Universidad de Alcalá
COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2016/17

DOCTORANDO: **SANTOS BARTOLOMÉ, JUAN LUIS**
D.N.I./PASAPORTE: ******4214X**

PROGRAMA DE DOCTORADO: **D331 DOCTORADO EN ECONOMÍA Y DIRECCIÓN DE EMPRESAS**
DEPARTAMENTO DE: **ECONOMÍA Y DIRECCIÓN DE EMPRESAS**
TITULACIÓN DE DOCTOR EN: **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

En el día de hoy 12/01/17, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de **TOMÁS MANCHA NAVARRO // VICENTE ESTEVE GARCÍA.**

Sobre el siguiente tema: *BIG DATA IN MULTIAGENT SYSTEMS: MARKET DESIGN SOLUTIONS*

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL[1] de (**no apto, aprobado, notable y sobresaliente**): Sobresaliente

Alcalá de Henares, 12 de enero de 2017

EL PRESIDENTE	EL SECRETARIO	EL VOCAL
Fdo.: José Carlos Arias	Fdo.: Federico Pablo Martí	Fdo.: Beatriz de Diego

Con fecha 25 de enero de 2017 la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- [x] Conceder la Mención de "Cum Laude"
- [ ] No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada

FIRMA DEL ALUMNO,

Fdo.: Juan Luis Santos Bartolomé

[1] La calificación podrá ser **"no apto" "aprobado" "notable"** y **"sobresaliente"**. El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:

No se produjeron incidencias.

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 25 de enero, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por *SANTOS BARTOLOMÉ, JUAN LUIS*, el día 12 de enero de 2017, titulada *BIG DATA IN MULTIAGENT SYSTEMS: MARKET DESIGN SOLUTIONS*, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve otorgar** a dicha tesis la

MENCIÓN "CUM LAUDE"

Alcalá de Henares, 25 de enero de 2017
EL PRESIDENTE DE LA COMISIÓN DE ESTUDIOS
OFICIALES DE POSGRADO Y DOCTORADO

UNIVERSIDAD DE ALCALÁ · COMISIÓN DE ESTUDIOS OFICIALES DE POSGRADO

Juan Ramón Velasco Pérez

Copia por e-mail a:
Doctorando: SANTOS BARTOLOMÉ, JUAN LUIS
Secretario del Tribunal: FEDERICO PABLO MARTÍ
Directores de Tesis: TOMÁS MANCHA NAVARRO // VICENTE ESTEVE GARCÍA

FACULTAD DE CIENCIAS ECONÓMICAS, EMPRESARIALES Y TURISMO

DEPARTAMENTO DE ECONOMÍA Y DIRECCIÓN DE EMPRESAS

Programa de doctorado en Economía y Dirección de Empresas

BIG DATA IN MULTIAGENTS SYSTEMS: MARKET DESIGN SOLUTIONS

Tesis Doctoral presentada por

JUAN LUIS SANTOS BARTOLOMÉ

Director: Dr. D. TOMÁS MANCHA NAVARRO

Co-Director: Dr. D. VICENTE ESTEVE GARCÍA

Alcalá de Henares, octubre de 2016

Universidad de Alcalá

DEPARTAMENTO DE ECONOMIA Y DIRECCIÓN DE EMPRESAS
Plaza de la Victoria 3
28802 Alcalá de Henares (Madrid)
Tel. 91 885 4293/4238
Fax. 91 885 4294
dpto.ecodiremp@uah.es

Dr. D. Tomás Mancha Navarro, Catedrático de Economía Aplicada del Departamento de Economía y Dirección de Empresas, de la Facultad de Ciencias Económicas, Empresariales y Turismo de la Universidad de Alcalá y Dr. D. Vicente Esteve García, Catedrático de Economía Aplicada de la Universidad de Valencia, tienen a bien

CERTIFICAR

Que: La tesis doctoral con el título "Big data in multiagent systems: market design solutions" elaborada por D. Juan Luis Santos Bartolomé, ha sido dirigida por los abajo firmantes cumpliendo con todos los requisitos exigibles a un trabajo de estas características, por lo que damos nuestra conformidad a la presentación de la misma para depósito y proceder a su lectura y defensa, de acuerdo con la normativa vigente.

Y para que conste donde proceda, firmo el presente documento en Alcalá de Henares, a 14 de octubre de dos mil dieciséis

LOS DIRECTORES DE LA TESIS

UNIVERSIDAD DE ALCALÁ - Departamento de Economía y Dirección de Empresas

Universidad de Alcalá

DEPARTAMENTO DE ECONOMIA Y DIRECCIÓN DE EMPRESAS
Plaza de la Victoria 3
28802 Alcalá de Henares (Madrid)
Tel. 91 885 4293/4238
Fax. 91 885 4294
dpto.ecodiremp@uah.es

Dr. D. José Antonio Gonzalo Angulo, Director del Departamento de Economía y Dirección de Empresas de la Facultad de Ciencias Económicas, Empresariales y Turismo de la Universidad de Alcalá, tiene a bien

CERTIFICAR

Que: La tesis doctoral con el título **"Big Data in Multiagent Systems: Market Design Solutions"**, elaborada por D. Juan Luis Santos Bartolomé, reúne los requisitos exigidos para su defensa y aprobación.

Y para que conste donde proceda, firmo la presente en Alcalá de Henares, a veinticinco de octubre de dos mil dieciséis.

EL DIRECTOR DEL DEPARTAMENTO,

UNIVERSIDAD DE ALCALÁ – Departamento de Economía y Dirección de Empresas –

Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve.

K. Popper

Essentially, all models are wrong, but some are useful.

G. E. P. Box

Everything simple is false. Everything which is complex is unusable.

P. Valéry

If you ever get close to a human and human behaviour be ready, be ready to get confused. […] There's definitely, definitely, definitely no logic to human behaviour.

B. Gudmundsdóttir

Acknowledgements

I am extremely indebted to my supervisor Prof. Tomás Mancha, for seeing me through the ups and downs of this lengthy pursuit, for encouraging me, bringing out the best in me, and most of all, for always having faith in me. I also wish to acknowledge my respect for my co-supervisor Prof. Vicente Esteve, who is a great example of how hard work in research delivers results. I also want to recognize my debt and gratitude to Prof. Federico Pablo, for nudging me in the direction of agent-based modelling, which I know it will become into a life-long passion.

I wish to express my warmest gratitude to all my friends for their unending encouragement and support, especially to Marta, my oldest friend, and Daniel, who proofread all the chapters. He helped me improving both the quality of the text and the contents. Jagoda Kaszowska has been both a close friend and a colleague since February 2008. Despite living in different countries, I am sure she will be my przyjaciółka for the rest of my life. This Ph.D. Thesis would definitely not exist without her constant support.

In Mosips project we faced a complex task and I would like to highlight how María Teresa del Val, Antonio García Tabuenca, María Teresa Gallo, Paolo D'Arminio, Alejandro Valbuena and Nagarajan Viswanathan taught me valuable things, not only in terms of specific tools such as object-oriented programming and database creation, but also I learned from them how to be a good professional and a better person.

I am also deeply indebted to other members of IAES: Juan Ramón Cuadrado, Antonio Mora, Teresa Fernández, Fernando Crecente, Daniel Martínez and especially to Laura Suárez for her continuous support.

I also would like to mention the professors and most of the students I met six years ago at CEMFI. Their behaviour showed me that being an honest person is more important than being successful at any cost.

And finally, but no less importantly, I wish to express my deepest gratitude and affection to my parents for instilling me with a life-long love for learning.

Thanks Gracias Dziękuję நன்றி

Alcalá de Henares (Madrid), October 2016

Resumen

El objetivo principal de esta Tesis es presentar un conjunto de novedosos y diferentes métodos en los que los sistemas multiagente pueden jugar un papel clave en predicciones y modelos económicos en un amplio conjunto de contextos. La hipótesis principal es que los sistemas multiagente permiten la creación de modelos macroeconómicos con microfundamentos reales que son capaces de representar la economía en los diferentes niveles de acuerdo con diferentes propósitos y necesidades.

Los sistemas multiagente solucionan la mayoría de los puntos débiles de los modelos macroeconómicos convencionales, lo que se pone de relieve en el Capítulo 1 de esta Tesis. La crítica de Lucas (Lucas, 1976) demostró que es ingenuo tratar de predecir el efecto de un cambio en la política económica que estudia las relaciones estadísticas detectados en datos históricos. La crítica reciente de Velupillai (2005) afirma que es inútil utilizar modelos ecuacionales convencionales para representar la economía y que se necesita una revolución algorítmica si la macroeconomía pretende realizar predicciones fiables.

Esta última crítica aún es desatendida por la mayoría de los macroeconomistas que se niegan a utilizar las herramientas de la economía artificial. Los sistemas multiagente se encuentran entre la herramienta más sofisticada y adaptable utilizada por físicos e ingenieros y deben sustituir a los modelos tradicionales, si se intentan predecir los efectos de las políticas económicas y asumir una dinámica más realista en los modelos.

Los sistemas multiagente son extremadamente flexibles y pueden incorporar grandes cantidades de datos, por lo que la mejor alternativa para hacer frente a la creciente cantidad de datos disponible. Sin embargo, la simplicidad es la única manera posible construir sistemas útiles y comprensibles.

Un sistema multiagente útil tiene que ser fácil de entender con el fin de que pueda ser mejorado y reutilizado por otros investigadores. De lo contrario, estos sistemas son cajas negras que se convierten en inutilizables, ya que se requiere conocer todas las interrelaciones entre sus módulos y algoritmos. Se pueden incorporar grandes volúmenes de datos en sistemas que incorporan información acerca de los agentes de mediante microdatos (principalmente sobre individuos, familias, establecimientos y empresas) con el fin de crear grandes bases de datos que imitan a los agentes de la sociedad modelada. Una vez que las bases de datos iniciales están preparadas para ser

utilizadas, los sistemas son capaces de incorporar diferentes niveles de racionalidad. Los agentes pueden tener racionalidad perfecta, racionalidad limitada o pueden decidir de forma aleatoria si no tienen suficiente información para establecer un rango de opciones.

Se puede afirmar que los sistemas multiagente son capaces de proporcionar soluciones a diversas necesidades en el diseño de mercados, incorporando grandes cantidades de información y manteniendo de un grado suficiente de simplicidad. El propósito principal de este trabajo es ofrecer un conjunto de nuevas aplicaciones que hacen evidente la afirmación anterior. Por lo tanto, los sistemas multiagente se deben tomar en cuenta con el fin de modelizar la economía y complementar o sustituir los enfoques tradicionales, especialmente en macroeconomía, pero también en otros campos especializados de economía como la política económica, organización industrial y economía del trabajo. La teoría de empresas también puede beneficiarse de la utilización de sistemas multiagente y la mayoría de los capítulos aplicados de estas investigaciones se centran en las empresas y las incorporan en los modelos desarrollados.

La investigación se estructura en seis capítulos. El Capítulo 1 es una introducción teórica al resto de los capítulos que presentan aplicaciones empíricas. En él se compara los sistemas multiagente con dos alternativas: los modelos de equilibrio general computable y la econometría espacial. Ambos son considerados por la mayoría de los autores en macroeconomía y economía regional como las mejores herramientas para encontrar respuestas a las preguntas en estas áreas de investigación. En este capítulo se afirma que los sistemas multiagente, a pesar de sus limitaciones, son preferibles a las alternativas presentadas.

El resto de los capítulos son intencionadamente diferentes en sus objetivos y sus contenidos. Es decir, tratan de mostrar cómo los sistemas multiagente se pueden adaptar a prácticamente todas las preguntas a las que nos enfrentamos en los ámbitos de la Economía y la Administración de Empresas. Las aplicaciones empíricas desarrolladas incorporan diferentes agentes; incluyen regiones y espacios de diferentes maneras; requieren diferentes cantidades de datos y modelizan diferentes tipos de mercados.

En resumen, los capítulos 2 a 6 se seleccionan como representaciones significativas de la versatilidad de los sistemas multiagente. Estas cinco aplicaciones incorporan diferentes tipos de agentes: incluyen individuos (2, 5, 6), familias (2, 5), empresas (3, 5, 6), establecimientos (5), instituciones financieras (6) y usuarios (4).

El capítulo 2 se centra en la población y el mercado de trabajo; el capítulo 3 se encuentra en el ámbito de la Organización Industrial; el capítulo 4 es, probablemente, la primera aplicación de sistemas multiagente en el campo aún poco explorado de la economía de la seguridad; el capítulo 5 analiza el efecto de una política pública; y, finalmente en el capítulo 6 se predice la inflación, siendo la aplicación que está más relacionada con la macroeconomía.

En el ámbito espacial, la desagregación espacial es deliberadamente diferente en cada aplicación: El capítulo 4 no incluye el espacio, El capítulo 6 es una aplicación para la zona euro en su conjunto y en el capítulo 3 se toma España en su conjunto. Los capítulos 2 y 5 exploran las dos de las principales posibilidades para la incorporación del espacio en los sistemas multiagente: el capítulo 2 incluye las regiones NUTS 3 de la Unión Europea y en el capítulo 5 se geolocalizan los agentes.

En el capítulo 2 se desarrolla un sistema multiagente que incluye a todos los individuos de la Unión Europea. El mercado de trabajo es un elemento clave en el sistema pero, con el fin de reducir la complejidad, las empresas están excluidas y la demanda de mano de obra artificial se utiliza en su lugar. De esta manera, se resalta la necesidad de hacer simplificaciones y se muestra la flexibilidad de los sistemas multiagente para ofrecer soluciones a los problemas que se plantean en su desarrollo. Con este sistema podemos predecir la población a escala regional para toda la Unión Europea y cómo distintos niveles de crecimiento económico repercuten asimismo sobre el empleo.

En el capítulo 3 se presenta un modelo de simulación con los principales puntos de vista de la teoría de negocios para estudiar el crecimiento empresarial y la demografía empresarial en un modelo evolutivo estocástico. Se determina el efecto del crecimiento económico sobre la dinámica de las firmas y, posteriormente, el efecto sobre el poder de mercado. El capítulo muestra que desde la aparición de la crisis actual tres sectores industriales han aumentado la concentración empresarial. Estos tres sectores son los que tienen la mayor concentración de los cinco sectores estudiados. La innovación de productos y la innovación de procesos también se incluyen en el modelo y se observa cómo modifican la producción y la demanda. El modelo que se presenta también muestra cómo las empresas se adaptan a los cambios en las características deseadas del producto y el efecto de la crisis sobre estas dinámicas.

