., Willumsen, L.G. (2001) *Modelling transport*, 3rd ed. J. Wiley, Chichester New York.

Ottensmann, J.R., Payton, S., Man, J. (2008) Urban location and housing prices within a hedonic model. *Journal of Regional Analysis and Policy* 38, 19-35.

Overnell, N. (2007) A Second Hedonic Longitudinal Study on the Effect on House Prices of Proximity to the Metrolink Light Rail System in Greater Manchester. University of Salford, Unpublished MSc Transport Engineering and Planning Dissertation.

Paelinck, J.H.P., Klaassen, L.H. (1979) *Spatial econometrics*. Saxon House, Farnborough, Eng.

Pagliara, F., Papa, E. (2011) Urban rail systems investments: an analysis of the impacts on property values and residents' location. *Journal of Transport Geography* 19, 200-211.

Pagliara, F., Wilson, A. (2010) The State-of-the-Art in Building Residential Location Models. *Residential Location Choice* eds Pagliara, F., Preston, J., Simmonds, D. Springer Berlin Heidelberg, 1-20.

Pellegrini, P.A., Fotheringham, A.S. (2002) Modelling spatial choice: a review and synthesis in a migration context. *Progress in Human Geography* 26, 487-510.

Pinjari, A., Pendyala, R., Bhat, C., Waddell, P. (2011) Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. *Transportation* 38/6, 1-26.

Putman, S. (1996) Extending DRAM Model: Theory-Practice Nexus. *Transportation Research Record: Journal of the Transportation Research Board* 1552, 112-119.

Quigley, J. (1976) Housing Demand in the Short Run: An Analysis of Polytomous Choice. National Bureau of Economic Research, Inc, 76-102.

Reilly, W.J. (1931) *The law of retail gravitation*. W.J. Reilly, New York,.

Rodríguez, D.A., Mojica, C.H. (2009) Capitalization of BRT network expansions effects into prices of non-expansion areas. *Transportation Research Part A: Policy and Practice* 43, 560-571.

Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82, 34-55.

Sener, I.N., Pendyala, R.M., Bhat, C.R. (2011) Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior. *Journal of Transport Geography* 19, 294-303.

Senior, M.L. (2009) Impacts on travel behaviour of Greater Manchester's light rail investment (Metrolink Phase 1): evidence from household surveys and Census data. *Journal of Transport Geography* 17, 187-197.

Shoup, D.C. (2005) *The high cost of free parking*. Planners Press, American Planning Association, Chicago.

Simmonds, D. (2001) The Objectives and Design of a New Land-use Modelling Package: DELTA. *Regional Science in Business* eds Clarke, G., Madden, M. Springer Berlin Heidelberg, 159-188.

Small, K.A. (1987) A Discrete Choice Model for Ordered Alternatives. *Econometrica* 55, 409-424.

Smith, J.J., Gihring, T.A. (2006) Financing Transit Systems Through Value Capture. An Annotated Bibliography. Victoria Transport Policy Institute, Victoria, Canada.

Srour, I., Kockelman, K., Dunn, T. (2002) Accessibility Indices: Connection to Residential Land Prices and Location Choices. *Transportation Research Record: Journal of the Transportation Research Board* 1805, 25-34.

Stephen, S. (1999) Chapter 41 Hedonic analysis of housing markets. *Handbook of Regional and Urban Economics* eds Paul, C., Edwin, S.M. Elsevier, 1595-1635.

Tobler, W. (2004) On the First Law of Geography: A Reply. *Annals of the Association of American Geographers* 94, 304-310.

Tobler, W.R. (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46, 234-240.

Torrens, P.M. (2000) How land-use transportation models work. Centre for Advanced Spatial Analysis (CASA), London.

Train, K. (2009) *Discrete choice methods with simulation*, 2nd ed. Cambridge University Press, Cambridge ; New York.

Vernon, H. (1997) Medium size cities. *Regional Science and Urban Economics* 27, 583-612.

Von Thünen, J.H. (1826) *Der isolierte staat in beziehung auf landwirtschaft und nationaloekonomie*, Jena. Translated by C.M.Wartenburg (1966). The Isolated State. Oxford. UK:Oxford University Press.

Vovsha, P. (1997) *Cross-nested logit model: an application to mode choice in the Tel-Aviv metropolitan area*.

Waddell, P. (1993) Exogenous workplace choice in residential location models: is the assumption valid? *Geographical Analysis* 25, 65-82.

Waddell, P. (1996) Accessibility and residential location: the interaction of workplace, residential mobility, tenure, and location choices. *Lincoln Land Institute TRED Conference*.

Waddell, P. (2002) UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association* 68, 297-343.

Waddell, P. (2010) Modeling Residential Location in UrbanSim. *Residential Location Choice* eds Pagliara, F., Preston, J., Simmonds, D. Springer Berlin Heidelberg, 165-180.

Waddell, P., Bhat, C., Eluru, N., Wang, L., Pendyala, R.M. (2007a) Modeling interdependence in household residence and workplace choices. *Transportation Research Record* 2003, 84-92.

Waddell, P., Ulfarsson, G.F. (2004) Introduction to Urban Simulation: Design and Development of Operational Models. *Handbook in Transport, Volume 5: Transport Geography and Spatial Systems* ed Stopher, B., Kingsley, Hensher. Pergamon Press, 203-236.

Waddell, P., Ulfarsson, G.F., Franklin, J.P., Lobb, J. (2007b) Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice* 41, 382-410.

Ward, M.D., Gleditsch, K.S. (2008) *Spatial regression models*. Sage Publications, Thousand Oaks.

Wegener, M. (2004) Overview of land-use transport models. *Transport Geography and Spatial Systems* eds Hensher, D.A., Button, K., Kidlington, 127-146.

Wen, C.-H., Koppelman, F.S. (2001) The generalized nested logit model. *Transportation Research Part B: Methodological* 35, 627-641.

Wheeler, D., Tiefelsdorf, M. (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems* 7, 161-187.

Wilson, A.G. (1970) *Entropy in urban and regional modelling*. Pion, London.

Wright, E.O., Rogers, J. (2011) Chapter 6. Transportation., *American society : how it really works*. W.W. Norton & Co., New York.

Zeiss, C. (1990) Incinerator Impacts on Residential Property Sales: Beyond Price Effects. *Journal of Urban Planning and Development* 116, 80-97.

Zhao, Z.J., Das, K.V., Larson, K. (2012) Joint Development as a Value Capture Strategy in Transportation Finance, *The Journal of Transport and Land Use* 5, 5-17.

Zipf, G.K. (1949) *Human Behavior and the Principle of Least Effort*. MA: Addison-Wesley, Cambridge.

ANEXOS

ANEXO A. PARÁMETROS ESTIMADOS EN LOS MODELOS DE GENERACIÓN / ATRACCIÓN DE VIAJES Y ELECCIÓN MODAL

A.1 Modelo de generación/atracción de viajes

A continuación se presentan los parámetros estimados en los modelos de generación y atracción de viajes. Los modelos de localización residencial y de localización de actividades económicas (véase el apartado 2 del cuerpo principal de la tesis) proveen endógenamente variables zonales referentes al número de residentes y de empleos presentes en cada zona. Son éstas por lo tanto las variables independientes básicas que deben introducirse en los modelos de generación/atracción para simular el número total de viajes generados y atraídos. El modelo de generación de viajes ha quedado especificado y estimado como:

$$y_i = 0.221 \cdot POB_i / HOG_i \qquad \text{(A.1)}$$

Dónde:

$y_i = O_i/H_i$ son los viajes generados por hogar en la zona i en hora punta

POB_i/HOG_i son los residentes por hogar

Variable	Parámetro	t	Sig.
POB/HOG	.221	17.84	.00
R^2		.84	
R^2_{adj}		.71	

Tabla A-1. Resultados del modelo de regresión para la generación de viajes

Por su parte el modelo de atracción de viajes ha quedado especificado y estimado como:

$$y_i = 0.199 \cdot POB_i / HOG_i + 0.59 \cdot EMP_i / HOG_i \qquad \text{(A.2)}$$

Dónde:

$y_i = D_i/H_i$ son los viajes atraídos por hogar en la zona i en hora punta

POB_i/HOG_i son los residentes por hogar en la zona i

EMP_i/HOG_i son los empleos en comercios y servicios existentes por hogar en la zona i

Variable	Parámetro	t	Sig.
POB/HOG	.199	14.47	.00
EMP/HOG	.059	1.95	.05
R^2		.71	
R^2_{adj}		.71	

Tabla A-2. Resultados del modelo de regresión para la atracción de viajes

A.2 Modelo de elección modal

El modelo de transporte ha sido calibrado para simular los siguientes cuatro modos de desplazamiento:

- A Pie
- Bici
- Transporte Privado Motorizado (TMot):
- Transporte Público (TPub):
    - Bus
    - Metro – Tren de cercanías

Los parámetros del modelo de elección modal han quedado estimados como se muestra en la Tabla A-3.

Variable	Parámetro	Test t
TV_Pie	-0.25	-16.79
TV*Edad>56	0.03	5.98
TV*Sexo	0.01	3.54
CE_BICI	-8.67	-15.03
TV*Ing<1200€/mes	0.07	6.13
CE_TPriv.Motorizado	-2.29	-10.55
TV_Motorizado	-0.50	-6.05
COSTE	-2.06	-24.23
CE_Tpúblico	-4.05	-17.33
TV_Tpúblico	-0.31	-6.33
TAcceso	-0.12	-5.93
TDestino	-0.09	-4.36
TEspera	-0.02	-1.94
Log C	-4129.13	
Log L	-2391.47	

Tabla A-3. Parámetros estimados en el modelo de elección modal

Dónde:

- TV: Tiempo de Viaje en el modo
- TVEdad>56: TV*Edad>56(1 si Edad >56, 0 si Edad < 56años)
- TVSexo: TV*Sexo(1 si es mujer, 0 si es hombre),
- TVIng<1.200€/mes: TV*Ing<1.200€/mes(1 si Ing<1.200€/mes, 0 si Ing.>1200€/mes)
- COSTETMot: Coste del Transporte Privado Motorizado
- COSTETPub: Coste del Transporte Público
- TA: Tiempo de Acceso desde el origen real a la Parada Origen
- TD: Tiempo de Acceso al destino desde Parada Destino
- TE: Tiempo de Espera en la Parada Origen

Finalmente el porcentaje estimado de elección en hora punta mañana de cada modo queda resumido en la Tabla A-4.

Hora Punta Mañana									
Modo	**Ingresos<1200 €**				**Ingresos>1200 €**				**% Total por Modo**
	Mujer		**Hombre**		**Mujer**		**Hombre**		
	Edad <56	**Edad >56**	**Edad <56**	**Edad >56**	**Edad <56**	**Edad >56**	**Edad <56**	**Edad >56**	
A pie	30.2	45.1	23.6	42.0	44.8	43.1	29.5	26.1	35.55
Bici	0.55	0.50	1.10	0.50	0.50	0.00	0.7	0.1	0.49
TMot	62.5	49.12	68.00	51.92	49.39	51.38	63.03	66.64	57.76
TPub	6.75	5.28	7.30	5.58	5.31	5.52	6.77	7.16	6.20

Tabla A-4. Porcentaje de elección de cada modo, si el viaje se realiza dentro del periodo de hora punta mañana en el área metropolitana de Santander

ANEXO B. EJEMPLO DE SIMULACIÓN DE LOS IMPACTOS PROVOCADOS POR LA IMPLANTACIÓN DE UN METRO LIGERO EN EL ÁREA METROPOLITANA DE SANTANDER

El proyecto de construcción de un sistema de metro ligero en la ciudad de Santander fue planteado en torno al año 2005 por el Ayuntamiento municipal y apoyado así mismo por otros agentes sociales (Fernandez, 2007). Este proyecto está así mismo recogido en el Plan General de Ordenación Urbana y en el Plan de Movilidad Sostenible del Municipio (aprobado en el año 2010).

El PGOU del municipio de Santander contempla la construcción de una línea de metro ligero que utilizaría un futuro túnel de conexión entre las áreas de Las Estaciones y la Avenida de los Castros como vía de acceso desde el centro a la zona norte de la ciudad.