El capítulo 4 discute el papel clave de los incentivos en la seguridad de los sistemas de información. Las vulnerabilidades pueden ser reducidas, e incluso eliminadas, si se tienen en cuenta las motivaciones individuales en el proceso de toma de decisiones sobre protección y el aseguramiento. Trabajos anteriores realizan este estudio utilizando un enfoque de teoría de juegos, pero el capítulo muestra que un modelo basado en agentes es capaz de incluir la heterogeneidad y las interrelaciones entre los individuos, y no se centra en el equilibrio alcanzado sino en la dinámica antes de su aparición.

El objetivo del capítulo 5 es el estudio de los efectos de la Ley para la Revitalización Comercial (Ley de Dinamización Comercial) que fue aprobada en la Comunidad de Madrid durante el año 2012. Se permite a las empresas abren las 24 horas del día los 365 días del año y comenzar la actividad empresarial sin licencia. Se desarrolla una base de datos orientada a objetos para estudiar el impacto de esta ley en el comercio de acuerdo con el tamaño y la ubicación de los establecimientos. La base de datos no es sólo espacial sino que el tiempo se incluye en el diseño como pieza principal para determinar los efectos fielmente. Los resultados muestran una extensión de tiempo para algunos establecimientos y áreas, así como una pequeña transferencia de la demanda a establecimientos más grandes que también se producía en menor grado sin la aplicación de esta medida.

Por último, el objetivo del capítulo 6 es explicar los determinantes de la inflación y pronosticar la tasa de inflación en la zona euro en los próximos cinco años. Los comportamientos de los agentes y sus expectativas están relacionados entre sí y explicados por los modelos macroeconómicos aplicados a agentes heterogéneos de tres clases: individuos, empresas e instituciones financieras. Además, se modeliza el comportamiento del sector público y el banco central con un solo agente de cada clase. Se predice una inflación para la zona euro creciente hasta 2018 con un límite cercano al 2,5% en tasa interanual siempre que no se produzcan perturbaciones externas relevantes.

Table of Contents

Table of Contents 1

Index of Figures and Tables 3

Preface 4

Chapter 1. Multiagent systems in Economics. A comparison with Computable General equilibrium and Spatial Econometrics. 11

1.1. Introduction 12

1.2. Shortcomings of Computable General Equilibrium and the shift to multiagent systems 16

1.3. Inadequacies of spatial econometrics analysis for policy forecasting 22

1.4. Multiagent systems features 25

1.5. Multiagent systems for policy forecasting 29

1.6. Artefacts in multiagent systems 34

1.7. Final discussion 37

Chapter 2. multiagent system to forecast employment and migration in the European Union at regional level 42

2.1. Introduction 43

2.2. System overview 47

2.3. Initialization and modules in the system 52

2.4. Calibration of the system 61

2.5. Visualization and scenarios 64

2.6. Final discussion 65

2.7. Appendix: Activity diagrams and algorithms of the model 69

Chapter 3. The effect of firm growth, innovation and firm dynamics on market concentration: an evolutionary simulation model 75

3.1. Introduction 76

3.2. The limitations of game theory for the analysis of firm dynamics and the needing of simulations 77

3.3. Market power in the Spanish manufacturing sector 80

3.4. A simulation model for firm growth and firm dynamics 84

3.5. Market concentration forecast. The effect of economic growth 95

3.6. Conclusions 97

Chapter 4. A multiagent system of insurance and protection decisions......................................99
4.1 Introduction 100
4.2. Role of incentives in information systems...101
4.3. An multiagent system for economics of security 108
4.4. Final discussion 121
Chapter 5. Consequences of trading hours deregulation. A spatio-temporal object-oriented data system for Madrid region...122
5.1. Trading hours in Spain and Madrid deregulation 123
5.2. Databases for trading activities. The design of a spatio-temporal database for Madrid region 126
5.3. Agents behavioural rules 132
5.4. Main results 136
Chapter 6. An Agent-Based Model to Forecast the Inflation rate in the Eurozone 139
6.1. Introduction 140
6.2. Background 141
6.3. Macroeconomic theories to explain inflation...144
6.4. An agent-based Model to forecast the inflation rate 151
6.5. Main results 163
6.6. Conclusions and future research directions 166
7. Final remarks 169
References 173

Index of Figures and Tables

Table P1. Characteristics of the MAS included in the Ph. D. work 6
Figure 1.1. Example of downscaling for population location 32
Table 1.1. Comparison of MAS, CGE and spatial econometrics models 39
Figure 2.1. Overview of the model and the needed data bases 47
Figure 2.2. Class diagram of the agent-based model 50
Table 2.1. Variables obtained from the population census 52
Figure 2.3. Education levels in the model 55
Table 2.2. Labsuit parameterization for ES617 (Málaga) 58
Table 2.3. Comparison of forecast result with and without weighted past regional performance 59
Figure 2.4. Calibration of the coefficients of aging and death for PT171 63
Figure 2.5. User interface of the system 65
Figure 3.1. Evolution of Herfindahl index 81
Table 3.1. Effect of the crisis on firms by size 82
Figure 3.2. Evolution of the average number of employees 83
Figure 3.3. Size distribution in established and new firms 92
Figure 3.4. Herfindahl index under different scenarios 94
Figure 3.5. Prices higher than the lowest in the sector (simulation for 2000-2020) 96
Figure 4.1. Effect of protection cost on protection rate 115
Figure 4.2. Effect of the amount loss on agents' behaviour 116
Figure 4.3. Effect of protection cost on protection rate 117
Figure 4.4. Effect of the implemented strategy on protection rate 118
Table 4.1. Evolution of the behaviour in strategies 2 and 3 with different L 119
Figure 4.5. Effect of the portion of desirable targets on protection rate (t=10) 120
Figure 5.1. Local units of sector 47 in Madrid region 128
Table 5.1. Number of retail establishments by subsectors 130
Figure 5.2. Database design 131
Figure 5.3. Main shops in which an individual purchases 135
Figure 5.4. Effect of the deregulation on the average number of employees 137
Figure 5.5. Effect of the deregulation on the extra opening hours per week (%) 137
Figure 5.6. Effect of the deregulation on the establishments that open on Sundays (%) 138
Figure 6.1. Agents and variables in the model 154
Figure 6.2. Agents and transactions in the model 158
Figure 6.3. Calibration and simulation of weights of macroeconomic theories for inflation expectations 161
Table 6.1. Values of theta coefficients in different stages of the calibration process 162
Figure 6.4. Inflation forecast of the agent-based model from 2008 to 2021 164
Figure 6.5. Inflation forecasts of the agent-based model from 2016 to 2021 165

Preface

The main aim of this Ph.D. Thesis is to present a set of different novel ways in which multiagent systems (MAS) can play a key role in economic forecasting and modelling in a wide panorama of contexts. Then, the principal hypothesis is that multiagent systems allow creating macroeconomic models with real microfoundations that are capable of representing the economy at different levels according to different purposes and necessities.

Multiagent systems overcome most of the weak points of conventional macroeconomic models as it is highlighted in Chapter 1 of this Thesis. The Lucas critique (Lucas, 1976) proved it is naive trying to predict the effect of a change in economic policy studying detected statistical relationships in historical data. Recent Velupillai critique (Velupillai, 2005) states it is worthless to use conventional equational models to represent the economy and an algorithmic revolution is needed if macroeconomics wants to make reliable predictions.

This last critique is still neglected by most of macroeconomists and they refuse to use the tools of Artificial Economics. Multiagent systems are amongst the most sophisticated and adaptable tool used by Physicists and Engineers and they should replace traditional models if Macroeconomics aims to correctly forecast the effects of economic policies and incorporate more realistic dynamics in their models.

Multiagent systems are extremely flexible and can incorporate large amounts of data, making them the best alternative to cope with the increasing quantity of available data. Nevertheless, as I learned –taking part of the Spanish team of IAES- during the three-and-a-half years MOSIPS project[1] (dozens of variables, hundreds of parameters and thousands of lines of code) simplicity is the only possible way to construct meaningful multiagent systems.

[1] MOSIPS (MOdelling and SImulation of public Policies on SMEs) was a project founded by the European Commission under the 7th Framwork Programme and developed by a consortium of organizations of United Kingdom, France, Germany, Austria, Italy and Spain. Its main objective was the building of a user-friendly policy simulation system allowing forecasting and visualization of the socio-economic potential impact of public policies on small and medium enterprises.

A meaningful multiagent system has to be easily understandable in order to be discussed, improved and reused by other researchers. Otherwise, these systems are black boxes that become unusable as they require knowing all the interrelations among their nonlinear modules and algorithms. Big Data can be conveniently used in multiagent systems incorporating microdata information about agents (mainly individuals, families, establishments and companies) in order to create large databases that mimic the agents of the modelled society. Once the initial databases are ready to be used, the systems are able to incorporate different levels of rationality. Agents can have perfect rationality, bounded rationality or they can decide randomly if they do not have enough information to establish a rank of choices.

It can be claimed multiagent systems are able to provide solutions to extremely diverse necessities in market design, incorporating large amounts of information and maintaining a sufficient degree of simplicity. The main purpose of this Ph. D. work is to offer a set of novel applications that sufficiently make evident the truth of the previous statement. Therefore, multiagent systems should be increasingly taken into account in order to model the economy and complement or substitute traditional approaches, especially in Macroeconomics but also in other specialized fields of Economics such as Economic Policy, Industrial Organization and Labour Economics. Business Theory can also benefit to a great extend from the use of multiagent systems and most of the applied chapters of these research are focused on incorporate companies into the developed models.

The research is structured in six chapters. Chapter 1[2] is a theoretical introduction to the rest of the chapters that present empirical applications. It compares multiagent systems with two alternatives: Computable General Equilibrium models and Spatial Econometrics. Both are considered by most of authors in Macroeconomics and Regional Economics as the best tools to find answers to the questions in these research areas. In

[2] The first version of Chapter 1 is published in Spanish in *Encrucijadas, Revista Crítica de Ciencias Sociales*. Nº 4, pp. 133-150. An English version was presented in the *European Regional Science Association Congress* in Palermo (Italy) and an updated version is under revision in *Argumenta Oeconomica*.

this chapter it is claimed multiagent systems, despite their limitations, are desirable to the presented alternatives.

The rest of the chapters are intentionally as different as possible in their aims and their contents. That is, they try to show how multiagent systems can be adapted to virtually all the questions we face in Economics and Business Administration. The developed empirical applications incorporate different agents; include regions and space in different ways; they require different amounts of data and they model different kinds of markets. Time is taken in different definition in these models; in one of them the system can incorporate decisions taken in less than 24 hours and in other case the model has a simulation period of months or years.

Table P1. Characteristics of the MAS included in the Ph. D. work

Chapter	Agents	Primary goal of estimation & forecasting	Other features	Spatial scope	Spatial disaggregation
2	Individuals Families	Population dynamics Labour variables	Artificial labour demand	European Union	NUTS-3
3	Companies	Firm dynamics Market concentration	Innovation Price and non-price competition	Spain	Country level
4	Users	Insurance vs. Protection decision	Imitation Free-riding	No	No
5	Individuals Families Companies Establishments	Effect of public policy on economic activity	Consumer preferences Transport networks	Madrid region	Geolocation
6	Individuals Companies Financial institutions	Inflation	Oil shock Public sector	Eurozone	Monetary Union

Source: Own elaboration

To sum up, the chapters from 2 to 6 are selected as meaningful representations of the versatility of multiagent systems. The number of examples could have been different, nevertheless it is considered that five is a good choice because it permits to show how this tool offers modelling solutions to a high number of different questions. Every chapter has already been published in journals or international peer-reviewed books and has been improved and adapted for being included in this Ph. D. work.

These five applications of multiagent systems incorporate different kinds of agents: they include individuals (2, 5, 6), families (2, 5), companies (3, 5, 6), establishments (5), financial institutions (6) and users[3] (4).

These models have several scopes as shown in Table P1. They are as different as possible in order to represent how multiagent systems can deal with real questions in different domains. Thus, Chapter 2 is focused on population and labour market; Chapter 3 is in the field of Industrial Organization; Chapter 4 is, probably, the first application of multiagent systems in the still little explored field of Economics of Security; Chapter 5 analyses the effect of a public policy; and, Chapter 6 forecasts inflation and it's the application that is more related to Macroeconomics. Each of them incorporates additional features that make them more attention-grabbing and contribute to the quality of the achieved results.

The spatial scope and spatial disaggregation are deliberately different in each application: Chapter 4 does not include space, Chapter 6 is an application for the Eurozone as a whole and Chapter 3 studies companies in Spain. Chapters 2 and 5 explore the two main possibilities for incorporating space in multiagent systems: Chapter 2 includes NUTS-3 regions of the European Union and Chapter 5 models the space geo-locating the agents. It has.

Finally, the summarized content of the five applied chapters is briefly presented:

In Chapter 2[4] a multiagent system that includes all the individuals of the European Union is presented. Labour market is a key element in the system but in order to reduce complexity firms are excluded and an artificial labour demand it is used instead. In this way, the necessity of making simplifications is highlighted and the flexibility of

[3] Users can be defined as individuals and organizations because the model has a high degree of flexibility.
[4] Chapter 2 was presented in the *European Regional Science Association Congress* in Palermo (Italy) and in the *European Conference of Complex Systems* – Integrated Utility Services: Smart Systems in Complex Models in Barcelona (Spain). The first version of this chapter is published in *Emergence, Complexity and Organization*, 17(2) C1.

multiagent systems to offer solutions is shown. However, this system has five modules that could be autonomous models and they incorporate a high amount of information about individuals.

The aims of this second chapter are to present and to discuss an agent-based model of population dynamics for the European regions at NUTS-3 level. It includes individuals that perform several activities with bounded rationality. The chapter briefly discusses the latest novelties on this topic and then describes the processes to prepare a database with the necessary information to feed and calibrate the model. Then the initialization module is presented. It generates individual heterogeneity according to average and marginal aggregate distributions of the included variables that characterize the agents. In order to simulate the mechanisms of migration the model creates an artificial labour demand at regional level using simple but effective rules based on mainstream economic theory. The rest of the model is also presented: education, pairing, aging and deaths. A set of scenarios is defined and the regional aggregates are computed. Results are visualized with tables, graphics and maps.

Chapter 3[5] presents a simulation model with the main insights from business theory to study firm growth and firm dynamics in a stochastic evolutionary model. Firm growth behavior and firm dynamics are defined according to the results of a panel data for a set of manufacturing markets for the Spanish economy. Then the effect of the economic growth on firm dynamics and subsequently the effect on market power are determined. The chapter shows that since the emergence of the current crisis three industrial sectors have increased business concentration. These three sectors were the ones with the highest concentration out of the five sectors studied. Product and process innovation are also included in the model and how they modify production and demand. The model presented also shows how firms adapt to changes in desired product characteristics and the effect of crisis on these dynamics.

[5] The first version of this chapter is published in *International Journal of Applied Behavioral Economics*, 5(3) pp. 31-49.

Chapter 4[6] discusses the key role of incentives in information systems security. Vulnerabilities can be reduced, and even removed, if individual motivations are taken into account in the process of protection and insurance design. The chapter first discusses the importance of externalities, free-riding behavior, uncertainty and the incentives mismatch between individuals and organizations involved in information systems security. Previous works perform this study using a game theoretical approach but the chapter shows that an agent-based model is capable of including the heterogeneity and interrelations among individuals, not focusing on the reached equilibrium but on the dynamics prior to its emergence.

The aim of chapter 5[7] is to study the effects of the Commercial Revitalization Act (Ley de Dinamización Comercial) that was passed in Madrid region during 2012. It allows businesses to open 24 hours a day 365 days a year and start business activity without a license. An object-oriented data model is presented to study the impact of this law on the trade according to size and location of the establishment. The database is not only spatial but time is included in the design as a main piece to determine the effects faithfully. The results show a limited extension of time for most establishments and areas as well as a small demand transfer small to larger ones that also occur in a lower degree without the application of this measure.

Finally, the aim of Chapter 6[8] is to explain the determinants of inflation and to forecast the inflation rate in the Eurozone for the next five years. The behaviours of agents and their expectations are interrelated and explained by macroeconomic models applied to heterogeneous agents of three classes: individuals, companies and financial institutions. In addition, the behaviour of public sector and central bank is also modelled with a single agent of each kind. Once the quantitative easing policy is implemented, the quantitative theory of money expects higher inflation rates in the long run. Inflation should remain low taking into account the Phillips-Curve. Last, according to the

[6] The first version of chapter 4 is published in *International Journal of Agent Technologies* and Systems, 7(3) pp. 1-17.

[7] The preliminary version in Spanish is published in *Documentos de Trabajo del Instituto Universitario de Análisis Económico y Social (IAES)*. Nº 7/2014. It was presented in the *II School on Complex Networks* in Lucca (Italy) and the first version of chapter 5 is published in *Journal of Socioeconomic Engineering*, 1, pp. 4-11.

[8] Chapter 6 is accepted for publication in *Applied Behavioral Economics Research and Trends*.

Aggregated Supply and Demand as well as to the Money Market equilibrium, the behaviours modelled allow forecasting low inflation. However, an external shock, as it would be an increase in the price of important commodities, can alter the inflation rate to a great extent.

CHAPTER 1

Multiagent systems in Economics.
A comparison with Computable General equilibrium and Spatial Econometrics

1.1. Introduction

There is a blurred consensus in the distinction between mainstream and periphery approaches. In fact, many assumptions have traditionally been shared by both of them. Notwithstanding, after the recent financial and economic crises, a great number of researchers have started to question the validity of core economics assumptions –clearly based on neoclassical paradigm- and conclusions extracted from the analysis carried out using the conventional models. Recently, it has been emphasized by authors in many research areas such as macroeconomics (Caballero, 2010), environmental economics (Janssen, 2005) as well as the regional economics (Schenk, Löffler and Rauh, 2007). Moreover, findings of the one particular area of research can be easily extrapolated to the other areas making the whole investigation process especially interesting. It can be noticed in the modelling process in economics aimed to forecast the effect of public policies and to make predictions under different scenarios.

Forecasting models in economics depend on the chosen paradigm in macroeconomics and the set of assumptions related to it. The failure of conventional macroeconomic models to predict the last financial crisis has made the whole field of study particularly prone to further research. A direct consequence of paradigm shift to so-called periphery approaches is the need to discuss changes in economics modelling.

Recently, the pitfalls of core macroeconomics have been debated by researchers and experts in many fields of studies but still further research is needed, especially in those that focus on spatial issues. In July 2010, The Economic Committee of the House of Representatives of United States discussed about Dynamic Stochastic General Equilibrium (DSGE) models[9], the approach that constitutes the core macroeconomics nowadays. During this meeting the researchers criticized the use of a single methodology in the economic policy (Solow, 2010). The main problem of the core economics methodology is its simplistic assumptions such as the efficiency of financial markets and rational expectations (Kirman, 2010) that do not capture the complexity of

[9] A special type of Computable General Equilibrium models that has been extensively used in Economics during the last decade by international institutions, central banks and finance ministries in most of the OECD countries.

socio-economic processes. Others well-known economist has even claimed that “for more than three decades, macroeconomics has gone backwards” (Romer, 2016).

However, these critical positions have received responses. One of the leading voices who argue that general equilibrium models will be able to predict major crises in the future is V.V. Chari (2010) that argues: “By construction, a model is an abstraction which incorporates features of the real world thought important to answer the policy question at hand and leaves out details unlikely to affect the answer much. Abstracting from irrelevant detail is essential given computational scarce resources, not to mention the limits of the human mind in absorbing detail! Criticizing the model just because it leaves out some detail is not just silly, it is a sure indicator of a critic who has never actually written down a model”.

Of course any model is an abstraction of reality, but CGE models have shown their predictive weakness with the financial crisis of 2008. They may not be the best option and therefore it can be more productive their replacement than an attempt to broaden their scope. Nowadays it is no longer possible to argue the choice of one alternative due to the scarcity of computational resources. The development of computers in recent years enables us to build complex models even in everyday laptops. Moreover, it falls into the common error of considering that the results provided by the model correspond to all the contents of the model. This occurs in models with just a few dozens of equations and variables, but in more complex models the user only receives aggregated information, this is the effect of the policy and the potential side-effects.

This kind of criticism is one of the usual response to MAS as stated in the article by Waldherr and Wijermans (2013) which identifies the most frequent criticisms received by social scientists presenting papers based on MAS: The models are judged to be too complex, too simple, not based on theories, unrealistic, arbitrary, use dubious assumptions and be black boxes. Many reviewers are not used to non-equational models in the field of economics while in other disciplines such as sociology are rejected for not being empirical. Anyway, the increment of MAS in social sciences in the last years shows a lower rejection of its use.

In turn, CGE provide the optimal solution mix of endogenous variables in response to an exogenous shock. CGE models contain explicit supply constraints, usually embedded in a neoclassical framework. The theoretical basis of CGE is complete Arrow-Debreu equilibrium under perfect competition. The specification restricts the number of parameters in a way allowing for a model calibration relying on a data base (Bröcker, 1998).

CGE approach has microfoundations and assumes economy being in the state of equilibrium and only exogenous shocks to the system exist. The effects of such external shocks can be temporal and permanent. In the first case the initial equilibrium is preserved while in the second case, after a short period of adjustment, the economy achieves a new state of equilibrium. In both cases, the adjustment mechanism of the adjustment process remains unexplained and as it is proved in this chapter it seems to be one of the main, though not the only, pitfalls of this approach.

The alternative to the CGE is multiagent systems. They neither assume the economy being in a state of equilibrium nor do they explain or concentrate on adjustment process. Instead, those models allow including the heterogeneity of their components what can be useful for instance in regional economics modelling because they put the emphasis on the spatial location and its importance on the overall economic performance.

MAS approach is completely different from the CGE as it takes into account the dynamics of the system as a central point of the study. Then, it makes it possible to find the intermediate determinants of the observed effects. The behaviour of the agents can be determined and fundamentally altered by direct interaction between them while the interactions in the conventional models happen only indirectly through pricing. This characteristic of the multiagent systems allows for changing behaviour through imitation in addition to learning and the results extracted from the model are not constraint by the definition of rationality degree.

Spatial econometrics is an alternative to CGE models when spatial accuracy is needed. In this context, the question of adjusting the shape of the modelling becomes even more complicated as not only computable general equilibrium and multiagent systems shall be compared but also a great number of models that use spatial econometrics techniques should be taken into consideration (Fingleton, 2000). Although econometric models theory and its related assumptions are widely known, the main features of such models used in economics are explained in this paper leading to the concept of spatial econometric models.

The most important conclusion extracted from the analysis is that CGE models and spatial econometric models do not seem to be such as useful tool to evaluate the regional impact of public policies and simulate their incidence. The drawbacks can only be overcome making use of multiagent systems. MAS allows including the institutional structure such as financial organizations and governments in the natural way while this feature is extremely difficult to be implemented in conventional models.

Moreover, assuming MAS approach we can learn in depth about relations between different agents, emphasizing the importance of heterogeneity, networks, location and learning by imitation. These models help us to understand not only the expected trend of macroeconomic variables but its impact on the various actors according to their characteristics. In this way, they are more useful than those widely used in spatial econometrics to make forecast. Moreover, they resist the Lucas critique and the more recent critiques made by Velupillai pointed out in the previous section.

The objective of this section is to compare multiagent systems (MAS), computable general equilibrium (CGE) approach as well as spatial econometrics, in the context of the assessment of the regional impact of policies. The section is structured as follows. After this introduction, in the part 1.2, the main shortcomings of CGE models are presented and a need of paradigm shift to the one assumed in agent-based modelling is emphasized. Then, in the part 1.3, inadequacies of econometrics analysis for policy forecasting are discussed. In the part 1.4, the multiagent systems features are explained including the area of research of agent-based computational economics, classification of agents as well as their number and the consistency of such models. In the part 1.5, the appliance of multiagent systems in spatial analysis is presented with emphasis on statistical matching, downscaling and location importance. The artefacts in multiagent systems are shortly described in the following part (1.6). Further research areas and conclusions are presented in the last part of the chapter.

1.2. Shortcomings of Computable General Equilibrium and the shift to multiagent systems

Until the mid-twentieth century macroeconomics studies were conducted without a complete formal theory to back up their assumptions. Macroeconomic policy was based on empirical knowledge without interconnections and without many of the theories that nowadays are considered basic. The field progressed through trial and error, and economists began using simple models, such as quantitative theory and Keynesian cross, which capture the essential elements of macroeconomics and provide qualitative guidance for policy decisions. Economic theory had not evolved to its mathematical axiomatic phase at that time, so the idea of a macroeconomic model based on axioms about individual behaviour had not been raised.

The development of macroeconomic models using econometrics permits many Keynesian models were presented based on microeconomic theory. They offered a formal model, the IS-LM (Investment-Saving / Liquidity preference-Money supply) conceived in 1936 during the Econometric Conference and presented later by Hicks (1937) who later recognized the inadequacies of general equilibrium models for assessing the impact of policy decisions (Hicks, 1980).

The microeconomic foundations were integrated into sector models to give them the appearance of being based on the axiomatic approach from the general equilibrium theory. This led to Keynesian economics to separate from the stance of its founder in several aspects. For example, Keynes stated agents are related in several different ways and he thought aggregate values are the sum of individual decisions, due to different motivations (Mancha and Villena, 1993).

These Keynesian models continued its progression until the end of the sixties. The criticism started to be widespread since the mid-fifties with Friedman (1957) and its permanent income hypothesis, which attacked the consumption function of these models, a key point to determine the multipliers in which fiscal policy is supported.

Monetarist economists continued to criticize the principles of Keynesian models arguing that economic instability was caused by an inadequate monetary policy and not by the *animal spirits* of investors (Friedman and Schwartz, 1963) serving as a precedent-based monetary policy control inflation. The Phillips curve in Keynesian models is used to explain how prices adjust over time with a trade-off between inflation and unemployment. Phelps (1968) opposed this assumption and explained this empirical unemployment from money illusion. He and Friedman stated expansionary economic policies cannot decrease the unemployment rate below the natural unemployment rate because workers would update their inflation expectation.

Robert Lucas in his 1976 seminal article argued that Keynesian models that were used up to that moment were not valid to predict the effects of policies as they were based on the optimal decision rules of economic agents, and these rules shift with changes in their environment. Hence, any policy change alters the structure of econometric models. According to Lucas, it is naive trying to predict the effects of a change in economic policy on the basis of relationships observed in historical data, especially using the aggregates. This argument puts into question the use of econometric models to explain the dynamics of economic processes.

The Lucas critique suggests that to predict the effect of a public policy, it is necessary to make use of models with parameters that take into account the relationship of preferences, technology and resource constraints that create individual behaviours. These models based on a new approach make feasible to predict the actions of individuals, taking into account the change in the public policy. Then they add individual decisions to estimate the overall effects of these variations.

The cycle theories of Kydland and Prescott (1982) and Long and Plosser (1983) differed from previous works: They did not highlight the role of monetary policy to explain economic cycles. These authors stressed the role of random disturbances in technology and intertemporal substitution of consumption and leisure that these shocks induce on the individuals.

These contributions played an important role to change the previous dominant view in macroeconomics to the current one. Then, the mathematical rigor and the conceptual difference with microeconomics theory became less significant. Business cycle models are the preeminent example of the dynamic general equilibrium theory.

This theory tries to give a comprehensive explanation of the behaviour of the production, consumption and price formation in an economy with one or several markets for goods, services and production factors. General equilibrium theory tries to explain from the particular to the general, starting with individuals while the Keynesian macroeconomics offered a vision from the general to the particular, where the analysis started with markets behaviour.

In the free market economies, prices and production of all goods, including the price of money, are interrelated. A change in the price of a good can affect other prices, such as wages of the producers of that good and the price of substitute and complementary goods. Hence, to calculate the equilibrium price of a single product, an analysis is required taking into account the rest of the different components that are available for production and exchange.

The first attempt to model the price formation of an entire economy was made by Léon Walras in 1874. His research was continued by the neoclassical economists over the second half of the twentieth century. In his work he provided several models, each of which takes into account an increasing number of aspects of a real economy: different types of goods, production, economic growth and money.