Adicionalmente el Plan de Fomento del Transporte Colectivo dentro del Plan de Movilidad Sostenible de Santander recoge "como una actuación esencial del futuro de Santander, la ejecución de una red de metro ligero, que independientemente del sistema viario, vincule áreas como las estaciones, la Universidad o el Sardinero, que son los puntos neurálgicos del modelo de ciudad que se planea" (APIAXXI, 2010a). El plan de construcción de esta red no contaría ya, como en la propuesta del PGOU, con una sola línea de metro ligero, sino con una red completa que articularía los principales espacios de la ciudad.

Estas políticas que a priori pueden considerarse muy beneficiosas para el conjunto del área metropolitana y para la economía de Cantabria, tienen como severa restricción el elevado coste del proyecto planeado. Según cálculos realizados por la empresa consultora encargada del diseño de la red, la puesta en marcha del sistema con cuatro líneas de servicio más el material móvil podría ascender a más de 113 millones de euros (EFE, 2009).

Se plantea por lo tanto la necesidad de evaluar ex – ante un proyecto de este tipo para conocer con cierto grado de confianza si los beneficios sociales aportados por el mismo son superiores a sus costes. Dentro de este tipo de evaluaciones es necesario además contar con modelos de simulación que permitan prever los impactos generados por la implantación de un nuevo modo de transporte en el conjunto del sistema urbano.

Estos impactos mientras que a corto plazo pueden afectar únicamente al funcionamiento del sistema de transporte (partición modal, asignación a la red etcétera), a medio – largo plazo pueden tener consecuencias en el conjunto del sistema urbano, incluidos la localización de la población, las actividades económicas y los valores inmobiliarios.

El presente anexo resume la simulación llevada a cabo con el modelo LUTI planteado en el apartado 2 del cuerpo principal de la tesis, para intentar prever los efectos de la implantación del nuevo sistema de metro ligero en el área metropolitana de Santander. Los resultados de esta simulación pueden usarse como estimaciones orientativas para la cuantificación de los costes y beneficios que podría implicar el proyecto de introducción de un nuevo metro ligero en la ciudad de Santander.

B.1 Codificación de la red de metro ligero

El Plan de Movilidad Sostenible de Santander (APIAXXI, 2010b) plantea la ejecución de cuatro líneas de metro ligero conectadas mediantes dos ejes transversales. Estas cuatro líneas permitirían formar una red urbana que conectaría varias de las zonas más importantes de la ciudad, tanto entre sí como con el centro urbano.

La primera de las líneas discurriría por el eje central de la ciudad entre el Hospital Valdecilla y el barrio de Puertochico así como en el eje longitudinal de la Avenida de los Castros (véase Fig B-1) hasta llegar al barrio de Cazoña. La ejecución de esta línea permitirá conectar varios de los nodos más importantes de la ciudad formando un anillo troncal para el resto de la red.

La segunda línea longitudinal propuesta por el Plan de Movilidad Sostenible, es el tramo Puertochico – Piquío. Esta pequeña línea partiría de la calle Castelar (véase Fig B-2) y articularía fundamentalmente el barrio del Sardinero conectando además con la línea 1.

La tercera línea partiría del área de Las Estaciones hasta llegar al barrio de Nueva Montaña. La ejecución de esta línea permitirá su posterior prolongación hacia los municipios de Camargo y El Astillero, llegando a conectar con la propuesta realizada por el Gobierno de Cantabria para ejecutar un tranvía entre las localidades de Sarón, Guarnizo y El Astillero formando así una autentica red ferroviaria de escala metropolitana.

La cuarta línea partiría desde Cazoña hasta alcanzar el barrio de El Alisal. Este eje en conexión con la línea 1 permitirá la movilidad entre los dos nodos de investigación y desarrollo más importantes de Cantabria, como son la Universidad y el Parque Científico y Tecnológico (PCTCAN), al mismo tiempo que mejoraría notablemente la relación con el transporte público de zonas como la ladera norte de General Dávila, La Albericia o El Alisal.

Esta red de cuatro líneas se acompañaría además de dos ejes transversales. El primero de ellos comunicaría el centro y el área intermodal de las estaciones con la Avenida de los Castros, a la altura de la Bajada de Polio, por lo que desde la zona centro se podrían habilitar movimientos tanto hacia El Sardinero como hacia el PCTCAN. Finalmente, el eje transversal de Nueva Montaña, uniría el PCTCAN con la línea tres a la altura de Nueva Montaña, permitiendo crear un trazado circular que albergaría todos los nuevos crecimientos en esta zona de la ciudad.

Con esta configuración, la propuesta de metro ligero avalada por el Ayuntamiento de la ciudad se prolongaría por algo más de 20,000 metros de red.

Fig B-1. Infografía del metro ligero de Santander a su paso por la Calle Castelar (Línea 2). Fuente: Ayto de Santander

Fig B-2. Infografía del metro ligero de Santander a su paso por la Avenida de Los Castros (Línea 1). Fuente: Ayto de Santander

De cara a la simulación, las nuevas líneas de metro ligero proyectadas se han codificado como una subred independiente en el software ESTRAUS. Esto significa que no presentan conectividad con la red ferroviaria existente encargada de posibilitar los desplazamientos dentro del área metropolitana. Sin embargo sí se ha habilitado la realización de transbordos en el área intermodal de las estaciones.

Las líneas de metro ligero introducidas en el modelo han sido las siguientes (véase también Fig B-3 y Fig B-4):

1. Línea Puertochico – Cazoña - Piquío
2. Línea Puertochico - Piquío
3. Estaciones – Corte Inglés

4. Cazoña – PCTCAN – Corte Inglés

Los datos orientativos que se han utilizado para realizar la simulación se muestran en la Tabla B-1.

Línea	Frecuencia(min)	Longitud de la línea (m)	Tarifa(€)	Capacidad
1	15	9760	2	220
2	15	2700	2	220
3	15	5610	2	220
4	15	5340	2	220

Tabla B-1. Características de las líneas de metro ligero proyectadas

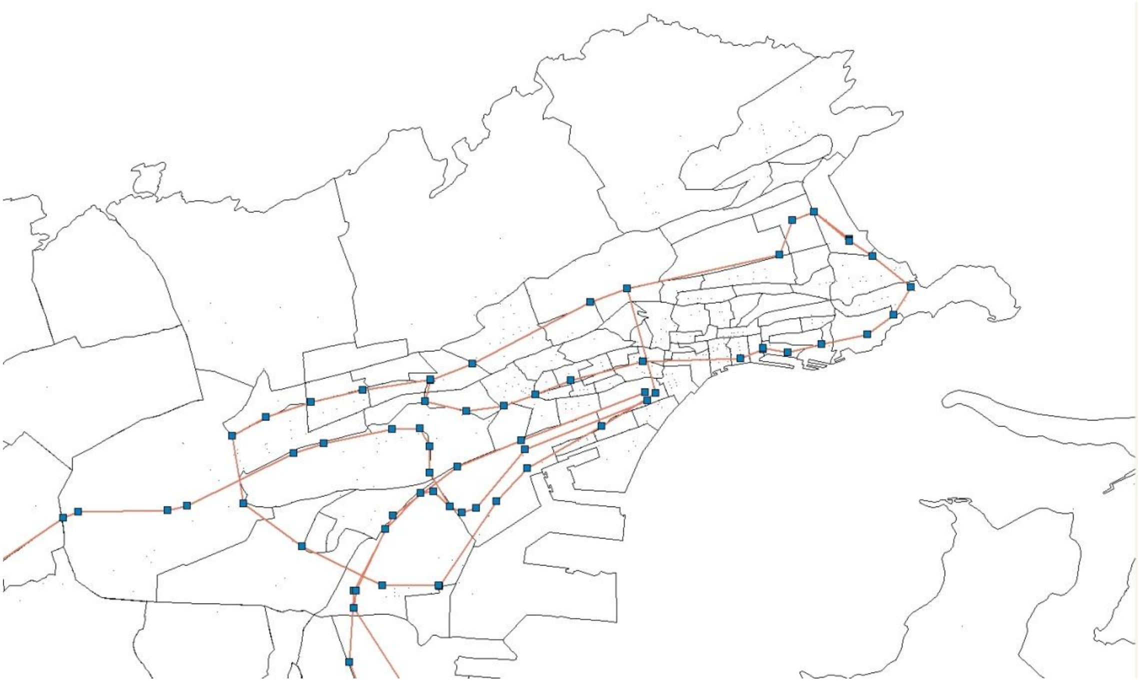

Fig B-3. Red de metro ligero codificada en el software ESTRAUS

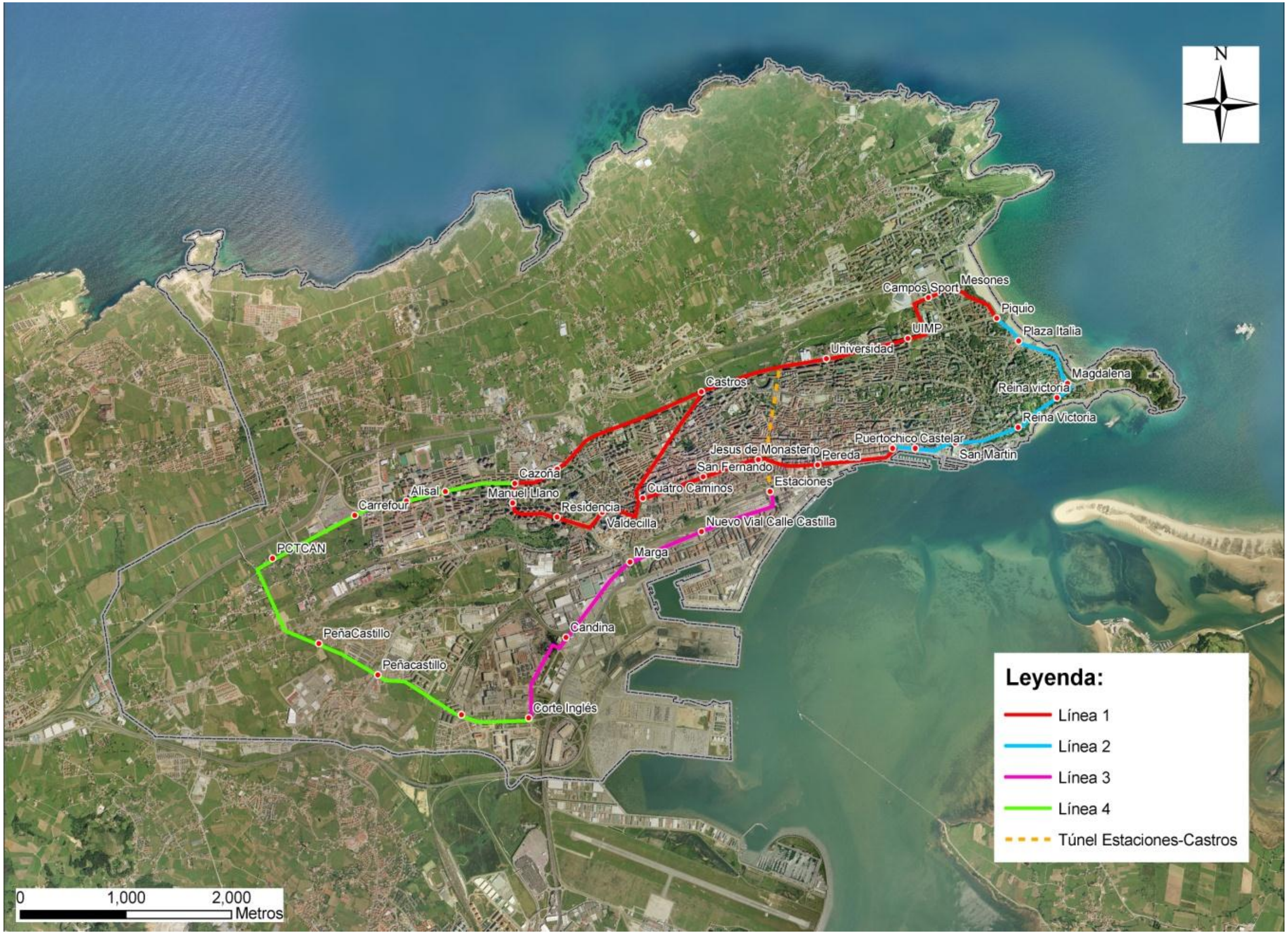

Fig B-4. Red de metro ligero planificada

B.2 Repercusión del nuevo modo en el conjunto del sistema de transporte

Para analizar la repercusión del nuevo sistema de metro ligero en el conjunto del sistema de transporte, pueden utilizarse los datos arrojados por el modelo ESTRAUS. Desde el punto de vista de la simulación se analizarán los datos relativos a la partición modal y a los niveles de servicio comparando la solución de equilibrio con la calibración para el año base.

Línea	Pasajeros	Pasajeros * Km
Línea 1	976.85	2216.10
Línea 2	32.28	33.88
Línea 3	414.00	887.97
Línea 4	122.01	257.62

Tabla B-2. Pasajeros transportados por línea de metro ligero en Hora Punta Mañana

En total el sistema de metro ligero transporta, según las estimaciones realizadas por el modelo en hora punta de mañana, más de 1500 pasajeros. Como puede observarse en la Tabla B-2, es la línea 1 la que desplaza más de la mitad de los viajeros seguida por la línea 3 y 4. La línea 2 tiene un número de pasajeros muy reducido lo cual es concordante con la zona a la que da servicio de funcionalidad casi exclusivamente residencial y de ocio por lo que la mayor demanda potencial de movilidad desde o hacia la zona debería producirse en horas fuera de punta y en días no laborables.

En cuanto a la partición modal, destaca el hecho de que la introducción del metro ligero supondría, según el modelo, una reducción del número de viajes en auto privado, algo acorde con los objetivos de movilidad sostenible y de fomento del transporte colectivo. El modo Bus urbano y el modo combinado Metro – Bus presentarían por el contrario una disminución ligera en la partición modal. La mayor parte de estas reducciones de viajes son captados por el metro ligero y el tren de cercanías los cuales pasan de suponer un 0.87% de la partición modal en el año base a un 4.4%. En el mismo sentido, el bus interurbano también aumenta ligeramente su número de viajes y su presencia en la partición modal seguramente debido a que la combinación bus interurbano – metro ligero supone para los usuarios una elección modal de mayor utilidad.

Los niveles de servicio (véase Tabla B-3, Tabla B-4 y Tabla B-5) muestran en la solución de equilibrio tras la implantación del metro ligero una mejoría bastante notable en varios indicadores para el conjunto del área de estudio. Si se tienen en cuenta los tiempos medios de viaje, la introducción del metro ligero supone una disminución de éstos especialmente en el área urbana de Santander, donde todos los modos incluidos el transporte privado y el autobús presentan reducciones de más del 10%. Sin embargo el aumento de los tiempos medios de viaje en el modo bus interurbano señala que fuera del núcleo de Santander pueden producirse situaciones de congestión, algo que se analizará con más profundidad al examinar la asignación a la red vial.

Las velocidades medias de desplazamiento experimentan así mismo un incremento especialmente dentro del transporte privado y los buses. La velocidad media del modo metro – cercanías sufre una ligera disminución, cercana al 2%, algo lógico ya que el metro ligero en comparación con los trenes de cercanías presenta velocidades comerciales algo inferiores.

Si se tienen en cuenta los tiempos de transbordo es notable como la introducción del metro ligero conlleva una reducción considerable de éstos, si bien también hay que tener en cuenta que el número de transbordos entre el modo metro – cercanías y el modo bus (urbano) experimenta una reducción importante debida a que gran parte de estos viajes combinados se realizan ahora usando exclusivamente la red férrea (cercanías + metro ligero). Analizando los tiempos de acceso, puede observarse como éstos se incrementan especialmente en el modo metro – cercanías (hasta algo más del 30%) lo cual es seguramente un fenómeno derivado de su mayor utilización incluso en zonas que no presentan una parada suficientemente cercana.

Modo	Tiempo Acceso (min)	Dist Acceso (km)	Tiempo Transb (min)	Tiempo Viaje (min)	Dist Viaje (km)	Tiempo Espera (min)	Veloc Media (km/h)
Transporte Privado	-	-	-	7.78	5.24	-	40.41
Bus Urbano	3.95	0.32	0	3.88	1.92	6.35	29.64
Metro-cercanías	7.75	0.57	2.49	5.31	4.24	6.17	47.87
Bus interurbano	6.84	0.97	0	5.84	3.32	5.79	34.17
Metro – Bus urbano	6.24	0.72	2.16	9.31	6.1	16.03	39.34

Tabla B-3. Indicadores del nivel de servicio medio de los distintos modos considerados en el Área Metropolitana de Santander. Solución de equilibrio. Modelo ESTRAUS

Modo	Tiempo Acceso (min)	Dist Acceso (km)	Tiempo Transb (min)	Tiempo Viaje (min)	Dist Viaje (km)	Tiempo Espera (min)	Veloc Media (km/h)
Transporte Privado	-	-	-	-1.56	-0.4	-	4.16
Bus Urbano	-0.12	-0.03	0	-0.44	-0.12	-0.29	1.26
Metro-cercanías	1.86	0.23	0.02	-2.86	-2.41	0.25	-0.94
Bus interurbano	0.21	0.15	0	0.34	0.44	-0.44	2.7
Metro – Bus urbano	-0.44	0.05	-0.93	0	1.62	2	10.48

Tabla B-4. Diferencias entre los indicadores de los niveles del año base y la solución de equilibrio aportada por el modelo ESTRAUS.

Modo	Tiempo Acceso (min)	Dist Acceso (km)	Tiempo Transb (min)	Tiempo Viaje (min)	Dist Viaje (km)	Tiempo Espera (min)	Veloc Media (km/h)
Transporte Privado	-	-	-	-16.70	-7.09	-	11.48
Bus Urbano	-2.95	-8.57	0.00	-10.19	-5.88	-4.37	6.86
Metro-cercanías	31.58	67.65	0.81	-35.01	-36.24	4.22	-1.93
Bus interurbano	3.17	18.29	0.00	6.18	15.28	-7.06	8.58
Metro – Bus urbano	-6.59	7.46	-30.10	0.00	36.16	14.26	36.31

Tabla B-5. Porcentaje de cambio entre los indicadores de los niveles del año base y la solución de equilibrio aportada por el modelo ESTRAUS

En cuanto a la asignación a la red (véase Fig B-5), en general se detecta una bajada de flujos en gran parte de las vías entre el año base y la solución de equilibrio aportada por el modelo. Sin embargo en algunos de los ejes de articulación principal como la S – 10 y la A – 8 puede detectarse también un incremento de los flujos lo cual puede producir congestión ya que son vías con un fuerte uso ya en el año base. Las mayores diferencias en la asignación se dan además dentro del núcleo urbano de Santander con una disminución en los flujos en determinadas zonas de la ciudad (barrio Castilla – Marqués de la Hermida, Sardinero) y un aumento en la zona centro y otros ejes longitudinales como General Dávila y transversales como Camilo Alonso Vega.

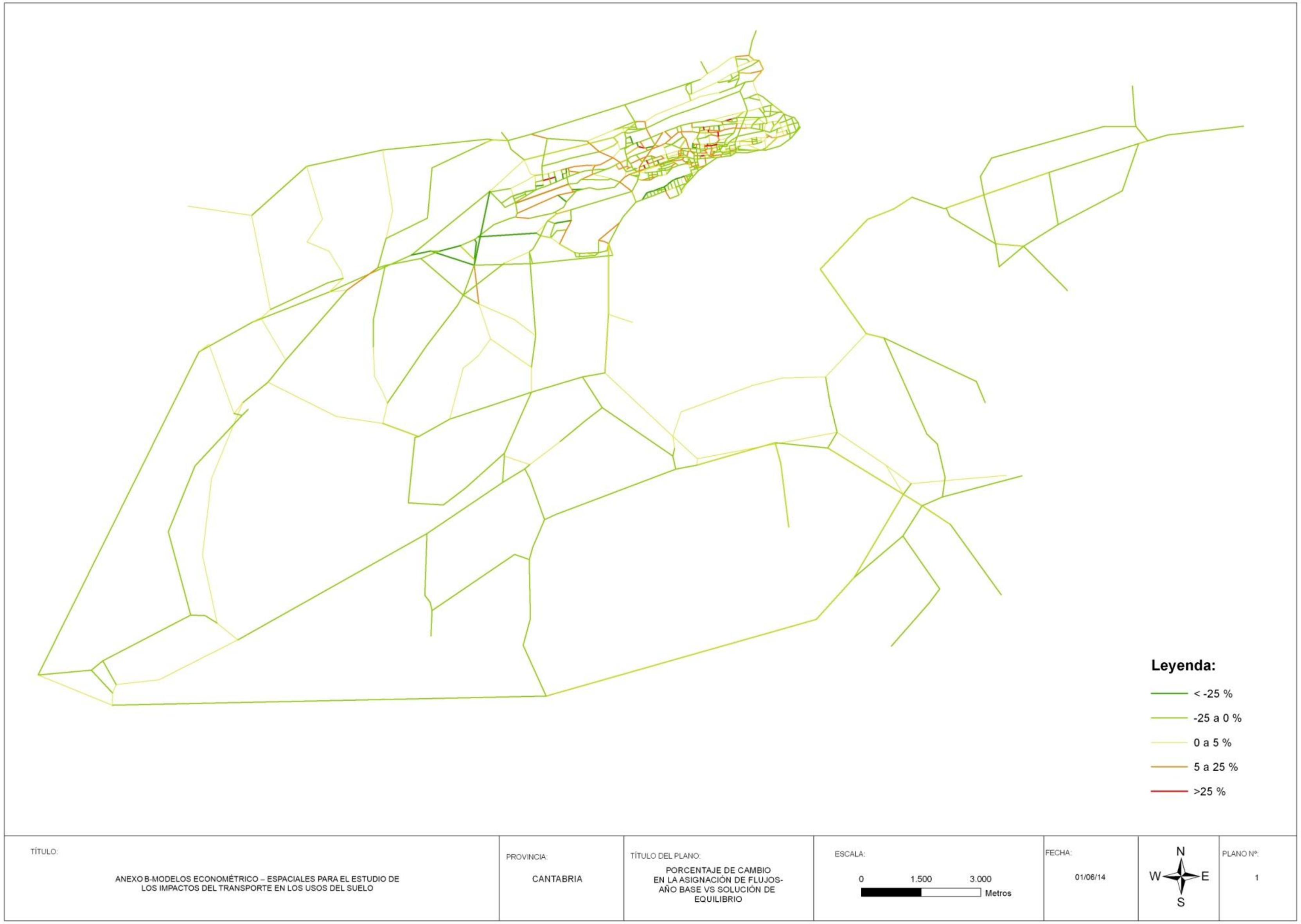

Fig B-5. Porcentaje de cambio en la asignación de viajes en los arcos de la red vial del Área Metropolitana de Santander. Año base vs Solución de equilibrio

B.3 Repercusión del nuevo modo de transporte en la localización Residencial

Para el análisis de los cambios en la localización residencial y de actividades en el área de estudio se examinará el porcentaje de cambio en el número de residentes/empleos en las 42 zonas de uso del suelo. Así mismo se ha dividido el área en 5 grandes zonas que agrupan las 42 áreas de uso del suelo para mejorar la legibilidad de los resultados y captar de forma más directa los patrones de cambio experimentados por el sistema territorial. Las cinco grandes macro – áreas han quedado definidas tal y como se presenta en la Tabla B-6.