Arrow and Debreu (1954) continued the investigation of the conditions under which the equilibriums are unique and stable. Thereby, any general equilibrium model accepts the assumptions that lead to the existence and uniqueness of this equilibrium. Some of these cases focus on individual preferences and production functions of companies, they have to meet certain mathematical conditions implausible in the real world. With a large number of consumers these conditions can be relaxed to require only positive endowments of goods to ensure this balance. A number of authors (Little, 1982; Bowles & Gintis, 1993) have criticized the lack of credibility in enforcing these conditions.

However, models based on the general equilibrium theory have demonstrated their unquestionable utility in recent years.

Simultaneously to the neoclassical revolution, Keynesian economists reacted and started to use general equilibrium models to understand the allocation of resources and the emergence of crises (Barro and Grossman, 1971). Afterwards the Keynesian continued to discuss the hypothesis of perfect rationality and pricing (Mankiw, 1985, Akerlof and Yellen, 1985). Keynesian contributions were included in the general equilibrium theory, correcting some of the most important problems and achievements of the Neoclassical Synthesis. Nevertheless, this position did not satisfy everyone and was criticized by Keynesians such as Joan Robinson, Nicholas Kaldor and Michal Kalecki.

In the last decade of the last century, a new synthesis was established, especially applied on monetary policy (McCallum and Nelson, 1999, Clarida, Gali and Gertler, 1999). This synthesis is constructed taking as starting point the from neoclassical models and using tools of general equilibrium theory, materialized in DSGE (Dynamic Stochastic General Equilibrium) models, a class within CGE models (Marimom and Scott, 1999; Canova, 2011, DeJong and Dave, 2011). The preferences of agents, the restrictions they face and the intertemporal optimization mechanism are the starting point and the analysis is constructed from these microeconomic foundations. One feature of these models is the rigidity of prices in the short term, one of the assumptions that most vehemently rejected neoclassical in their first attacks on Keynesianism.

These models have been imposed as the paradigm within the macroeconomic analysis of short and medium term. Central banks and finance ministries have shown increasing interest in its use for policy analysis. Today many institutions in both developed and emerging economies have developed their own models and many others are starting or planning to. DSGE models have been directly applied to policy evaluation and simulation although their validity was not sufficiently demonstrated from historical series.

DSGE models are powerful tools that provide a coherent framework for analysis and policy evaluation. They help to identify the sources of fluctuations, answer questions about structural changes and predict the effects of policy changes, enabling counterfactual experiments.

An important feature of general equilibrium models is what Walras described as a condition of market equilibrium (*market clearing condition*), which was given by a central authority which proposes a set of prices, excess demand determine these prices, and adjusts to their equilibrium values. However, now we know that the dynamics implicit in an exchange economy is generally unstable, even chaotic (Fisher, 1983; Saari, 1985). Furthermore, this mechanism has not a counterpart in the real market economy, and violates the spirit of decentralized exchange, in which the general equilibrium model is based.

A more reasonable approach would recognize that the behaviour of the aggregate does not have to correspond to the behaviour of the components added, and cannot usually be derived from consideration of the latter only. Any economic model theoretically well founded shall analyze not only the characteristics of individuals, but also the structure of their interactions. This view is widespread in other disciplines such as biology, physics and sociology. These sciences recognize that the aggregate behaviour of systems of particles or molecules, neurons or the society cannot be deduced from the characteristics of a representative of the population. The same can be taken into account for economic systems, where the fallacy of composition[10] exists, and should be treated (Howitt, 2006).

The macroeconomic variables arise from the aggregation of the activities of the agents that perform economic actions. To get the best possible results it would be useful to understand these actions individually. Hence, in addition to aggregated data at national and regional it would be possible to obtain aggregates at lower level.

One of the most important functions of macroeconomics is the evaluation of public policies and the forecast of the macroeconomic outcome. Even if the overall result obtained with CGE models is accurate, it would not be enough because the different

[10] The fallacy of composition takes place when the result of the whole is inferred when only a part is known. The example of the so-called Condorcet paradox taken from the voting theory could clarify the issue: rational voters could apply the majority rule and arrive to a non rational decision due to a lack of transitivity.

effects would not be observed on the heterogeneous agents. Hence, a model that adds the agents into an aggregate cannot report the characteristics of the agents who are affected by a policy. Thus, it is unachievable to report which ones change their labour status as a result of a labour policy variation. This should be the ultimate goal of macroeconomic analysis in addition to knowing the total effect on the whole economy.

There are determinant factors in the real world as subsistence needs, incomplete markets, imperfect competition and strategic behaviour interactions, which cause difficulties in the analytical formulations and therefore are not incorporated. Several researchers in the middle-80s were aware of the limitations of general equilibrium models, and they started to develop agent-based tools in the area of computational economics. These tools were able to capture part of the complexity of real-world economic phenomena. Before long they began debating whether the application of these instruments could provide more evidence based economic modelling. Three decades later, thanks to the great increase in computing capacity and the proliferation of micro-data sources, it seems that it is time to overcome the drawbacks of CGE models and study a new methodology that is far from perfect, but that will continue to advance the scope of macroeconomics, especially the studies devoted to assess the spatial distribution of the economic activity.

Furthermore, as stated by Diebold (1998)[11] the match goal of DSGE modes was to represent several variables at equilibrium rather than represent a complex economy. However, these models started to get more complex and nowadays it is nearly impossible to understand what actually happens inside them. The result is the situation described by Robert Solow (2006), when he wrote that in human nature there is a deep pleasure to adopt and defend a doctrine totally counterintuitive, that leaves the ones uninitiated wondering what planet they are on.

The previous stated reasons can justify the need for a paradigm shift in macroeconomic forecasting. As noted, multiagent systems could be one alternative, but not only because there are other models out of the mainstream that can also predict major crises

[11] In the session "Monetary and Fiscal Policy: The Role for Structural Macroeconomic Models" at the annual meeting of the American Economic Association

(Caballero, 2010). In fact, recent studies have questioned the validity of any economic method based on equations instead of algorithms. K.V. Velupillai (2005, 2012a, 2012b) argues mathematically that any model based on equations or what he calls *classic* mathematics can be consistent with itself and correct from a formal point of view, but the results cannot be extrapolated to the reality. Only algorithmic models can be used to compare their results with reality. Everybody knows macroeconomics does not fulfil its purpose and we would need to start almost from scratch in order to say something relevant. The only possibility is to use algorithms, operations and rules that allow us to find the solution of a problem. These algorithms must be implemented on decision-makers units that perform actions: agents.

1.3. Inadequacies of spatial econometrics analysis for policy forecasting

The aim of this section is to clarify the concept of spatial econometrics and to compare the arguments in favour and the ones contrary to the use of spatial econometrics in regional economic studies, especially if the aim of the research involves public policies.

The term spatial econometrics appeared for the first time in the late 70s during the annual meeting of the Dutch Statistical Association (Paelinck and Klaassen, 1979). In the opinion of the authors, the *spatial econometrics* is “blend of economic theory, mathematical formalization and mathematical statistics” and its main features are related to “the asymmetry of spatial relations, the role of spatial interdependence in spatial models, the importance of explanatory factors located in other spaces, the differentiation between ex-post and ex-ante interaction and the explicit modeling of space”.

The contributions of spatial econometrics to regional economics are remarkable. Hence, it is not surprising that many international experts have dedicated their work to spatial econometrics modelling (Isard, 1960; Maddala, 2001). Notwithstanding, still a few areas of research are left unexplored. A good example of the problem in spatial econometrics is the one of spatial dependence among contiguous residuals of a linear regression.

Maddala analyzed this problem that falls in the category of *spatial correlation*. In practice, there are two ways the error terms might be correlated. One of them is related to omitted variables and the so-called "*keeping up with the Jones*" effect[12]. The second way is related to the problem of error terms for contiguous states that usually tend to be correlated. Hence, spatial correlation models are aware of this important issue, usually ignored in the analysis undertaken with other tools in applied economics. (LeSage, 2008)

The appearance of New Economic Geography, became spatial econometrics tools as a part of mainstream economic analysis. It contributed to broaden the scope of this approach and an increasingly number of techniques were incorporated to the popular the measurement of spatial variables.

Frequently, spatial econometric models are used to represent a narrow and specific part of the economic reality. Hence, they do not worry about the remaining processes not taken into account. An example can be found in García-Milá and Montalvo (2007). The authors try to evaluate the impact of new highways on business location. The creation of new highways constitutes a public policy and in addition to the increase in infrastructures availability, business owners might vary their behaviour due to a change in their expectations. Traditionally, this kind of infrastructures is developed to spur growth.

A significant increase in local economic activity may result on a lack of additional public investments. On the other hand, a smaller increment in the number of businesses located near new highways would change the expectations of rational businessmen, who would anticipate new investments. In any case, the expected result would be a slightly non-significant increment in the number of businesses, but the analysis undertaken with spatial econometrics tools cannot conclude the ultimate causes of the observed result.

Besides of its scarcity, several examples can be found that aim to depict all the economy using spatial econometrics in addition to other tools, commonly input-output matrixes to assess the sector relations and interactions. Regional IOE (Input-Output Econometrics)

[12] This effect describes the correlation of variables (e.g. for households) in the same neighbourhood.

"models can and generally do have a wide range of specifications, depending on the region type and purpose of design impact analysis, forecasting, demo-economic, income distributions fiscal policy evaluations, etc." (West, 1995). They do not consider endogenous government revenue and expenditure, and they usually introduce impacts on the demand, as it is done in standard IO models. This kind of approaches represents an attempt to overcome Lucas critique and achieve attention-grabbing results, prevailing over CGE models in the study of small areas.

Taking into account the previous comments, it can be assumed spatial econometrics cannot constitute the new paradigm in macroeconomic forecasting. However, new approaches to the concept of region broaden the scope of this approach and make its results more reliable and accurate. A key reference to this topic is the book of Fernandez and Rubiera (2013) where the potential of spatial economics in the socioeconomic analysis of territory is explored.

Traditionally databases have been available only at the regional level but nowadays the number of geo-located microdata and data sources at municipal level is growing. This novelty makes possible to complement the results obtained using traditional regional economics tools. The study of the concept of region appears to be a fundamental issue to analyse prior to the specification of the model.

Moreover, the restriction imposed by the level of disaggregation of data available is becoming less frequent in the analysis. Hence, not only shall the adequate methodology be chosen in each case of study, but also the most convenient spatial unit for the purpose of analysis. The first step of the research carrying out is to discuss the importance of election of spatial unit and the redefinition of the concept of region that is to be made by aggregation of local units that cease to have administrative character but instead gain additional economic content. Beside of the fact that there are an increasing number of publications of spatial time series at lower (micro) level, there are occasions when no such data is available. In such case, particular statistics and econometric tools shall be used to estimate disaggregate data.

In the first part of Fernandez and Rubiera (2013), F. Sforzi redefines the concept of region for regional and geographic economies, analysing it also from the political point of view. The administrative regions have traditionally been used to carry out most of the studies in the regional economies. Notwithstanding, even in the early traditional studies, new concepts, such as Functional Economic Areas or Local Labour Markets, have already appeared. These changes have been used to improve a lot the results extracted from the analysis and their usage is becoming more and more widespread in the spatial economics.

This new approach, when combined with the appropriate analysis carried out with spatial econometrics can lead to powerful models. These models are starting to estimate effects and determinants at the lowest possible scale, and its usefulness in the regional analysis will increase without any doubt. However, they would not integrate all the actual heterogeneity of individuals and the complexity of their underlying behaviour.

1.4. Multiagent systems features

The social sciences are aimed to understand the behaviour of individuals but also to explain how their interactions lead to the observed aggregate outcomes. Understanding the economic system requires more information and knowledge that understanding only the behaviour of individuals who compose the system. The overall effect of modelling interactions shows that the results can be more than the simple sum of its parts.

The MAS are adequate for studying macroeconomic and regional economics as they work on methods that assume the complex interactions between agents, they describe emergence properties of simpler interactions between them that cannot be deduced only from the aggregation of properties of all individuals.

When the interaction of agents depends on the past experience, and especially when they adapt permanently to new scenarios on the basis of previous experience, expectations and imitation, the mathematical analysis and its tools have very limited ability to derive dynamic consequences. In such case, MAS may be the only one existing practical method of analysis.

As explained in Epstein and Axtell (1996) and elaborated in Tesfatsion (2007), the Agent-Based Computational Economics (ABCE) is the calculation of economic processes modelled as dynamic systems of interacting agents.[13]

The ABCE is a relatively unexplored approach that tends to study economic system holistically. Once the behavioural rules are specified by the researcher, all subsequent events are driven by the interactions between agents. These interactions are determined dynamically by the internal structures of behaviour, the level of information, beliefs, motivations and methods of data processing. A crucial point of multiagent systems is that there are not aprioristic theoretical restrictions that could limit interactions between agents and that usually are subjective. The agents in the MAS world shall tend to be completely free to act and interact with their virtual realities as it is done in the real world by its counterparts.

In order to create a multiagent system to facilitate understanding of economic processes, three criteria have to be followed:

- The MAS model should include the correct classification of agents that has the empirical reality they have to replicate.
- The scale should be adequate to the objectives that it is tried to achieve.
- The specification of the model shall be subjected to the empirical validation in order to ensure final and intermediate causal mechanisms are being modelled adequately.

The classification of phenomena into groups or categories is mandatory in a scientific research process correctly carried out. It also facilitates the detection and correction of errors while it allows the comparison of results. The classification should be performed paying attention to the empirical findings highlighting its importance in the process of theorizing and modelling. Before proceeding to the classification, the researcher should know the answer to the following issues:

[13] In this chapter, the term 'agent' refers to a collection of individualized data and the methods of representation of an entity defined in the virtual system. These agents may be living beings, social groups, institutions or companies.

- Types of human needs which are relevant for understanding the economic phenomena of interest.
- Types of goods and services that satisfy these needs.
- Types of facilities that exist or may exist to produce and distribute these goods and services, and that participate in production activities.
- Possible presence of agents that supervise the design and operation of institutions, and where applicable, their motivations.

The ABCE as such does not offer answers to these questions as it is rather an approach than a theory. But without doubts, MAS models provide a systematic way of incorporating any classification that is relevant in the study of any particular economic phenomena. There is no need to focus on the general equilibrium theory and any *artificial* classification derived from the acquired knowledge of economic theory and related disciplines should be rejected.

In fact, the MAS allow the researcher to include all the wide spectrum of decision functions: from the neutral ones to the environment carried out by agents without cognitive functions to decisions that require specific and sophisticated cognitive abilities. These decisions are undertaken by agents who collect and process data actively. For instance, MAS can include structural agents (e.g. land, buildings), institutional agents (e.g. legal system, public sector, companies and markets) and cognitive agents (e.g. employers, consumers and financial intermediaries).

These agents may have an extensive list of characteristics but the widely accepted synthesis is the one described by Wooldridge and Jennings (1995). The main characteristics are:

- *Autonomy* to interact independently
- *Social skills* to interact through some communication system
- *Reactivity* to perceive changes in the environment
- *Proactivity* as a capacity of making decisions on their own.
- The degree of *rationality* is described by the probability of making the optimal decision taking into account information available.

Once the classification of agents is specified, data and relationships between each type of them can be defined using the available evidence from the empirical studies, results of econometric analyses, as well as surveys and interviews.

MAS are typically implemented using the object-oriented programming (Nguyen, 2008) to develop the skills of each agent and their relations. Hence, they allow starting the modelling process and lately introduce further required changes in order to match the model to the studied economic reality.

The model scale is a critical feature of MAS. This approach does not make use of the assumptions of standard economic theory (agents' aggregation and the representative agent paradigm) to simplify the specification and emerge the underlying processes. Prior to making the choice of the appropriate model scale, a number of issues have to be taken into account in order to obtain consistent results. First of all, we should answer the question how different needs and desires of individuals should be represented. Then, the procedure of fulfilling needs and desires by goods and services should be explained. Finally, the institutional aspects of production and distribution should be explicitly incorporated to the analyses.

The choice of the number of agents is critical for their impact on the economic activity. On the one hand, when market participants are sufficiently numerous, there is perfect competition among buyers and sellers. On the other hand, when a market consists of a seller facing many buyers the monopoly emerges, in which the only seller sets the market price to obtain maximum profit. There are plenty of intermediate positions, where the rivalry between buyers and sellers leads to imperfect competition. Modern economic theory is largely based on the assumptions of perfect competition. This modelling strategy has been used because of its convenience in analytical terms rather than its empirical foundations.

One of the greatest contributions that MAS could make to economic theory is the exploration of the effects of scale without the external imposition of artificial coordination devices. Nowadays, without defining in advance the behaviour of agents in the setting of prices and quantities, it is possible to simulate what would occur when the economy is facing a few thousand to several million participants. We should define and contrast whether the different economic effects on the regions are satisfactorily met by small-scale models, and which requires a scale closer to empirical reality.

The increasingly powerful computing abilities enables this kind of simulations to be carried out, what is needed now is to understand the reasons that lead us to use these large computing capabilities and discern when it is appropriate to treat a more limited number of agents to foresee the actual characteristics of the studied economies. In practise it turns into finding the maximum ratio, if not 1:1, with the same results and interactions pattern than in the system with the actual number of agents.

1.5. Multiagent systems for policy forecasting

Recently, an increasing number of multiagent systems have been used to predict and simulate the impact of public policies. Among the first works on this topic are those of Gintis (2006) that questioned the functioning of markets assumed in the general equilibrium theory and the core paradigm. The first methodology presented by the author was deficient in many aspects but certainly implies a remarkable advance in the right direction.

Some of the undersupplied issues were subsequently rectified in the group of models *Lagom* - Swedish word that means the *right quantity*- Since then, many articles about the MAS methodology were published and huge advances in the prediction methods have been made (Wolf et al., 2013). Further research on the potential use of MAS was also initiated by the publication (Dosi, Fagiolo and Roventini, 2010) that tried to consolidate Keynes and Schumpeter theories on business cycle and edogeneous growth through the dynamic properties study of economic aggregates and the public policies impact on the different components of economies.

Nowadays, this approach is becoming even more attractive as current advances in the computation capacity permits modelling a great number of agents and relations between them. Noteworthy examples of robust models are those such as Eurace@Unibi

(Deissenberg, Van der Hoog y Dawid, 2008, Cincontti, Raberto y Teglio, 2010; Dawid et al., 2011) and MOSIPS (Mancha et al., 2012; Pablo et al., 2014)[14].

MAS used in prediction and simulation of the impact of policies focused on the regional impact have to be developed in parallel to the elaboration of an adequate data base or designed to use an available big data source. The main characteristic of the data base to feed the model should be the assignation of the values to any agent and any relevant variable of the model, including agents' precise location and definition of relationship between agents in order to create realistic networks. If the available data do not accomplish the requirements it will not be possible to classify potential inaccuracies between model and data errors.

Currently, there is no public data base that could meet those requirements. For that reason, the database has to be constructed by two procedures presented in the following part of the chapter. Those procedures are still uncommon in the social sciences. The first one includes *statistical matching* and it allows joining micro data from various sources such as the population census or labour force survey. The second one, *downscaling*, becomes crucial due to the location importance, allows understanding how the territory and the spatial dimension of the model has to be incorporated.

Statistical matching is used to join multiple data sources into a single one combining their variables whenever there is no common identifier in the data. A good example is the combination of data sources containing two complementary variables such as income and family expenditure. Notwithstanding, in the case of more sophisticated models, the elaboration of data bases for the purpose of MAS for regional and local forecasting and simulation, is much more complicated. A great number of data sources have to be joined into one. Moreover, different data sources have their own different

[14] The aim of MOSIPS, previously mentioned, is to model regional economies in the more realistic way, and to evaluate the impact of public policies on SMEs. Then, not only agents are described but the emphasis is put on relations that are determined basically through spatial variables. For that reason, not only shall the main characteristics of the agents be defined but also their location, as well as the possibility of sharing information and exchanging products, services, inputs and financial assets should be taken into account.

characteristics. For all that reasons, the process of statistical matching has to be carried out with great precision and taking into account several issues (Sutherland, 2002).

- Two samples are essential to have a significant number of variables in common and then also others not necessarily directly comparable. Then, the researcher should be able to distinguish between primary sample and the complementary one. The first one gives us a structure, data and the sample size while from the second one, only additional information is extracted.
- The next step is to homogenize data to correct the heterogeneity problems in different aspects and dimensions (units, territory, population definition etc.). The appropriate linkage variable between both databases has to be established to allow matching observations.
- Once the two previous steps have been carried out, the merger itself can be done in two phases. In the first one, cells of similar observations are created which are grouped according to predefined criteria for matching variables. In the second one, observations in the same cell are grouped using other matching variable.

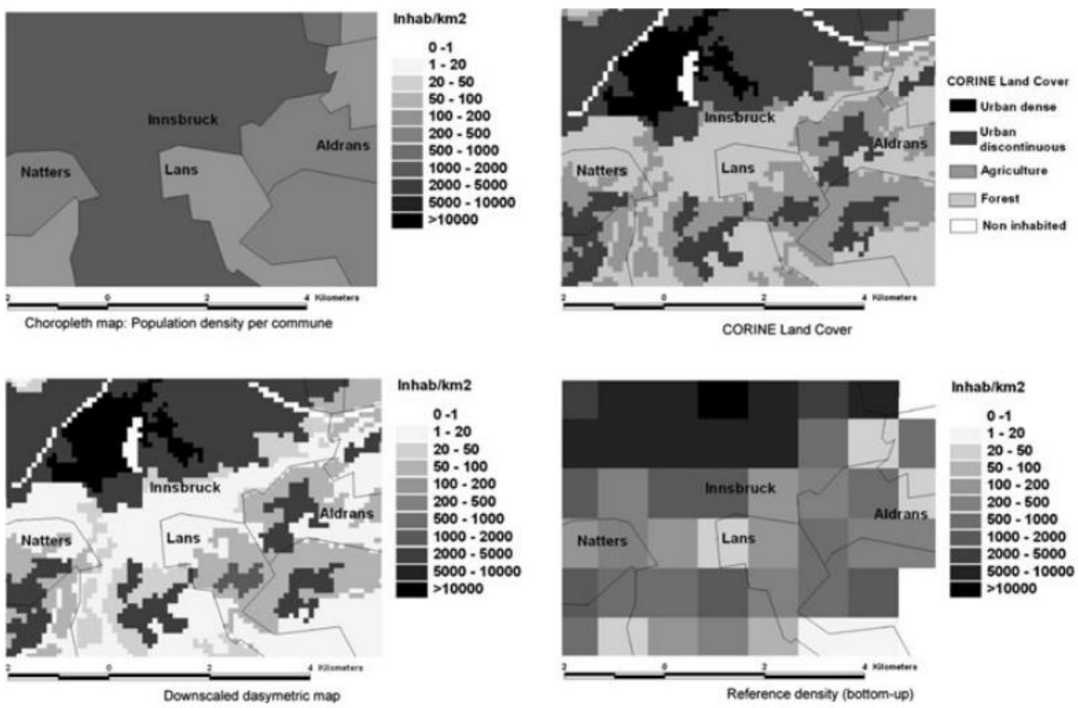
Matching variables are those variables that are common to both samples and they are used to match observations. In order to carry out the matching process correctly, variables that are not common to two samples shall not be related to each other only through the matching variable what is known in the literature as a conditional independence (D'Orazio, Di Zio and Scanu, 2006).

Joining the population census that provides data on residence with data on mobility from additional survey reporting where individuals work can help to verify buying patterns and to explain how they can be modified by taking into account that additional information (Schenk, Löffler and Rauh, 2007). In this paper it can be seen the change in a household expenditure distribution before and after performing statistical matching between population census and labour force survey. After including the workplace, the consumption in every near shop change significantly.

Downscaling is a statistical technique that unifies spatial data on different levels of definition to obtain the results that quality exceeds the one in any of the sources of data used in the procedure. This technique was firstly used in studies on climate, but now it has shown its relevance in other fields such as economics.

In order to represent relations between individuals properly, information about their geographical location shall be included. The European Union has already noticed the importance of implementation of geo-reference in such analyses and in the upcoming years the obligation of collecting such individual data will be introduced. However, as currently no such database exists; the most widely used solution consists of joining several sources by the down-scaling procedure to obtain a grid to set the location of agents and which data can be extracted to the model from Gallego (2010).

Figure 1.1. Example of downscaling for population location



Source: Gallego (2010)

The agent behaviour is based on their characteristics among which information about their place of work, residence or mobility pattern are included. For instance, if an individual has a property in the luxury area, there is a greater probability that this individual could easily borrow from banks or other financial institutions. Companies differ in terms of effective demand and costs depending on location. In addition, environment is determined by the characteristics of individuals and companies. These characteristics as well as the environment define and change their behaviour. Data on location of economic agents shall be complemented by information about the use of roads, public facilities, as well as price of the land.

While in conventional models all actors participate in making transactions on one market, in the correctly defined MAS, there are plenty of markets and agents. Each individual has no access to all companies in the market, but only to the closest ones and the largest ones, that is the shops she knows about. For that reason, it is essential to define the concept of visibility. Normal type functions can be set depending on the locations of agents, the sizes of companies and distance to them. The visibility changes depending on distance.

In the literature, this concept is illustrated with retail shopping. Individuals use to know the prices of *all* neighbouring shops while only some of the big supermarkets far away from their residence. Additionally, in the second case the cost of transport shall be added to the prices according to the agents´ degree of rationality. For that reason, every individual does not face all the supply, but its demand can only be positive in some establishments. The introduction of transport infrastructure alters distances and in the consequence visibility and purchasing patterns. The same happens when more data sources are added with statistical matching.

Hence, the visibility of companies along with the location of agents is what makes it possible to establish commercial relations and exchange products in the markets defined in the multiagent systems. In opposition to MAS, in the CGE models all the agents trade with everybody at the same time. This assumption is unrealistic and results can differ from genuine ones.

On the other hand, MAS developed the heterogeneity of agents without taking into account their location and their relations would lead to the results that can be obtained by assuming the general equilibrium theory approach. For instance it would lead to the same results as the ones obtained in CGE. Therefore, these widely used models can be included within the wider group of multiagent systems taking into account that in this case simplistic and unrealistic assumptions are assumed and in consequence results would be different to those obtained while the correctly developed MAS are used.

Using the downscaling and describing properly markets in the MAS, the spatial levels of economic integration can be defined within certain autonomous units. At the local level there are areas where determined group of individuals and companies made a large proportion of its economic and financial operations with other agents from the same integrated area. The importance of spatial information underlies in large part in the fact that it allows identifying which of these areas are likely to be affected to greater extent by changes in policies or the implementation of infrastructure projects of any kind.

In MAS, the concept of agglomeration economies can be introduced. Two areas where no economic activity is undertaken are different in terms of being appealing to the potential investors as it could be used in future as enterprise area or residential one, allowing forecasting the spatial limits of urban agglomerations. The technique of downscaling can be used to determine different prices to areas that seem to be similar in many aspects but that have different locations, for instance agrarian areas close to a large city and those far away from urban areas.

1.6. Artefacts in multiagent systems

The empirical validation of the model is a relevant part of the research process in both cases: traditional models and multiagent systems. Notwithstanding, researchers are aware of the existence of some problems that may occur during the validation process, particularly if we are using MAS methodology (Galán et. al., 2009).

First of all, there is a problem related to the degrees of freedom. MAS usually have many parameters that define the system. In the consequence, researchers are able to obtain the pre-defined result or desired feature by adjusting the degrees of freedom. This untruthful procedure is getting more widespread as more and more learning algorithms and functions are available for researchers to be modified.

If it is used a more generic model that enables agents to learn, one shall decide if information should be stored in a neural network or whether a prediction model should be used. Consequently, on the one hand, the modelling tools can be used to model a wide range of economic phenomena and a fit of data can be improved notably. On the other hand, it is not all about the flexibility. The model shall be always chosen after the detailed analysis of behaviour rules and relations between the agents that appear in the context of the economic activity.

The proper definition of the MAS seems to be another problem. Most of the special properties of the MAS are often not correctly defined and they do not reflect appropriately the human behaviour observed. Despite of the problems previously mentioned, there are reasons to expect that researchers will be able to develop a robust methodology of empirical validation. Currently, the empirical validation process seems to be a greater challenge of the MAS approach than the construction of such models. In the future, once the necessary procedures of validation will be developed, researchers shall proceed to connecting the agents´ behaviours to the real ones in order to conduct experiments and extract conclusions about the real world.