Denominación Macro – Área	Zonas que la forman
1. Centro Santander	1 y 2
2. Barrios Santander	3 a 8 y 11 a 21 y 23, 24
3. Periferia Municipio Santander	9, 10, 22,25 y 26
4. Primera Corona Metropolitana	27 a 35
5. Segunda Corona Metropolitana	36 a 42

Tabla B-6. Agrupación de las áreas de uso del suelo en Macro - Áreas

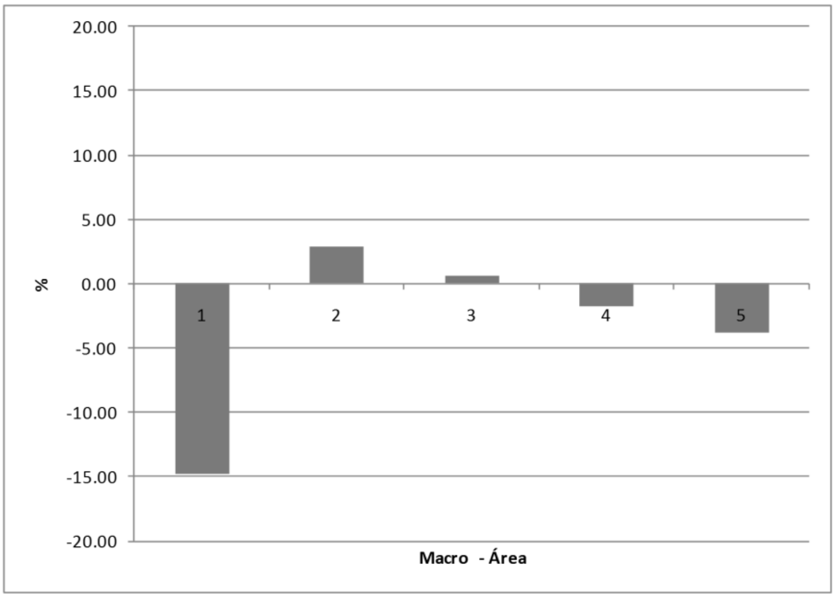

Fig B-6. Porcentaje de cambio en la población residencial (Macro – Áreas)

En general puede afirmarse, según los resultados de las simulaciones (véase Fig B-6), que la implantación del metro ligero supondría un cambio significativo en la localización de la población a medio – largo plazo. El centro urbano sería el área más perjudicada seguramente porque la implantación del metro ligero no implicaría una mejora significativa en términos de su accesibilidad a los centros de trabajo. De hecho la propia caída de población podría llevar aparejada también una caída en la localización de actividades comerciales y de servicios (véase apartado siguiente). En cambio las zonas 2 y 3 pertenecientes al municipio de Santander y especialmente la primera de ellas, es decir, la correspondiente a los barrios que rodean al centro de la ciudad, podrían verse beneficiadas por una mayor demanda de localización (incrementos de 2.9% y 0.66% respectivamente). Por otro lado las zonas 4 y 5 correspondientes al área metropolitana también podrían ver reducida en cierta medida su población, concretamente en algo más de un 1.5 % y 3.5 % respectivamente.

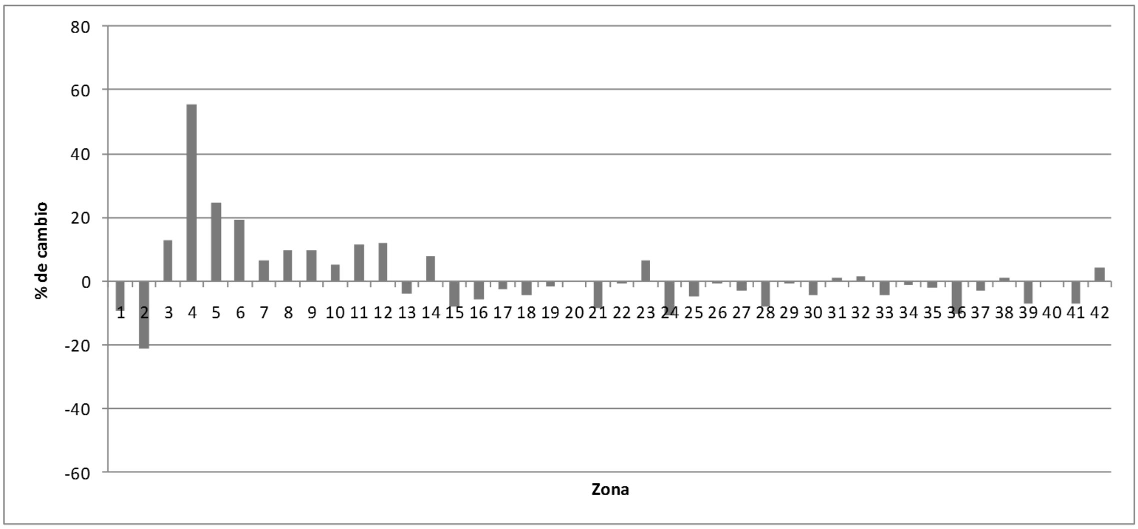

Fig B-7. Porcentaje de cambio en la población residencial (Zonas de Uso del Suelo)

Desagregando los datos por las zonas de uso del suelo (véase Fig B-7 y Fig B-10) puede observarse como las zonas que mejorarían en mayor medida su utilidad de cara a la localización de población serian todas las comprendidas entre el área 3 y el 12 y especialmente las áreas 4, 5 y 6 que corresponden con la zona este y noreste de la ciudad (Barrios de El Sardinero, Menéndez Pelayo y General Dávila en su tramo inicial). Estas áreas, como se verá posteriormente, son las que presentan mayores incrementos en la accesibilidad tras la implantación del metro ligero. Actualmente siendo áreas que se encuentran plenamente integradas en la ciudad, presentan una menor accesibilidad a los centros de actividad y comercio dada su especialización funcional en el sector residencial y debido a la menor presencia de oferta de transporte público.

Estos datos también pueden desagregarse según la clase socioeconómica de los residentes (véase Fig B-8 y Fig B-9). En general puede decirse que los patrones de cambio son muy similares en ambas clases socioeconómicas, si bien los cambios porcentuales en los residentes de clase alta son más acusados.

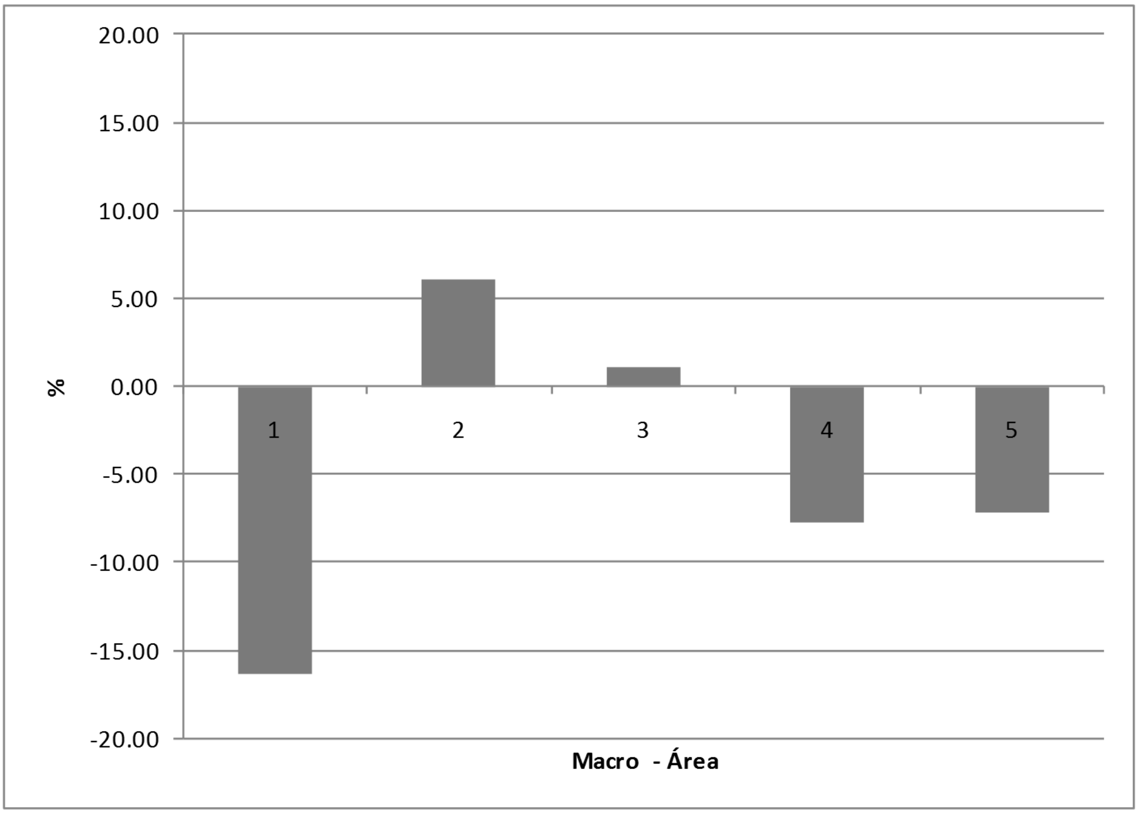

Fig B-8. Porcentaje de cambio en la población residencial de clase alta (Macro – Áreas)

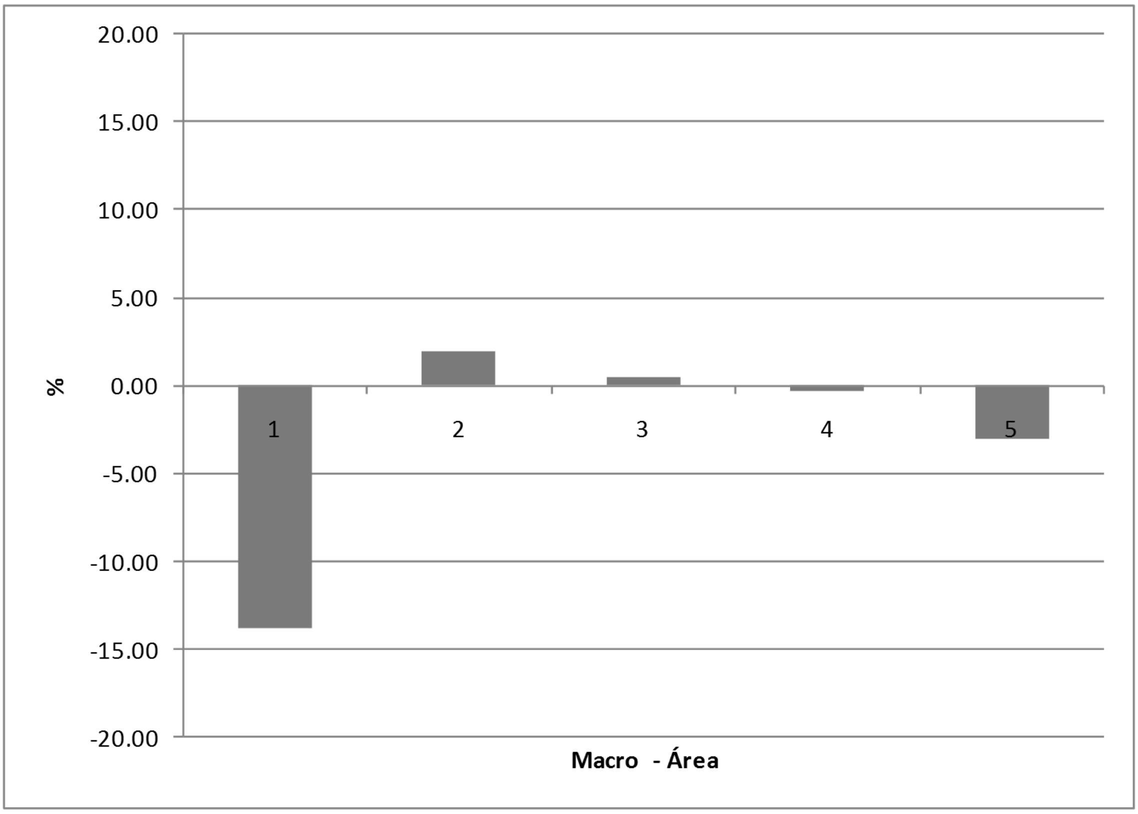

Fig B-9. Porcentaje de cambio en la población residencial de clase media y baja (Macro – Áreas)