There exist other empirical validation methods behind the statistical data processing. One of them is the replication of the empirical characteristics at many levels and at multiple scales. In such context, MAS offer the possibility of obtaining new empirical evidence, since the economy is described by micro data that is previously generated from macro data. This procedure is not available and cannot be used in case of other standard economic models.

Certainly, the MAS approach offers much more accurate validation methods than the traditional modelling one. For instance, there is a possibility of testing whether the simulated behaviour of the agents is well aligned with the real subjects. Therefore, it is essential to note that, beyond simple time series, the MAS allow observing and analysing the complete distribution dynamics of the economy. The features of the MAS system such as the wealth distribution or company sizes can be compared with the corresponding real distributions of the economy. The micro-level distribution could be treated as an approximation to the empirical validation, but it requires data information that is not always available (Takadama et al., 2008). However, even limited data samples provide an important feedback about the empirical plausibility of the model and the simulated distributions at micro level.

At the same time, it is worthy to note that MAS can display complex dynamics, but non-linear behaviours that appear while representing the interconnections of agents´ activity; they may raise difficult questions to traditional econometrics. Summarizing, the properties of the MAS approach previously mentioned make the empirical validation of such models interesting but at the same time much more complex than in the case of traditional ones.

It is also essential to raise the issue of potential applicability of MAS in the public-policy impact analysis. Complex agent-based models with dimensions and relations between agents not based on the reality contrasted with the data cannot be developed *ad hoc* by the researchers. The potential of the MAS approach is huge and policy-makers may try to use the MAS to expand their knowledge and to evaluate the incidence of the public policy on the society, economic activity, as well as to expand their power. Policy-makers could use the MAS to explore the important changes in the political configuration, taking into account behavioural rules of the agents that are defined in the system and that may be affected by the incidence of policies and public decisions.[15]

Another main problem of the MAS is related to the violation of the law of one price, but as it has been shown it occurs only in the short term. In the long term, the average price equals the historical equilibrium price. Moreover, there is empirical evidence that there exists a significant dispersion of prices on the markets, including the markets deeply integrated, such as national markets within the European Union interior market. This feature of the MAS could also be seen as a positive characteristic of this modelling methodology.

The process of imitation is fundamental in MAS, particularly it is an essential part of evolutionary models, but it is considered a second-class learning in economic theory and applications. For instance, in case of the genetics, the offspring inherits all the genes

[15] The remarkable example is the one of the Obama campaign during the U.S. elections in the 2012. In this case, the use of micro data elaborated and processed by the researchers using the techniques previously described, that model and analyze main behavioural rules, was crucial to winning the elections focusing only in those agents who were prone to change their choice (Beas, 2012).

of the parents, with a very low rate of mutation. The cultural transmission is the dominant process in biology, anthropology and sociology, and the importance of culture depends largely on the ability of agents to imitate. This mimetic capacity is much higher in human beings than in the case of other animals. By contrast, the individual learning, especially merely making use of the personal experience, is the most widespread mechanism in economic theory, though it is usually slow and inefficient. Moreover, markets volatility makes the imitation highly ineffective (Conlisk, 1988).

In order to define the imitation patterns correctly, the networks of agents shall be developed adequately. The establishment of such networks between agents is highly subjective and no sufficient empirical backup has been given since there are no statistics and studies sufficiently developed to provide information about the way individuals convey information and how they respond to economic issues in a determinate moment of time, such as job searching, investment and consumption.

There also have been other types of problems indicated such as those related to definition of capital accumulation mechanisms. Notwithstanding, sooner or later they will be solved as it also occurred in the case of the general equilibrium models. On the other hand, the validation of MAS fully based on probabilistic determination will take longer to occur, but there are some steps in the right direction (Grazzini & Richardi, 2015). This difficulty comes from the multitude of sources used, and the process to join and link agents with the space correctly. The aggregate behaviour will be sufficiently close to reality from a certain level of disaggregation (region, municipality and even neighbourhood) and it never will be at the individual level due to the randomness of the processes carried out within the model, and the random construction of the agents. The challenge is to find the point from which the results can be considered valid, as this point depends on the variables involved and the results to be achieved.

1.7. Final discussion

In order to compare the three types of methodologies: MAS, CGE and spatial econometrics models, and to extract the main conclusions from the performed analysis, the main features and characteristics of them are described in Table 1.1. Although all of them have their strong and weak points, the superiority of the MAS is evident and the main conclusion of the analysis is that MAS can be used to analyse complex economic phenomena and to obtain reliable results.

With respect to the famous Lucas critique, both computable general equilibrium models and MAS are not affected by this critique and the incidence of public policies potentially can be predicted adequately by each one of them. On the contrary, spatial econometrics models do not adjust according to the behavioural rules changes and it is considered to be one of the main weaknesses of this approach. At the same time, one shall be aware of existence of the Velupillai critique, especially in the context of regional and spatial economics. Both CGE and spatial econometrics use equations and are directly affected by this critique while MAS approach uses the algorithms implementation and it overcomes this already not been replied critique.

Another feature that allows comparing the models is the level of spatial disaggregation. The CGE models are developed at national or regional level while spatial econometrics and multiagent systems use disaggregated data that may have any scale. Regional CGE are not infrequent but a higher spatial disaggregation difficult their operation. Several attempts have been done in order to introduce a higher spatial disaggregation in CGE. A prototypical example can be found in Bröcker (1998). However, regional CGE ares still infrequent.

Simultaneously, the adjustment capacity to represent realistically the socio-economic relations and the economy is low in the case of spatial econometrics models, medium in case of the CGE and higher in the last case of MAS. The lower capacity of other models is related to the econometric specification that can be adapted only slightly or to the modular and the limited redefinition of equations in case of the computable general equilibrium models.

Table 1.1. Comparison of MAS, CGE and spatial econometrics models

Characteristic	Computable general equilibrium models	Spatial econometrics	MultiAgent Systems
Affected by the Lucas critique	NO, they can predict adequately the incidence of public policies	YES, their parameters do not adjust according to the behavioural change	NO, they can predict adequately the incidence of public policies
Affected by the Velupillai critique	YES, they use equations	YES, they use equations	NO, they are implemented with algorithms
Spatial disaggregation	LOW, they are built at national or regional level	HIGH, the spatial disaggregation can have any scale	HIGH, the spatial disaggregation can have any scale
Adjustment capacity to represent realistically the economy	MEDIUM, they are modular and their equations can be redefined	LOW, the econometric specification can be adapted slightly	HIGH, these models are not based in any pre-defined theory and can take any specification
Prime learning process	EXPERIENCE, importance of long time series	EXPERIENCE, importance of long time series	IMITATION, importance of networks
Heterogeneity representation capacity	LOW, the representative agent paradigm can be removed and it is possible to include agents with different characteristics (e.g. income level)	MEDIUM, the spatial heterogeneity can be included in the models	HIGH, the actual degree of heterogeneity can be included in the models
Verified analysis capacity	NON CONSENSUS, their overall performance is accurate in growing periods but they fail in economic depressions	YES, they accomplish their aim and they are able to extract the effect of every determinant involved in the studied process	NON CONSENSUS, they have revealed their capacity in other sciences but their use in economics it is incipient
Standard evaluation method	YES, there are methods to evaluate their performance and adjustment to real data	YES, there are plenty of econometric tools to evaluate their performance and significance	NO, these models cannot be evaluated following a standard method yet

Source: Own elaboration

One the main characteristic of the MAS is the definition of the prime learning process. In any conventional model, the behaviour is assumed to be based on experience and the estimation technique requires the use of long time series. On the contrary, in the MAS approach the importance of networks is emphasized. Moreover, to the MAS the actual degree of heterogeneity can be included in the models while in case of the spatial econometrics models only the spatial heterogeneity can fully be taken into account.

The heterogeneity of agents is completely omitted by the CGE models that usually assume the representative agent paradigm. There are new advances in this context and this paradigm can be removed to make possible including agents with different characteristics such as income levels but this field of studies is still not sufficiently explored.

The next feature is the possibility of empirical verification. In case of CGE there does not exist a wide consensus on their overall performance. Their accuracy is outstanding under normal conditions but they have no capacity to predict or perform well during stress periods such as during the recent financial and economic crises. The spatial econometrics models have been taken under empirical verification frequently. They accomplish their aim and they are able to extract the effects of every determinant involved in the studied process.

The most unexplored field is the empirical verification of MAS. There is no wide consensus on this theme though the MAS have revealed their capacity in many fields of studies (social science, biology, politics, etc.). The main difficulty of MAS is still its reduced applicability and empirical verification in economics in a standardized way. Hence, one of the main drawbacks of the MAS applied to economics, as it occurs to every single new methodology recently incorporated in any field, is the lack of existence of the standard evaluation method.

However, this methodology allows carrying out the validation at multiple levels: aggregate results, spatial results at multiple scales and individual performance, having into account that in the last case it is unfeasible to achieve complete accuracy. On the contrary, CGE and spatial econometrics models have already defined standard evaluation methods that permit to access the performance and significance. Nevertheless, if the objective is evaluating the impact of public policies adequately, MAS emerge as the convenient alternative.

In conclusion, the need to assess and predict the effects of public policies made necessary the adoption of CGE models. However, these models have some drawbacks such as the homogeneity of agents and their inability to predict abrupt changes. Hence, they cannot report on the spatial effects and the resulting distribution of the included variables. These models do not where and whom public policies will impact positively or negatively and in what amount. Moreover, by their nature tending to equilibrium, they cannot correctly predict the path in recessions like the current one that is affecting a large number of countries in Europe.

Due to these serious shortcomings, it is necessary to promote a new type of model that can predict the crisis and the effect of policies on individuals. Because of the drawbacks of spatial econometric models, it is essential to find a new kind of approach in macroeconomics. This novel approach seems to be the MAS, in which the emphasis is on individual agents, the differences between them and their relationships. The behaviours of the agents are determined by the environment and relations with other individuals, they can learn by imitation, as it happens in the reality, although economic theory has hardly been taken into account. Nevertheless, MAS have their own problems and needs: the difficulty in the validation process and the choice of the convenient scale. Furthermore, in terms of needs, statistical matching techniques and downscaling should be developed according to the objectives of each model.

CHAPTER 2

A multiagent system to forecast employment and migration in the European Union at regional level

2.1. Introduction

The aim of this chapter is to present a multiagent system for predicting the population of the European regions at NUTS 3 level[16]. Before proceeding to the presentation of the model details and description of the preparation of the databases needed to feed and calibrate this model, the latest advances on this topic are presented. Thus, it will be introduced how multiagent systems (MAS) have been used to model the population dynamics, paying also attention to the importance of labour market.

Nowadays, population dynamics is a field of study of the highest interest. It joins the results extracted from the economics, demography studies and sociology. All of them have their own methodologies of studies and all of them present new techniques of population dynamics and related phenomena modelling (Stillwell & Clarke, 2011, Wilson, 2011).

Silverman et al. (2014) present a seminal paper of an agent-based model of the dynamics of mortality, fertility, and partnership formation. They “propose that directly linking demographic methods with MAS frameworks will allow us to produce models which increase our understanding of population change, while simultaneously helping us to avoid the pitfalls of an over-dependence on empirical data. MAS allow us to produce models which have a greater explanatory capacity, while the demographic components allow to use the inherent flexibility of the MAS approach to generate plausible scenarios within a given parameter space”.

In the article they present an agent-based model that emphasizes the partner formation and they combine it with survivorship and birth rates both empirical and projected. This agent-based model, developed at the same time, follows the same path and develops an agent-based model with the same features as the cited one, but this one is focused on the importance of the labour market and migration patterns. These two relevant parts are not taken into account by Silverman et al., but they may be fundamental in economies such as the ones of Southern Europe countries, where long-term unemployment and

[16] NUTS (Nomenclature of Units for Territorial Statistics) are the territorial units in the European Union. NUTS 3 are equivalent to the Spanish provinces.

inexistent recovery expectations urge a part of the population to change their location. Hence, migration patterns within European countries also affect the overall population dynamics. These patterns have been incorporated in MULTIPOLES model. It has represented an important step in the next generation of models to study the population dynamics, and it includes births and deaths and migrations, improving the results obtained by the official statistics services (Kupiszewska & Kupiszewski, 2005).

The key feature that has not been taken into account in most of MAS of population dynamics is migration. The MAS technique is one of the most promising methods that have been used in population and labour market modelling. In the report by Zhang and Jager (2011), the model allows the agents to decide whether to migrate or not. This approach undertaken in the PRIMA project follows partially the same track undertaken by the one used in MOSIPS project (Mancha et al., 2012; Pablo-Martí, 2013; 2014) and in this chapter.

The scope of PRIMA and this MAS are widely divergent, and this difference is reflected in the variables and methods used in both of them. The aim of the PRIMA model is to study the migration dynamics in 14 wards[17] in the UK while the model developed in this section covers all the European regions and its aim is not only to model migration but the interrelations between labour market interactions, migration decisions, education demand and demographic processes in order to achieve a correct knowledge of how these different decisions are interconeected and how they affect the population dynamics.

One of the most important techniques that has been used in study of population dynamics and especially labour market policies incidence are micro simulation technique and agent-based modelling. As Rahman et al. (2013) underline in their paper "these days spatial micro simulation modelling plays a vital role in policy analysis for small areas. Most developed countries are utilizing these tools in more effective ways to make acknowledgeable decisions on major policy issues at local level".

[17] The ward is the primary unit of English electoral geography for civil parishes and borough and district councils

The authors apply this technique to study the housing stress phenomena and demonstrate its variability in accordance to changes in the geographical units under study. The main characteristic of this paper are not only the results extracted from the analysis, but the whole technique of micro simulation as well as a proposed validation technique to perform the statistical test of the spatial microsimulation modelling estimates. The concept of housing stress pointed out by the authors has plenty of definitions in the literature; the most common is to consider that it happens to a household that spends more than 30 percent of its income in housing.

This concept highlights the importance of the high unemployment rates in many European regions in recent years. This high unemployment has meant a dramatic decrease in family income for a significant number of households. Moreover, in a number of countries a major portion of these families have to make mortgage repayments and subsequently the housing stress is a relevant phenomenon in countries affected by the current downturn.