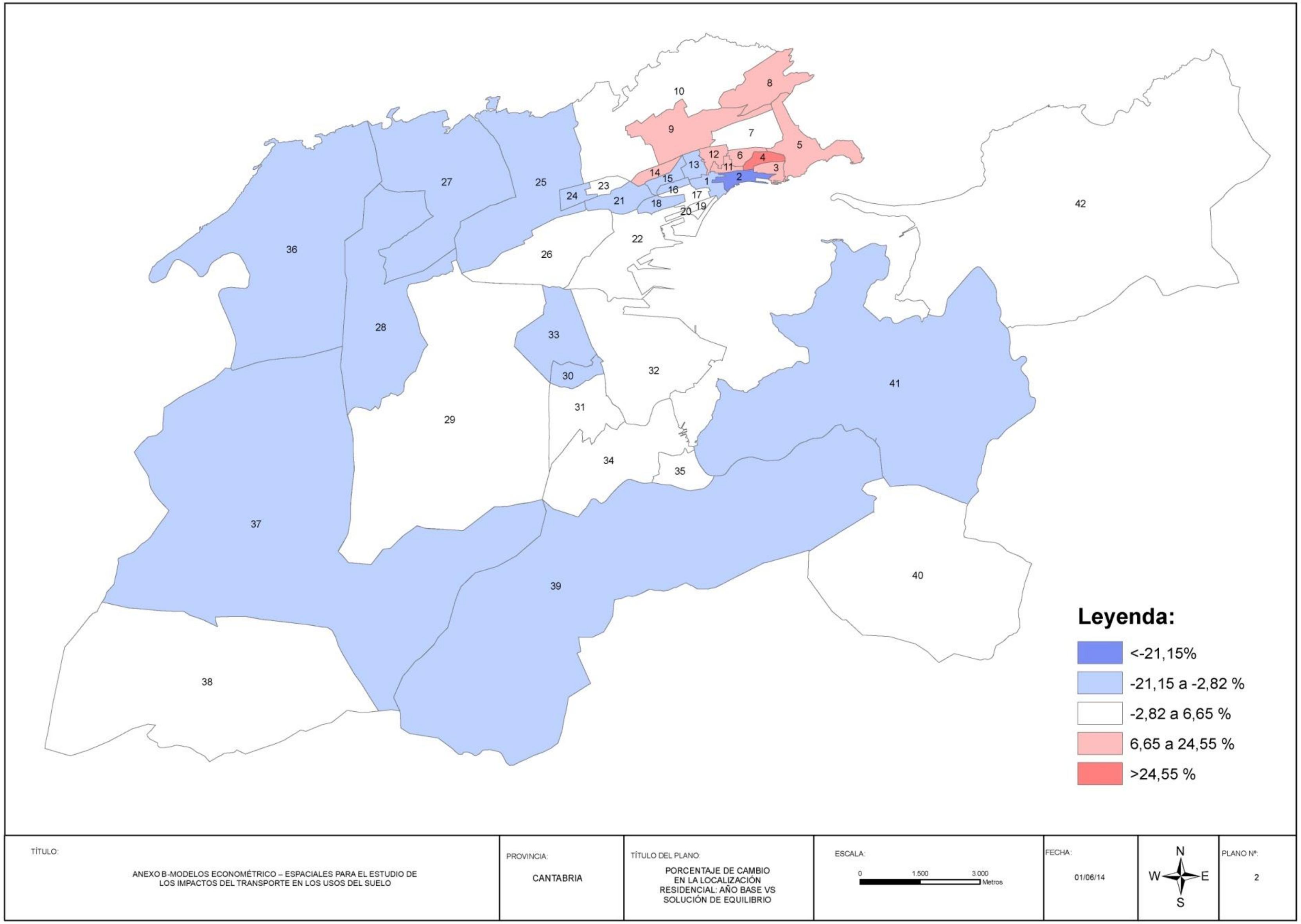

Fig B-10. Porcentaje de cambio en la localización residencial: año base vs solución de equilibrio

B.4 Repercusión del nuevo modo de transporte en la localización de actividades económicas

Los cambios en la localización de actividades económicas entre el año base y los resultados de la simulación reflejan cierta descentralización, especialmente a favor de los barrios de la ciudad de Santander y en menor medida del resto de localidades que forman el área metropolitana (véase Fig B-11). En el caso del centro de la ciudad la disminución en el número de empleos podría ser de algo más de un 3% mientras que en la macro – área 2 podría darse un incremento de hasta el 2.2%. Las macro – áreas 3 y 4 prácticamente no presentarían cambios en la localización de empleos, mientras que el área más periférico del sistema podría incrementar su número de empleos ligeramente (en torno al 0.25%).

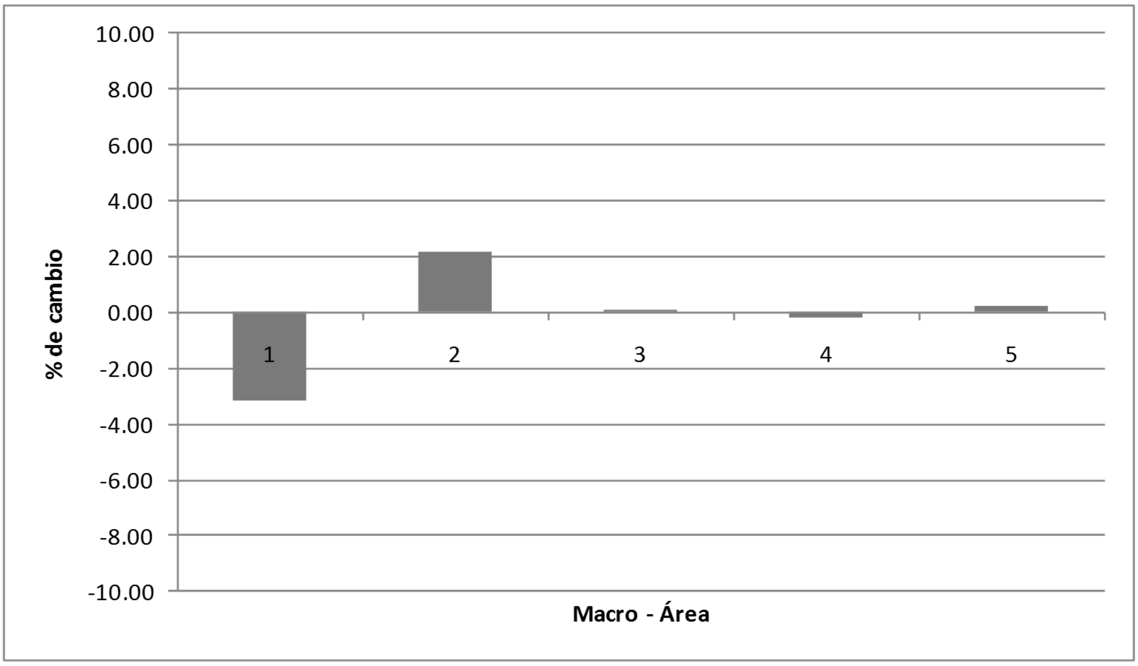

Fig B-11. Porcentaje de cambio en la localización de actividades (Macro – Áreas)

Si se analizan los cambios en las 42 zonas de uso del suelo (véase Fig B-12 y Fig B-15) puede detectase el patrón de cambio en la distribución de los empleos. Mientras que la zona 1 y especialmente la zona 2 experimentan una caída en la localización de actividades, las zonas 3 a 12 del sector noreste de la ciudad presentan ganancias de hasta el 13%. En general puede observarse como el patrón de relocalización de actividades es muy similar a los cambios ya señalados en la distribución de la población algo lógico si se considera que el modelo de localización residencial y el modelo de localización de actividades están conectados siendo muy dependientes uno del otro.

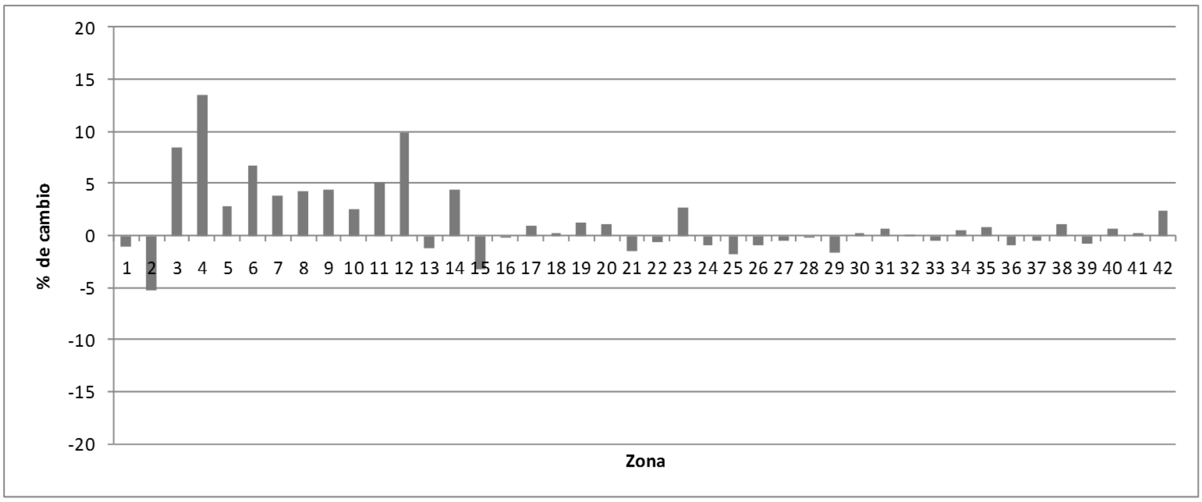

Fig B-12. Porcentaje de cambio en la localización de actividades (Zonas de Uso del Suelo)

La tendencia detectada en los cambios en la distribución de empleos se repite en el sector minorista (véase Fig B-13) aunque de forma más acusada ya que el centro urbano muestra una deslocalización de prácticamente el 13% de sus empleos. Nuevamente las zonas más beneficiadas de esta descentralización son los barrios que rodean el centro urbano de la zona noroeste con un aumento en el número de empleos comerciales de 7.5%.

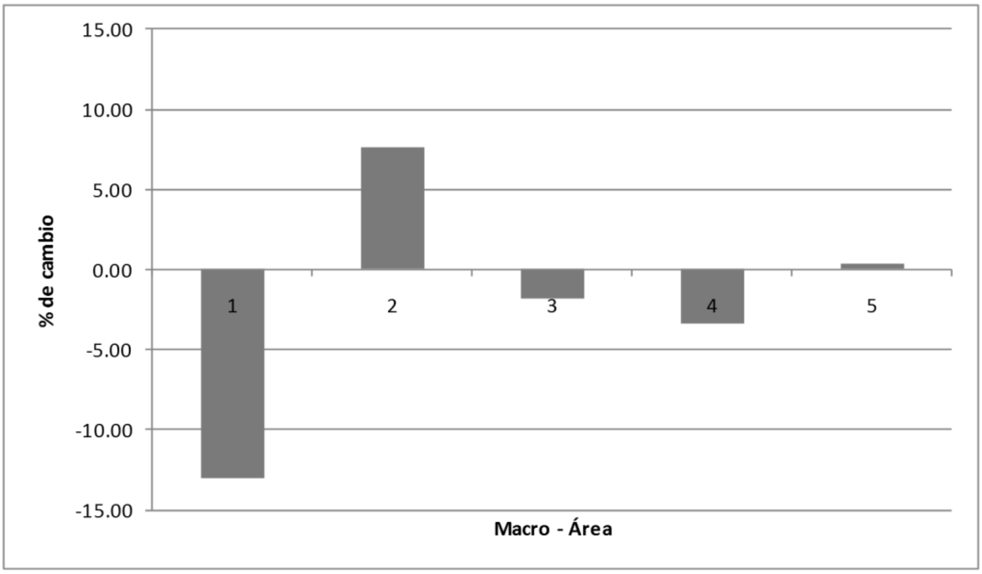

Fig B-13. Porcentaje de cambio en la localización de actividades comerciales (Macro – Áreas)

El sector servicios (véase Fig B-14) demuestra ser sin embargo menos sensible a los impactos provocados por la implantación del metro ligero. La caída simulada del número de empleos en el centro urbano es de algo más del 3%, porcentaje similar al incremento experimentado en el número de empleos en la macro – área 2. Además en este caso la descentralización de las actividades de servicios se produciría por todo el área metropolitana a diferencia del sector comercial que tendería a crecer sobre todo en las áreas alrededor del centro urbano. De hecho la macro – área 5 y más concretamente los municipios de Astillero, Marina de Cudeyo y Ribamontán al Mar el crecimiento modelado alcanza porcentajes superiores al 1.5%.

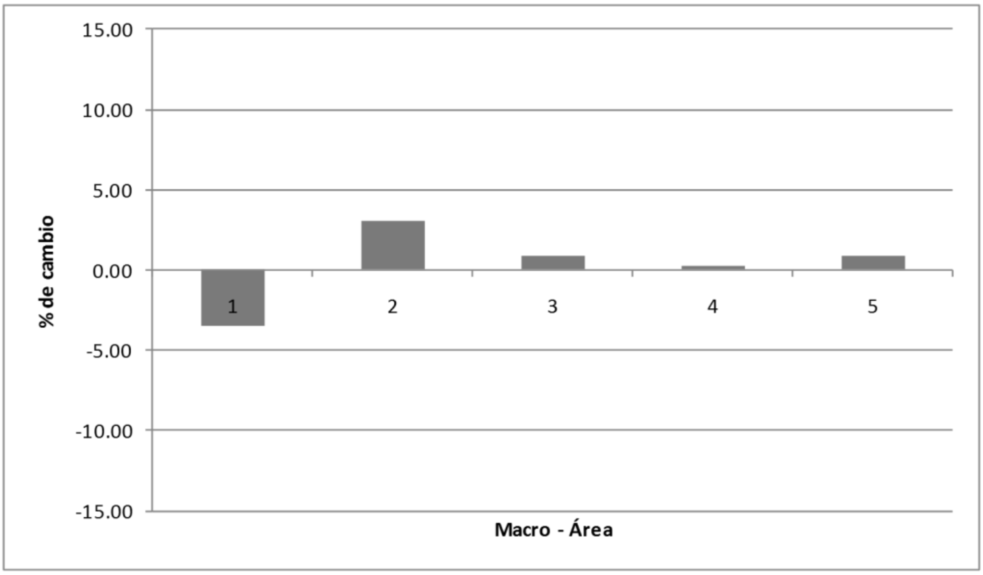

Fig B-14. Porcentaje de cambio en la localización de actividades de servicios (Macro – Áreas)

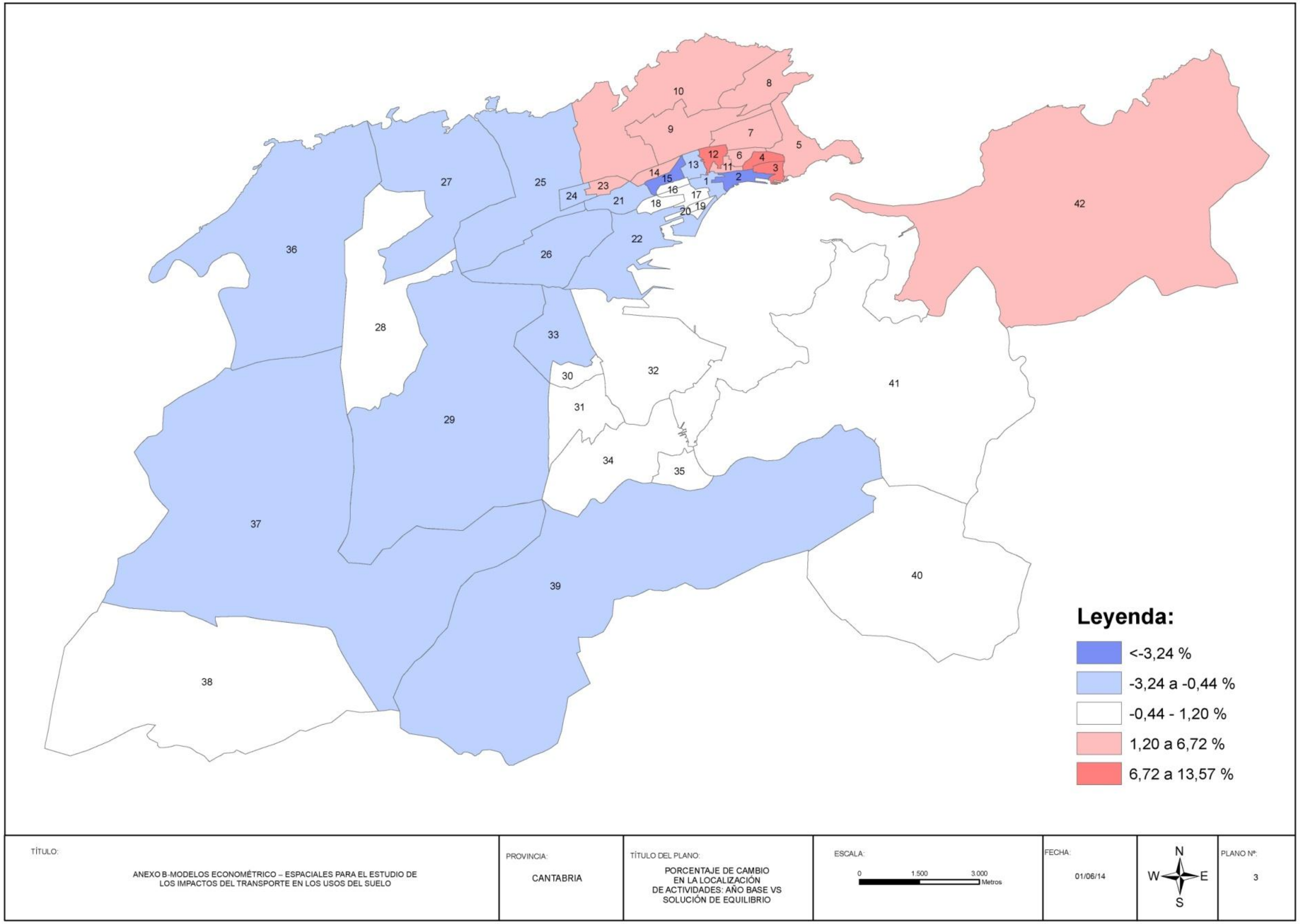

Fig B-15. Porcentaje de cambio en la localización de actividades económicas: año base vs solución de equilibrio

B.5 Repercusión del nuevo modo de transporte en los precios inmobiliarios zonales

El impacto de un nuevo modo de transporte en los precios inmobiliarios es un tema fundamental ya que plantea la pregunta de quién gana y quién pierde como consecuencia de una medida de este tipo. Además es de interés para la administraciones públicas ya que éstas pueden recuperar en parte la capitalización de los beneficios de las actuaciones a través de impuestos a los valores inmobiliarios (Ortúzar y Willumsen, 2001).

El modelo especificado estima estos cambios a través de la técnica de la regresión hedónica (Malpezzi, 2008). En la se recogen los cambios porcentuales por Macro – Áreas. Según el modelo únicamente el área 2 experimentaría una ligera subida en los valores inmobiliarios derivada fundamentalmente de la entrada en servicio del metro ligero lo que implicaría menores tiempos de viaje al centro urbano y menores tiempos de espera al transporte público. En cambio el resto de áreas y especialmente las más periféricas del sistema metropolitano podrían experimentar caídas en los valores inmobiliarios de más de un 10%. Desagregando los datos por la zonificación de uso del suelo (véase Fig B-17 y Fig B-18), puede observarse como en el área centro la caída de los precios inmobiliarios se da sobre todo en la zona 2. De la zona 3 a la zona 9 hay un incremento notable de los valores inmobiliarios. Esto es especialmente cierto en el área 9 (con un aumento de más del 15%) correspondiente al barrio de Valdenoja II actualmente en proceso de desarrollo urbanístico y con un servicio escaso de transporte público.

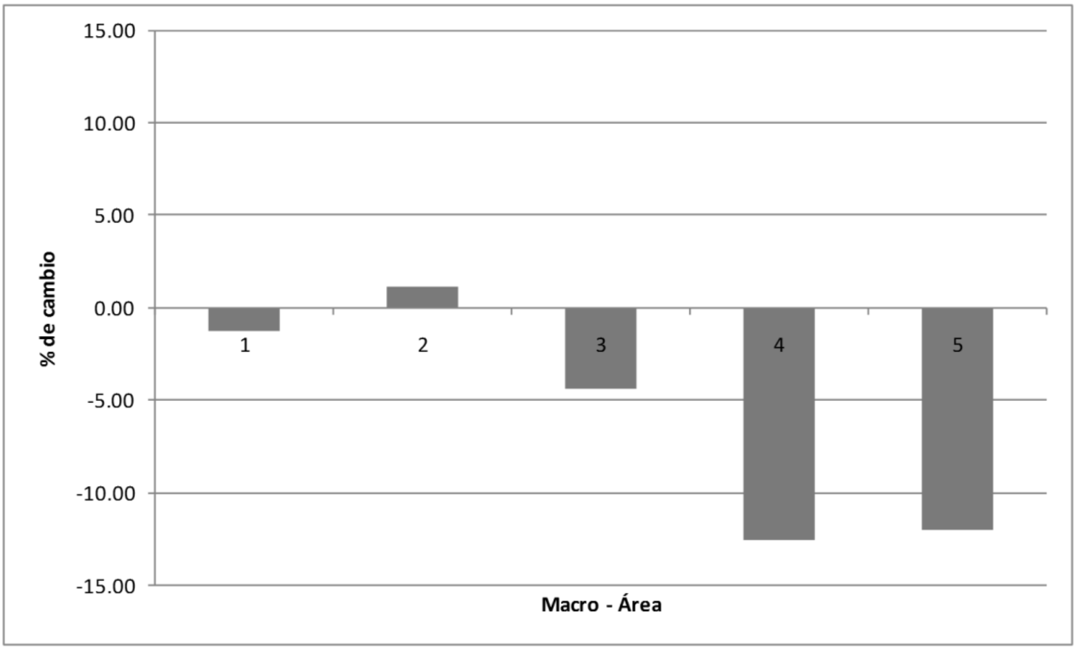

Fig B-16. Porcentaje de cambio en los precios medios inmobiliarios (Macro – Áreas)

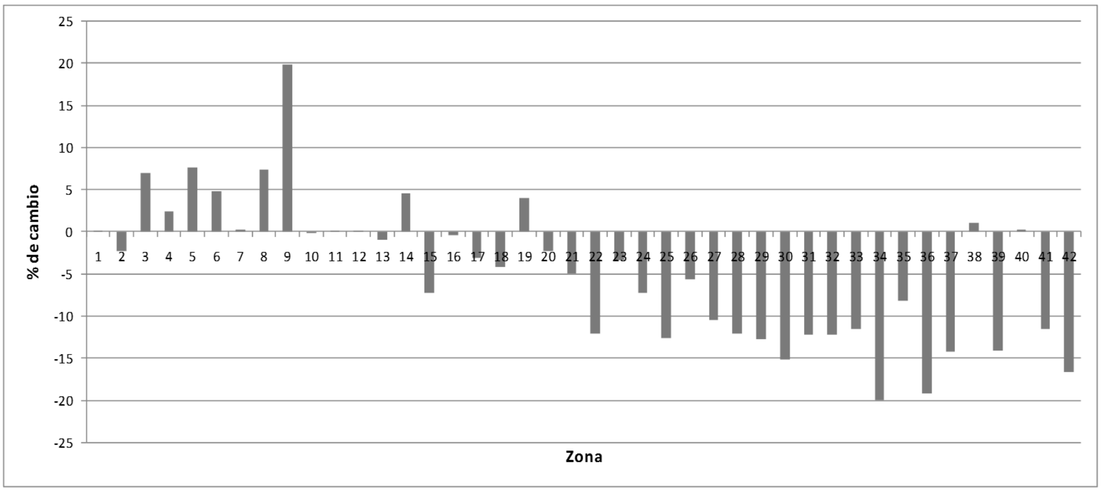

Fig B-17. Porcentaje de cambio en los precios medios inmobiliarios (Zonas de Uso del Suelo)

A partir del área 20 y, especialmente, a partir del área 26, es decir, fuera del municipio de Santander, se producen descensos generalizados en los precios medios inmobiliarios (en algunos casos superiores al 20%). El factor principal de estos descensos es el aumento de los tiempos de viaje al centro metropolitano. A pesar de que el metro ligero supone una mejora en los tiempos de viaje y espera para los habitantes de la capital, sobre todo en aquellos barrios que bordean el centro en su sector este, esto no es ciertamente así para los habitantes de otros núcleos del área metropolitana. De hecho, como ya se ha citado, el aumento de los flujos en ciertas vías principales de articulación del sistema (S – 10 y A – 8 fundamentalmente) podría generar congestión lo que a su vez se traduciría en mayores tiempos de viaje y por lo tanto en pérdidas de accesibilidad desde ciertas zonas (véase el apartado siguiente). Estas pérdidas de accesibilidad debidas a los mayores tiempos de viaje provocarían una disminución de los precios medios inmobiliarios en áreas como Astillero, Ribamontán al Mar o el norte del municipio de Piélagos. Estos efectos señalan que si el metro ligero provocara una relocalización de actividades a favor del centro y de los barrios colindantes, algo a priori positivo desde el punto de vista de limitar un proceso excesivo de metropolización, podrían generarse situaciones de congestión excesiva o de deseconomías de escala (Camagni, 2005) que no sólo podrían afectar al propio núcleo de Santander sino sobre todo a aquellas áreas funcionalmente vinculadas con él por flujos hogar – trabajo.