Although pure demographic topics and population dynamics in general seem to be the areas of research of the prime importance, the interaction between demography and the economy growth and development is not less relevant. Elgin and Turnen (2012) raise the questions of whether sustained economic growth and declining population could coexist and if there is a possibility to reverse the decline in fertility. In fact, this is the question that has been raised in this chapter. The main conclusion obtained is that returns to human capital in production can help us to understand that contradiction.

Whenever the degree of increasing returns to human capital falls, developed economies can take advantage of the so-called endogenous efficiency-augmenting mechanism. This mechanism describes the possibility of switching to "human capital oriented technologies that support aging population". Taking into account that possibility, the coexistence of sustained economic growth and a declining population is allowed in the long run.

However, "the degree of increasing returns to human capital has been falling over time throughout the world along with population rates". The MAS developed in this chapter takes into account this recent empirical finding and it is built allows a simultaneous increase of human capital stock when the long-term unemployment rises, as found in the empirical data, while the regional population can decrease due to the dynamics of mortality, fertility and migration. Thus, the micro analysis includes certain aspects of major importance in the analysis of population and the impact economy and the role of agents' organization on its dynamics (Geard et al., 2013).

One of the most valuable articles related to the topic of artificial labour market creation is Chaturvedi et al. (2005). These authors emphasized the usefulness of multiagent simulations in many areas of research and their huge applicability, including the labour market and population dynamics. In the article the creation of an artificial labour market (ALM) as an agent-based simulation model was discussed and two phases of ALM development were indicated: modelling and simulation phases. In the modelling one, the representation and specification are defined, while in the simulation, the scenarios are built to allow computational experimentations.

The example implemented by the authors was the one of the military recruit market. The goal of this paper has been to underline the benefits of the ALM development that are, without doubts: supporting market segmentation and facilitating an integrated decision support functionality. Moreover, the agent-based simulations can be used to develop a platform for building an ALM that is commonly used to test different labour market policies incidence.

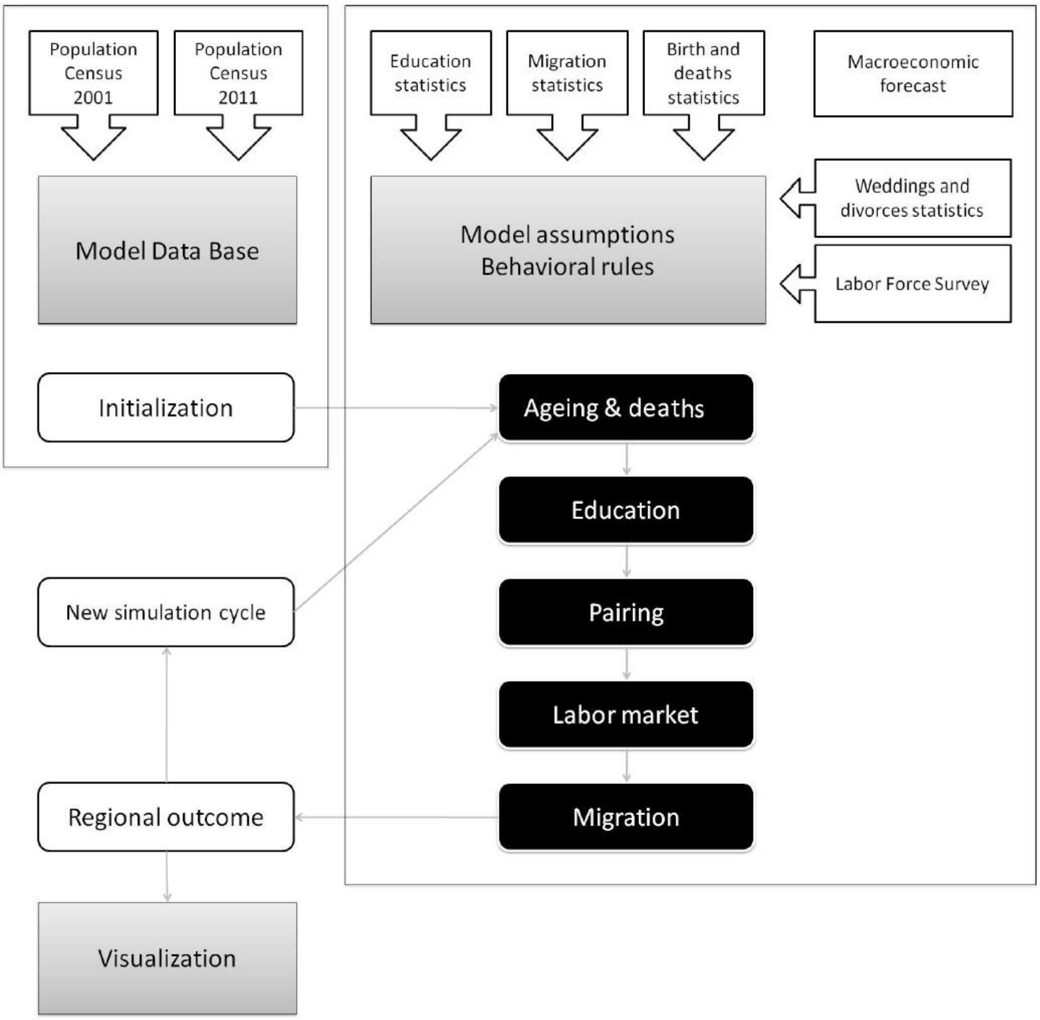
Hence, the final goal is to program micro-level agent behaviour that would have macro-level effects and that would, through specific calibration, reflect economic reality. Moreover, the use of the so-called parallel worlds, specific interfaces used to switch from the real data to the virtual world, is especially useful during the scenario development and then during the decision making process for each of those predefined scenarios. The add value of this research was to elaborate a new methodological approach for building a large scale, macro-level synthetic economy, which works from artificial simulation life cycle principles, as well as to develop an effective representation techniques for synthetic economies, which supports market segmentation.

Then, we consider the methodology applied by Chatuverdi et al. (2005) is an advance in the right direction in labour market and population dynamics studies, especially because it allows focusing the main attention on individuals.

2.2. System overview

The aim of the multiagent system hereby presented is to forecast and simulate the population dynamics of the European regions at NUTS 3 level. In addition it makes it possible to forecast the education demand of individuals and the labour market outcomes. These results can be displayed for different kinds of individuals depending on their age, gender, nationality, education level and labour status.

Figure 2.1. Overview of the model and the needed data bases



Source: Own elaboration

In order to make possible these tasks, it is necessary to build a specific database. We use the population census carried out in 2001 in order to build the population and the microdata of the census of 2011 for executing the calibration of the parameters.

This database built from the population census is used to initialize the model. Then every agent passes sequentially through the five implemented modules:

1. Ageing and deaths: Every agent computes a probability of death and the module determines the ones who decease. Widowed and orphans agents are taken into account. The rest of the population increases its age by one year.

2. Education: The individuals can start, end and continue different levels of studies depending on their age, gender and labour status.

3. Pairing: Adult agents of one gender decide if they want to divorce, get married and have children.

4. Labour market: Individuals have the choice of turning into inactive or starting looking for a job. The labour demand is aggregately computed for every region and the hiring and firing processes are carried out accordingly the defined labour suitability of the agents.

5. Migration: The agents compute their satisfaction with their labour status and the probability of housing stress is calculated. Depending on the result the individual decides whether he/she stays in the same region or changes location. In the second case, the family also modifies its location.

An iteration of the system represents a year. After these five modules have been carried out for all the individuals the regional outcomes are computed in order to carry out the visualization. Then a new simulation cycle starts.

The system is stopped after twenty iterations, making it possible to simulate the actual results from 2001 to 2015 and forecast the trends from 2016 to 2021. The end point is set to 2021 because the MAS can use a pool of macroeconomic forecasts of international institutions up to this date. However, as is well known, these forecasts have been wrong several times in the recent past because of the drawbacks of the used methodology (Solow, 2010).

In order to reduce the problem we have computed the sensitivity of a deviation of +/- one percent in the growth of the GDP in all the regions during the first eleven iterations. Then, in the worst case it had an impact of 5.23% of the overall population in a period of eight years (the highest error was in region R0321 -Bucarest-for the period 2004-2012). Consequently, eight years is set as the maximum forecast scope of the model in order to control the uncertainty derived from the use of the macroeconomic forecast.

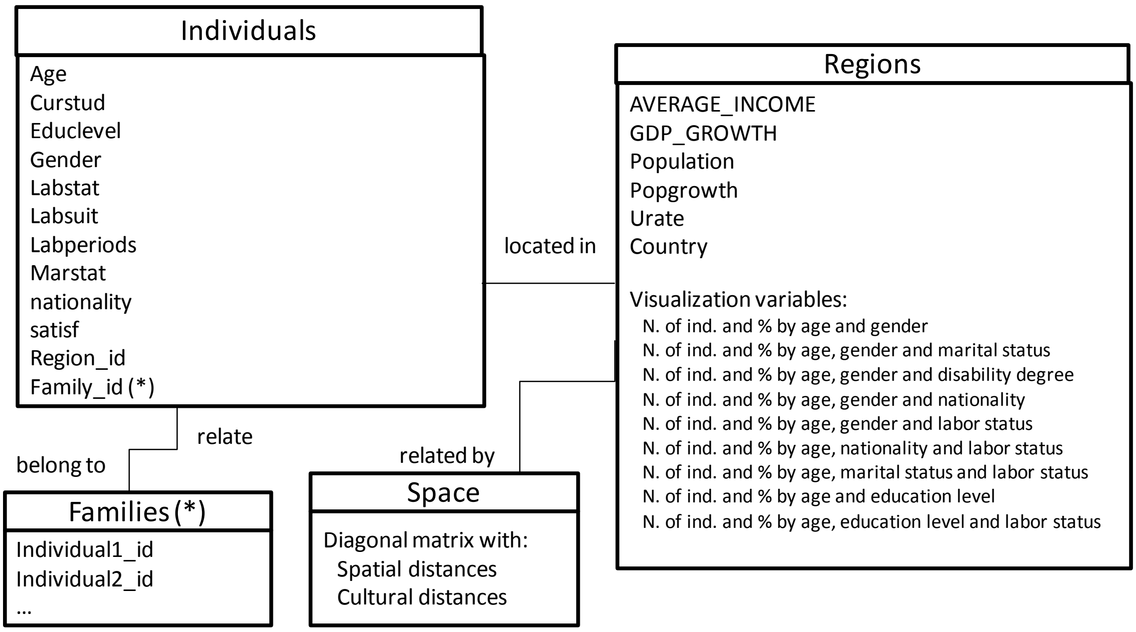
Nevertheless, other variables of the model, especially the labour market results are more dependent on this forecast. The associated uncertainty necessarily has to be higher than the boundary of five percent. The author is aware of the efforts to increase the predictions time scope, but researchers must be conscious of the uncertainty of their presented results.[18]

The model has the following objects:

- Individuals: These agents carry out the actions and populate the system.
- Families: Individuals can be aggregated in families in a graph database. Families can make decisions together such as migrate.
- Regions: Individuals are located in one region and they can change their position. Regions include variables of aggregate outcomes and macroeconomic forecast.
- Space: Matrix that includes the spatial distances between the centroid of every NUTS-3 region and cultural distances.

[18] A paradigmatic example of the pointed issue is the project "Population Pyramids of the World from 1950 to 2100"(United Nations, 2011) that after only five years since its second version release it is incorrect in plenty of its first results for the year 2015.

Figure 2.2. Class diagram of the agent-based model



Source: Own elaboration

INDIVIDUALS are the agents of the model. Their actual heterogeneity is preserved and included in the system in the following variables as shown in Figure 2.2:

- Age: Age is expressed in years. It allows identifying dependent children: those individuals whose age is lower than the minimum legal working age. Age is the main variable that permits identifying different behavioural patterns in the system.
- Curstud: It includes the education level in progress, otherwise is zero. When its value is positive the individual is studying and so cannot be part of the potential labor force.
- Educlevel: It informs about the highest awarded diploma of every individual and it is used as a proxy of the human capital in the computation of labour suitability in the case of unemployed agents.
- Gender: The Boolean variable adopts 0 for women and 1 for men.

- Labstat: The labour status identifies if the individual is working or not (labstat=3 for employed agents) whether looking for a job (labstat=2 for unemployed) or not (labstat=1 for students and labstat=4 for inactive agents)

- Labsuit: This variable is used to compute the labour suitability and hence reflect the probabilities of getting hired and fired in the iterations of the simulation.

- Labperiods: The number of years since the last change in the variable labstat.

- Marstat: It is another source of behavioural variation: 1 for single, 2 for married, 3 for widowed and 4 for divorced agents.

- Nationality: The nationality is divided into native, other EU-28 and non EU-28 agents.

- Satisf: The overall satisfaction of the individual current location. It is compared with the expected one in other locations.

- Region_id: Region when the individual is located.

- Family_id : Family identifier of the individual. It is not used in the current version of the model in order to simplify the presented model algorithms. This variable allows reducing the uncertainty in migration decisions, where all the individual of a household can change their location together when a family is defined as a couple and their children under the legal working age.

Every REGION has a set of variables that affects the behaviours of the individuals:

- AVERAGE_INCOME: This variable is estimated using times series analysis and macroeconomic forecast information.

- GDP_GROWTH: Actual data for the period 2001-2012 and macroeconomic forecast estimations for the remaining iterations.

- Population: Total number of individuals in the database.

- Popgrowth: Change in the population between the current and the previous iteration.
- Urate: unemployment rate of the region.
- Country: Country identifier.

The remaining objects are the FAMILY that includes identifiers of the individuals contained in them and the SPACE used to save spatial and cultural distances used to determine the migration patterns.

2.3. Initialization and modules in the system

The model can be initialized in 2001 and 2011, corresponding to the times when the last two population censuses were carried out in the European countries. In addition to these databases, a set of additional data sources are needed to carry out the calibration of the agents behavioural rules, incorporated in the model with set of coefficients. The macroeconomic forecast it is computed with a collection of forecast changed as stated below in the section titled "Labour market".

Table 2.1. Variables obtained from the population census

Description	Name in the database
Region	NUTS-3 (reg)
Gender	gender
Age in years	age
Nationality	nationality
Marital Status	marstat
Number of children with age<16	children
Education level	educlevel
Current studies	curstud
Labour status	labstat

Source: Own elaboration

National Population Censuses offer information of the population at municipal level for all the European countries, National Statistics Departments also supply microdata for a percentage of the population. The purpose of the model is to forecast the result at NUTS 3 level. Hence the available data are aggregated into NUTS 3 regions. The MAS makes use of only a small part of the variables included in the microdata source.