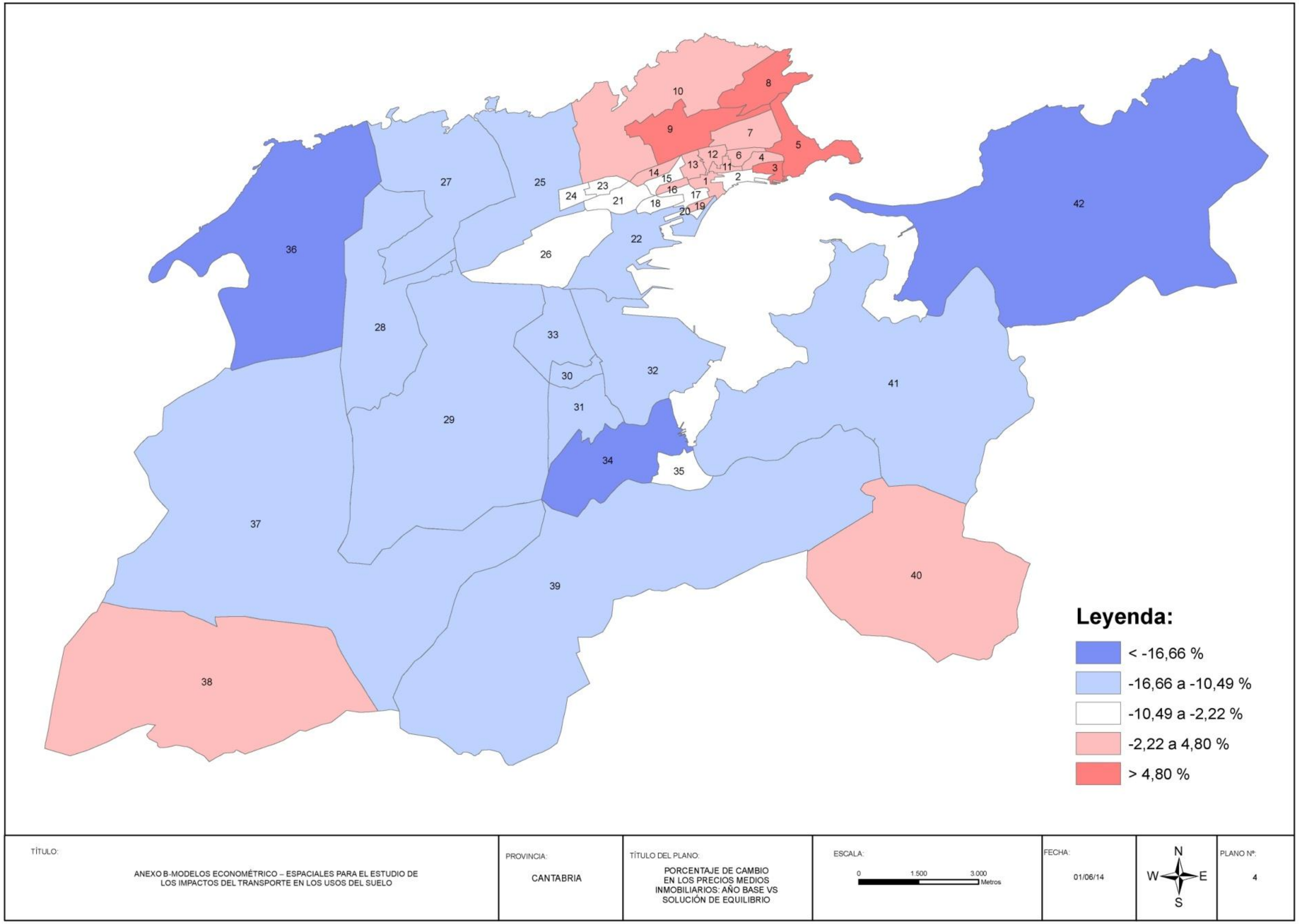

Fig B-18. Porcentaje de cambio en la predicción de precios inmobiliarios: año base vs solución de equilibrio

B.6 Repercusión del nuevo modo de transporte en la Accesibilidad Zonal

La implantación del metro ligero en la ciudad de Santander puede traer cambios en la accesibilidad de las distintas zonas del área metropolitana. Estos cambios pueden deberse tanto a la relocalización de actividades y población como a la variación de los costes de viaje entre zonas.

Examinando en primer lugar los cambios experimentados por las distintas zonas en su accesibilidad pasiva (véase Fig B-19 y Fig B-22) puede afirmarse que las áreas más beneficiadas por la implantación del metro ligero serian nuevamente los barrios que rodean al centro de la ciudad en su sector noreste (áreas 3 a 9). En cambio las zonas en el sector este se verían menos beneficiadas y especialmente las áreas fuera de la capital podrían experimentar disminuciones en la accesibilidad pasiva de hasta el 30%. Estas pérdidas de accesibilidad son debidas conjuntamente a los dos factores antes señalados. La relocalización de población a favor del centro urbano y la mayor congestión en las vías principales del área metropolitana, provocaría pérdidas de accesibilidad importantes especialmente en municipios como Piélagos, Villaescusa o Marina de Cudeyo. Algo similar ocurriría también en el centro de Santander donde también se da cierto descenso de población y un aumento de los tiempos de acceso a otras zonas. Así pues, desde el punto de vista de la accesibilidad pasiva y, por lo tanto, de la facilidad de la población para alcanzar estas zonas, el metro ligero beneficiaria fundamentalmente al sector este de la ciudad.

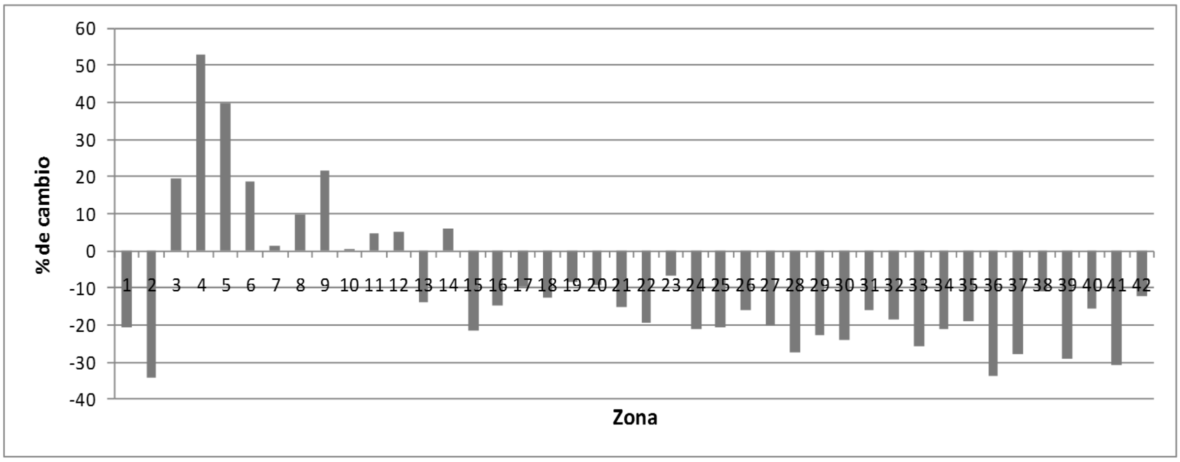

Fig B-19. Porcentaje de cambio en la accesibilidad pasiva (Zonas de Uso del Suelo)

El patrón de cambio en la accesibilidad activa (véase Fig B-20 y Fig B-21), es decir, de facilidad de alcance de empleos, es muy similar al de la accesibilidad pasiva sin bien con cambios porcentuales algo más moderados. Al igual que en el caso anterior, la implantación del metro ligero y la relocalización de actividades provocada por éste, beneficiaria sobre todo a las zonas del área este de la ciudad (3 a 12) mientras que las zonas del sector oeste y especialmente los municipios colindantes como Astillero o Piélagos, perderían en cierta medida accesibilidad a empleos.

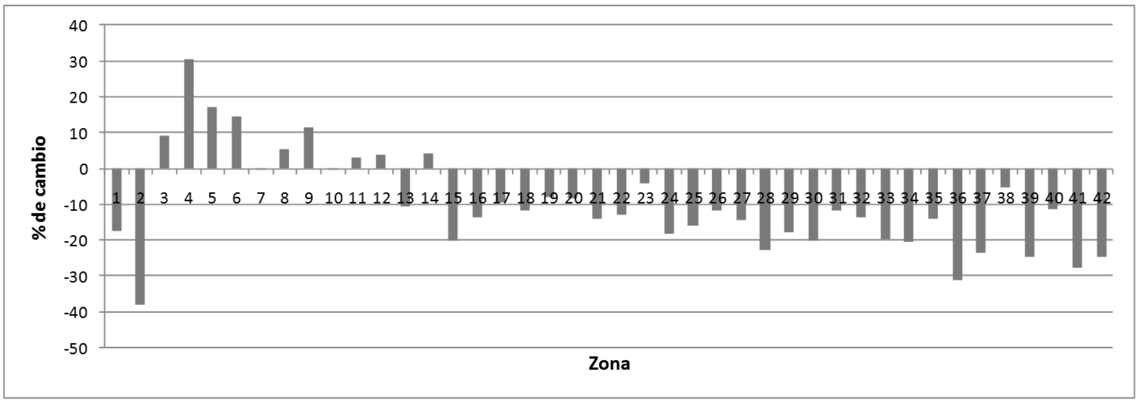

Fig B-20. Porcentaje de cambio en la accesibilidad activa (Zonas de Uso del Suelo)

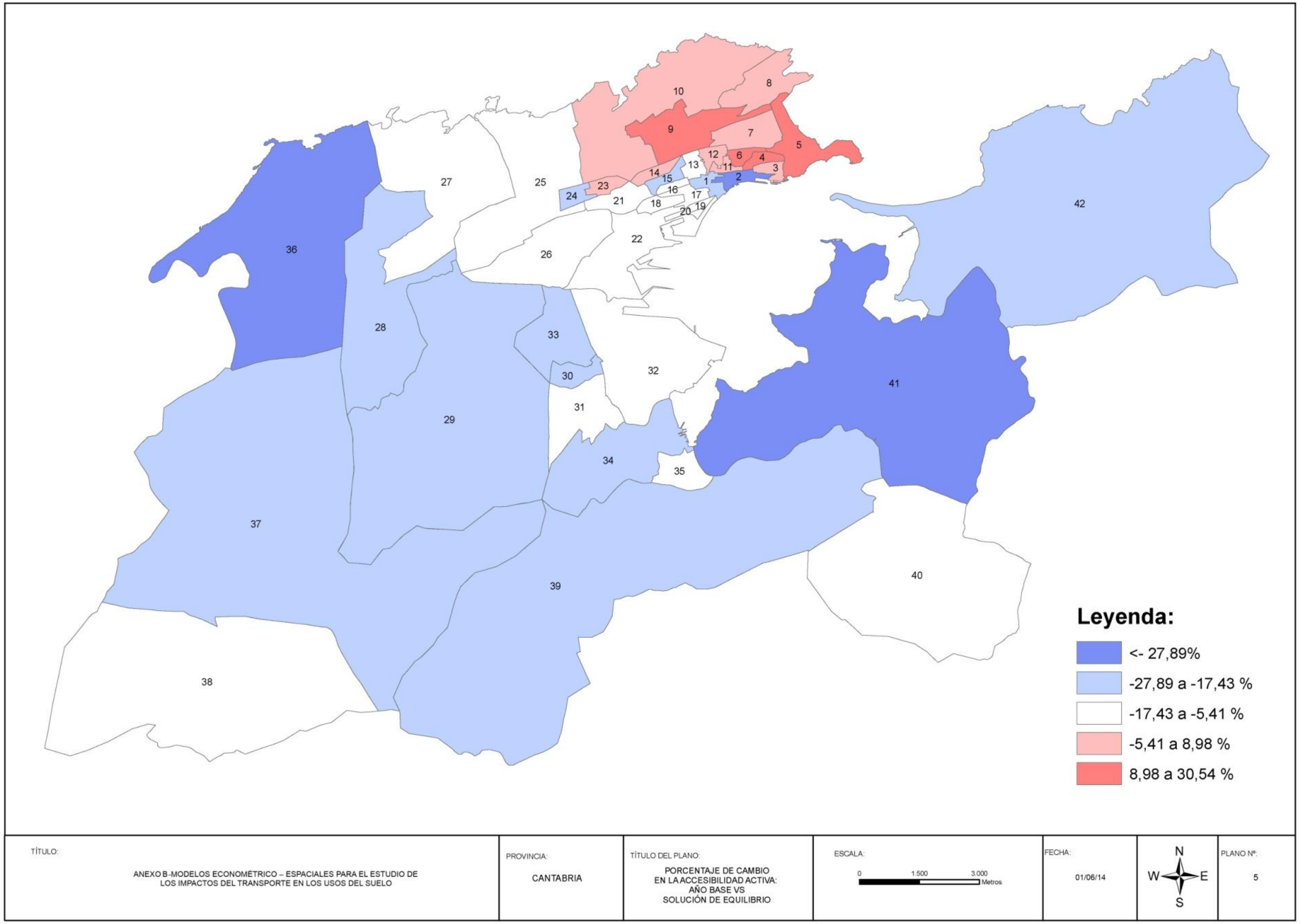

Fig B-21. Porcentaje de cambio en la accesibilidad activa: año base vs solución de equilibrio

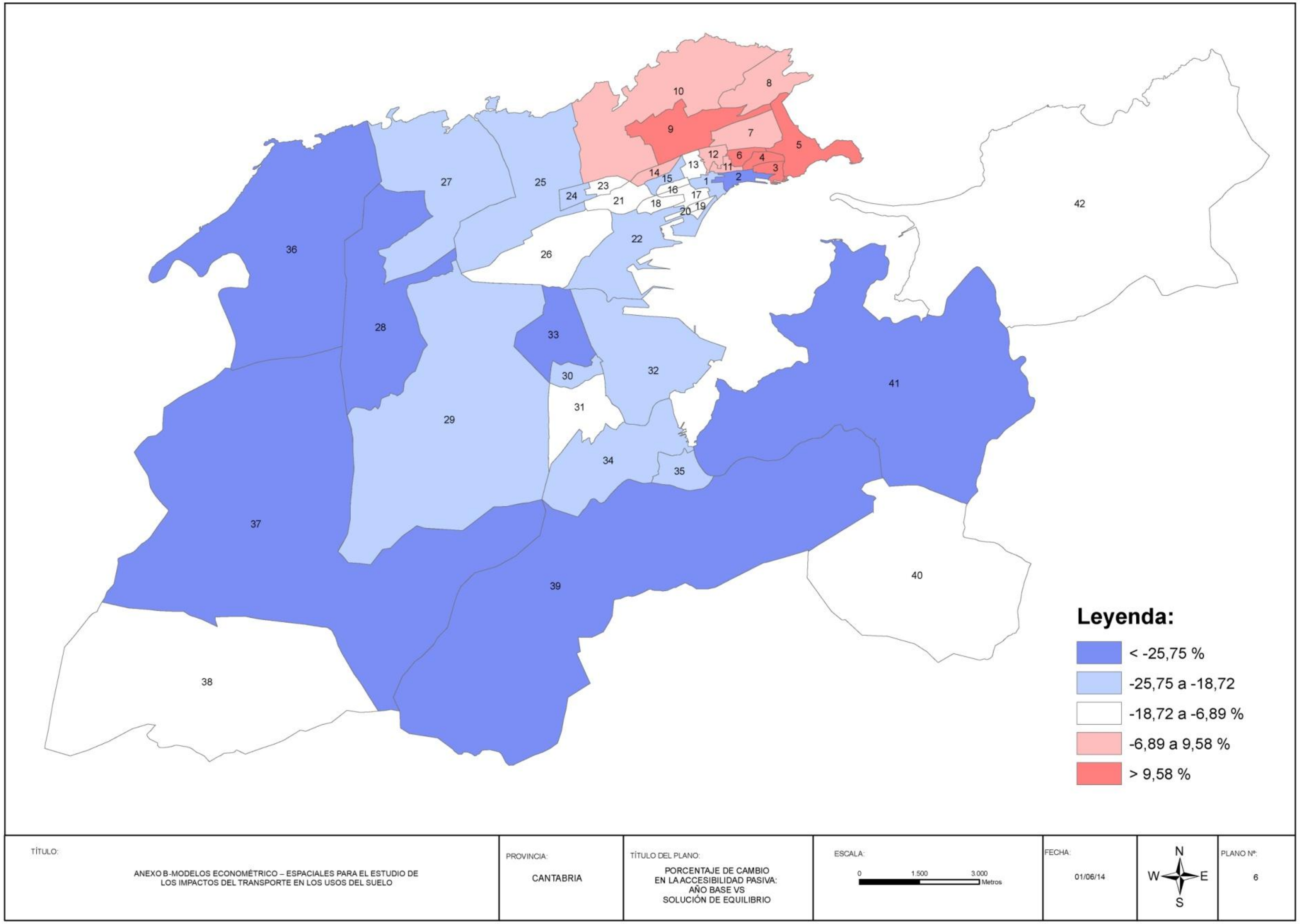

Fig B-22. Porcentaje de cambio en la accesibilidad pasiva: año base vs solución de equilibrio

B.7 Conclusiones

En la Tabla B-7 puede verse un resumen de algunos de los cambios más relevantes simulados por el modelo LUTI tras la implantación del metro ligero.

Ámbito de análisis	Efectos detectados
Partición modal	• Fuerte incremento en la elección modal del modo metro – cercanías y en menor medida del bus interurbano. • Decrecimiento del uso del transporte privado y del bus urbano.
Niveles de Servicio y asignación a la red	• Aumento de las velocidades medias en todos los modos (excepto en el metro – cercanías). • Reducción de los tiempos medios de viaje dentro del área de Santander en todos los modos. Aumento de los tiempos de viaje en el modo bus interurbano y por lo tanto en las coronas metropolitanas. • Reducción de los flujos en gran número de vías dentro del área urbana de Santander. • Incremento de flujos y cierto grado de congestión en vías de articulación metropolitana como la S – 10.
Localización de la población	• Reducción de la población en la zona centro (zona 2 especialmente). • Incremento de la población en el sector este y noreste. • Reducción moderada de la población en diversos municipios del área metropolitana.
Localización de actividades económicas	• Reducción del número de empleos de comercios y servicios en el centro de la ciudad. • Incremento del número de empleos en ambos sectores económicos en el sector nororiental de la capital. • Incrementos moderados de localización de actividades (especialmente en el sector servicios) en la zona este del área metropolitana (municipios de Astillero, Marina de Cudeyo y Ribamontán al Mar).
Precios inmobiliarios	• Aumentos notables de precios medios inmobiliarios en el sector noreste de la capital. • Reducciones de precios medios en las zonas de la parte oeste de la ciudad de Santander y especialmente en el resto del área metropolitana.
Accesibilidad activa	• Incremento de la accesibilidad activa (a empleos) en el sector noreste de la capital. • Reducción de la accesibilidad activa en el centro de la ciudad. • Fuerte reducción de la accesibilidad en diversos municipios del área metropolitana por efecto del nuevo patrón de localización de actividades y de los mayores tiempos de viaje.

Accesibilidad pasiva	• Incremento de la accesibilidad pasiva en el sector noreste de la capital. • Reducción de la accesibilidad pasiva en el centro de la ciudad. • Fuerte reducción de la accesibilidad pasiva en diversos municipios del área metropolitana.

Tabla B-7. Resumen de los efectos simulados por la implantación del metro ligero en el municipio de Santander

Según los efectos detectados, las áreas más afectadas por la implantación del metro ligero serian el centro urbano y el sector noreste de la ciudad de Santander. Las mejoras de accesibilidad, tanto activa como pasiva, experimentadas por los barrios orientales al centro como Menéndez Pelayo, Sardinero y Valdenoja podrían provocar que hogares y empresas comerciales y de servicios tendieran a preferir estas localizaciones para establecerse. Esto podría suponer cierto procedo de desconcentración de empleo y población desde el centro de la ciudad hacia esas zonas.

Este proceso podría ir así mismo acompañado de cierta tendencia a revertir, o al menos limitar en cierta medida, el proceso de dispersión de población y actividades desde el núcleo santanderino a los municipios cercanos.

Desde el punto de vista de los valores inmobiliarios, éste patrón de relocalización de actividades y población se vería acompañado de un patrón similar de incrementos y decrementos de precios medios en las zonas del área metropolitana. Si bien el modelo hedónico especificado no es propiamente un modelo de simulación del mercado inmobiliario, estas alzas en los precios medios son coherentes con los cambios detectados en el patrón de localización de residentes y empleos. Es válido suponer que una mayor demanda de localización debería provocar un crecimiento de los precios independientemente de los cambios detectados en la oferta de transporte de cada zona.

En cuanto a los cambios provocados por el nuevo modo en el sistema de transporte, el submodelo ESTRAUS simula una captación por parte del metro ligero de algo más del 4% de los viajes en hora punta de mañana. Paradójicamente el mayor uso de un modo de transporte público como es el metro ligero y el cambio del patrón de población y actividades por él provocado, podría incrementar la congestión en el área metropolitana y especialmente en los ejes de articulación principales actualmente ya con notables problemas (vías de acceso a Santander como la S – 10 especialmente). Este fenómeno podría reducir las condiciones de accesibilidad de varios de los municipios que rodean la capital y en definitiva podría generar deseconomías de escala perjudiciales para el conjunto del área metropolitana. Por lo tanto la implantación del metro ligero en el caso de producir efectos de relocalización de población y actividades a favor de la capital debería estar acompañada de un plan de movilidad a escala metropolitana que tuviera en cuenta estos efectos y los mitigara en la medida de lo posible.

Finalmente hay que detallar una serie de precauciones sobre los resultados simulados por el modelo ante la implantación del nuevo modo de transporte sostenible. Como ya se ha expuesto (véase apartado 2 del cuerpo principal de la tesis) el modelo de uso del suelo parte de una serie de hipótesis de carácter teórico y práctico lo que condiciona las interpretaciones de los resultados de las simulaciones. En primer lugar, el modelo considera el sistema urbano analizado como un sistema cerrado, con lo que es difícil plantear los efectos de atracción de población y actividades desde fuera del área de estudio o, al contrario, de pérdida de población y actividades provocadas por la implantación de una medida.

En segundo lugar, la localización de las actividades económicas clasificadas como pertenecientes al sector básico se considera exógena al modelo al no depender tan fuertemente de las condiciones de localización. Sin embargo diversos agentes económicos y, nuevamente, los poderes públicos pueden favorecer el establecimiento de equipamientos y empresas en determinadas zonas del sistema urbano modificando así fuertemente las condiciones de accesibilidad y de generación/atracción de viajes.

Por último, el modelo no simula el mercado de suelo al carecer de un submodelo de oferta inmobiliaria. Esto quiere decir que el cambio en los patrones de población y empleos detectados son inicialmente una demanda potencial que puede ser cubierta, o no, por el sector inmobiliario y por los poderes públicos. Estos últimos son especialmente importantes en la dinámica del sistema urbano, debido a su capacidad de intervención tanto activamente como más indirectamente, a través de la planificación urbanística, en el mercado de inmobiliario.

Referencias del Anexo B

Ayuntamiento de Santander (2010a) *3ª Información Pública del Plan General de Ordenación Urbana de Santander*. Santander.

Ayuntamiento de Santander (2010b) *Plan de Movilidad Sostenible del Municipio de Santander*. Santander.

Camagni, R. (2005) Economía Urbana. ed Bosch, A., Barcelona.

EFE (2009) Las obras del metro ligero ascenderían inicialmente a 113 millones de euros. *El Diario Montañes*, Santander.

Fernandez, J. (2007) El Ayuntamiento proyecta un metro ligero con 19 kilómetros de trazado. *El Diario Montañes*, Santander.

Malpezzi, S. (2008) Hedonic Pricing Models: A Selective and Applied Review. *Housing Economics and Public Policy* ed Tony O'Sullivan, K.G., 67-89.

Ortúzar, J.d.D., Willumsen, L.G. (2001) *Modelling transport*, 3rd ed. J. Wiley, Chichester New York.