In addition to these variables the Population Census has the household identifier. The current model does not require identifying families as groups of parents and dependent children (those who are under the legal working age) but the randomness level is reduced when this information is included. Although, in the population census this information is not directly available, it is possible to identify couples and their children, and then arrive to the correct number and composition of families according to the definition.

After the error and null values replacement in the microdata, the rest of the population is initialized in order to represent the actual population in every NUTS-3 region. Nevertheless, it would be incorrect to create multiples of each observation because it can be easily proved that, against the statement of the Statistics Offices, the microdata are representative of the population if I only take into account variables such as age and gender, but microdata do not represent the population when education level and labour status are considered.

Due to the previous mentioned reason, in the initialization process the agents are created and the values for the variables are defined according to this procedure:

1. Compute the number of individuals for each region in the year 2001 in the following groups: age (five years ranges), gender (male and female), marital status (single, married, widowed and divorced), nationality (native, rest of EU28 and others), labour status (unemployed, employee, inactive).

2. Each individual included in the database is subtracted from the aggregate value of his correspondent group.

3. The process ends when the results in each of the 1,440 groups of population are selected.

It is not possible to compute the marginal groups in every case at NUTS-3 level. In these cases it is assumed a distribution equal to the one computed for the NUTS-2 or NUTS-1 level.

2.3.1. Ageing and deceases

The ageing and deceases module has two methods[19]. In the first one every individual computes its mortality probability and according to this probability he deceases or not in the subsequent time period. Otherwise, the individuals who continue in the population increase their age. This module takes into account the differences between genders, and how these differences vary with the age. It also includes the high morbidity probability that affects new-borns.

In the second method when a person who had children passes away another individual in the same region and higher age become the new parent of the children. When a child is eliminated from the population its parents decrease by one their number of children.

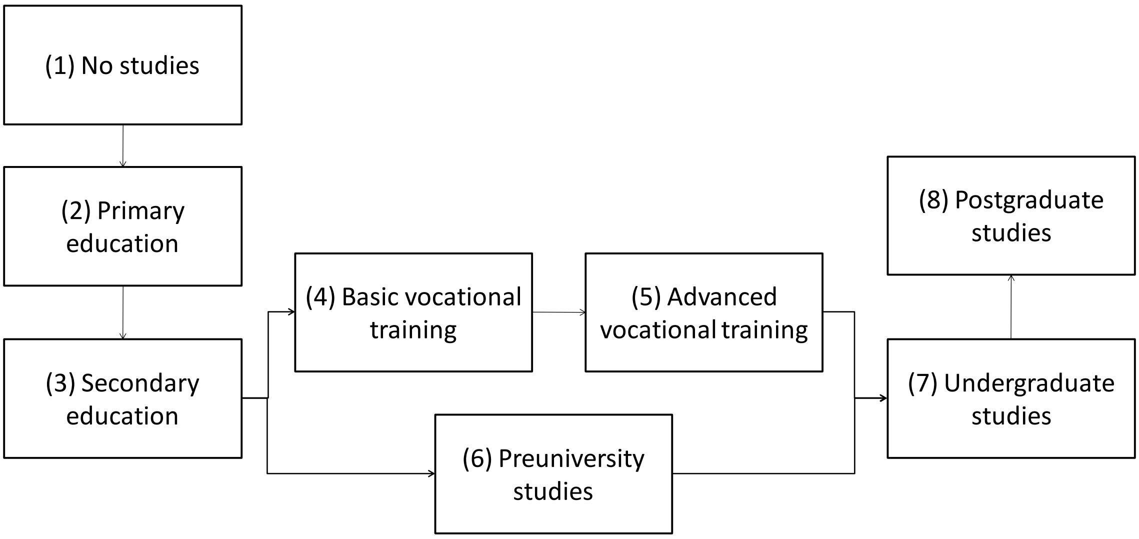
2.3.2. Education

The education module has two methods. In the first method if the individual is studying and he ends his studies he increases his education level and changes his labour status to unemployed. In the cases where he is studying and does not end the formation period, he just increases his counter of education periods and continues. Children start the compulsory education when they are five years old.

In the second method the agents who are not studying decide if they want to start the next level of studies or not. There are two possibilities: The usual path is to start the next level of studies just after ending the previous one. In addition, other people, mostly long-term unemployed, decide to increase their human capital to increase their chances to get employment. Both circumstances are taken into account in this second method. In this way, a region with crises will increase its stock of human capital and it even could increase its GDP in the long run as stated in 11. The model assumes the education is a full-time occupation and it does not allow working and studying at the same time.

[19] The algorithms and the activity diagrams are presented in the appendix in the last part of this chapter in order to make the flow of ideas more direct.

Figure 2.3. Education levels in the model



Source: Own elaboration

In the figure 2.3 it is shown how an individual can pass from the lower level of education to postgraduate studies. As it can be seen there are not backwards paths, as it is usual in the actual education trajectories. Once the individual ends the compulsory education (level 3: secondary education or high school) he can select vocational training or pre-university studies. The model takes into account the different probabilities according to gender: Women start pre-university studies more often than men in most of the countries of the European Union according to Eurostat information. Men prefer to undertake vocational training with a higher probability.

2.3.3. Pairing

In this module individuals decide if they want to divorce, get married and have children. The model assumes a woman and a man form a couple. This assumption is made to simplify the model. Nevertheless, in next versions of the model this constraint will be removed in order to study in depth the migration and reproductive behaviour of same gender couples. In the pairing module every married woman decides if she wants to divorce, then every non married woman decides if she wants to get married and if the choice is affirmative she determines the characteristics of her partner. Finally, every woman decides if she wants to have a new descendent. The specification of the module could be changed to include both gender decisions but the computations of the coefficients would be more complex and the results would not vary overall. In addition, non-married women take the decision of having children in practice while non-married (non-paired) men who have children are still infrequent.

Nationality is taken into account through the different calibration of coefficients appointed as α_{xx}(nac) in the code. In this way the model takes into account the different probabilities of divorce, marriage and parenthood that every nationality has.

The three activities undertaken in this module follow normal distributions according to the age of the ex-wife, fiancé or mother respectively. In order to compute the probability in each activity it is computed the area under the normal distribution in the range (age-1/2, age+1/2). The result of this defined integral is the probability to perform the indicated activity.

The nationality is taken into account through the different calibration of coefficients pointed as $\alpha_{xx(nac)}$ in the code. In this way the model takes into account the different probabilities of divorce, marriage and parenthood for every nationality.

Finally, the model also takes into account the differences in the probabilities of getting married according to the previous marital status. At low ages, single, divorced and widowed have a similar behaviour, but this similarities start slowly to diminish and divorced and widowed individuals tend to get marriage again with lower chances than single individuals. This fact takes place especially between women in countries of eastern and southern European countries.

2.3.4. Labour market

The labour market is composed of two different parts: a real labour supply integrated by individuals who decide whether to work in the market or be inactive, and an artificial labour demand, computed by adapting a set of standard macroeconomic rules.

Additional to the hiring process, a number of workers are fired every period, becoming unemployed. This process is also determined aggregately through the expected macroeconomic outcome.

This module is composed of the following methods: labour supply update, labour suitability update and labour market.

In the first method every individual who is currently working decides if they are going to retire. The main difference in this behaviour comes from the gender difference. In most of the regions women tend to leave their jobs at a lower age than their male colleagues.

This different behaviour is also considered when unemployed people cease to look for an employment. This behaviour is especially remarkable when the woman is married and has children. Unemployed individuals, independently of their characteristics, get inactive more often when the unemployment rate is high. In the same way, inactive individuals start looking for a job again in the expansionary phase of the business cycle when the unemployment rate is low.

In the second method, after updating of labour demand, every unemployed individual updates the value of the variable LABSUIT. It tries to faithfully represent the attractiveness of every individual for the labour demanders. A higher value increases the chances to be hired and decreases the probability of losing current employment if the individual is employed. Nevertheless, in the third method, when the labour suppliers and the labour demand meet every individual has positive probabilities of been hired and fired, as it happens in the reality.

The variable labsuit has been calibrated with microdata of the Labour Force Surveys. It uses the values of the following variables of the individual:

- Age: The labour suitability increases from the minimum legal working age to thirty/thirty five years by approximately twenty percent. Later it decreases at a higher rate and a worker becomes completely unable to find a job in the range sixty to sixty-five years old.
- Gender: Males are hired with a higher probability. Male unemployment rate is lower and the labour suitability takes into consideration this empirical finding.
- Nationality: Foreign workers usually have more difficulties to be hired.
- Labour periods: Long-term unemployed population can hardly find a new job in

most of the European labour markets, arriving to the so-called hysteresis. On the other hand long-term employees have fewer chances to be fired than the temporary workers, recently incorporated to the activity.

The variable labsuit is defined for employees and unemployed individuals, and it adopts zero value for students and inactive agents, making them unable to find a job. Its specification is different depending on the labour status. For unemployed individuals it includes gender and nationality but these two variables are excluded in the case of current workers, who are assumed to not be discriminated for their personal characteristics.

Table 2.2. Labsuit parameterization for ES617 (Málaga)

Coefficient	Estimated value (2001-2012)
α_{67}	63.842
α_{68}	-0.137
α_{69}	-0.086
α_{70}	-0.116
α_{71}	-0.028
α_{72}	0.074

Source: Own elaboration

Table 2.2 results show that in this NUTS3 region the unemployed foreigners (α_{68}) have 13.7% less chances to be hired than a native individual with the rest of the characteristics equivalent. Women (α_{69}) have 8.6% less chances to get employed than men. These effects can be compounded: A foreign woman would have 23.4% less chances to be hired than a native man with the same age and periods unemployed. A higher education level, acting as a proxy of human capital stock, has also a positive impact on the chances of getting employment. Finally, the coefficient α_{70} in each equation show that each year unemployed decreases the chances of getting a job by 11.6% while each year being employed in the same company (α_{72}) decreases the chances to lose the job by 7.4%.

Finally, the labour demands at regional level are computed using a macroeconomic forecast (FC) pool and adapting it to the past regional performance, overweighing the recent years.

$$FC_{reg} = FC_{country}\left(\left(\sum_{1990}^{2012}\frac{\Delta GDP_{reg,year}}{\Delta GDP_{country,year}}(year-1989)\right)\left(\sum_{1990}^{2012}year-1989\right)\right)^{-1}$$

Table 2.3. Comparison of forecast result with and without weighted past regional performance

A	A1	A2	A1/A	A2/A	W(A1/A)	W(A2/A)
1.0	0.6	1.4	0.6	1.4	0.6	1.4
1.5	1.4	1.7	0.93	1.13	1.87	2.27
2.0	2.1	1.8	1.05	0.9	3.15	2.7
1.8	1.9	1.5	1.06	0.83	4.22	3.33
2.3	2.2	0.8	0.96	0.35	4.78	1.74
Average Ai/A for the period			0.92	0.92	0.97	0.76

Source: Own elaboration

In this example, the importance of the weight of historic regional performance with respect to the national aggregate GDP increment (A) is assessed. The two presented regions (A1 and A2) grew at a rate of 92% of the national level. However, the first one increased its average while the second region results turned down. Hence, the weighted past performance (w(Ai/A)) makes use of this trend and predicts a grow forecast of the region 1 at 27.8% higher than the one of the Region 2.

With the regional forecast the change in the rate of unemployment is computed according to the well-known Okun law (1962), computed at regional level. This overall change in unemployment is divided into the regional figures for the number of employees to be fired and the number of unemployed individuals to be hired. The relation between these two quantities is considered fixed at regional level and is computed with the microdata of the Labour Force Survey.

Finally, in the artificial labour market mechanism, workers are selected to be hired and fired taking into account the relative value of their labour suitability and the number of vacancies and workers to be fired. It is assumed that the enterprises are forced to fire their less productive employees, which are those that have a workforce with lowest values of labsuit. In addition, labsuit includes the number of periods the employee worked in the enterprise. Hence, recently hired workers are more likely to be fired. The workforce of new firms covers this kind of employees and the model includes the well-known empirical finding of new enterprises have the lowest survival probability (Audretsch & Mahmood, 1995). Nevertheless, the matching process is randomized at individual level, but preserves the preference for individuals with a higher labsuit aggregately.

In the example in the table 2.3, it is assessed the importance of weight the historic regional performance with respect to the national aggregate GDP increment (A). The two presented regions (A1 and A2) grew at a 92% of the national level. However, the first one increased its average while the second region results turned down. Hence, the weighted past performance (w(Ai/A)) makes use of this trend and predicts a grow forecast of the region 1 a 27.8% higher than the one of the Region 2.

With the regional forecast it is computed the change in the rate of unemployment according to the well-known Okun law, computed at regional level. This overall change in unemployment is divided into the regional figures for the number of employees to be fired and the number of unemployed individuals to be hired. The relation between these two quantities is considered fixed at regional level and is computed with the microdata of the Labour Force Survey.

Finally, in the artificial labour market mechanism, there are selected the workers to be hired and fired taking into account the relative value of their labour suitability and the number of vacancies and workers to be fired. It is assumed that the enterprises forced to fire their employees are the less productive ones, those that have a workforce with lowest values of labsuit. In addition, labsuit includes the number of periods the employee worked in the enterprise. Hence, recently hired workers are more likely to be fired. The workforce of new firms is included in this kind of employees and the model includes the well-known empirical finding of new enterprises have the lowest surveillance probability. Nevertheless, the matching process is randomized at individual level, but preserves the preference for individuals with a higher labsuit aggregately.

2.3.5. Migration

The migration module has four methods: in the first one, adult individuals with housing stress decide whether they want to migrate or not. In the second method any adult individual might decide to migrate in order to increase income. Then they decide the destination region. The last method in the model takes account of the immigration phenomena.

Long-term unemployed individuals have the highest probability of emigrating, especially if they are young, single, they do not have children and they live in regions with a declining populations and high unemployment rates. Other kinds of migration occur in regions with low average income. In this case some individuals want to change their location in order to increase their expected income. Individuals who live in areas with declining populations, an unstable work position, low age and no children are the agents who undertake this kind of migration more often.

Individuals who decide to migrate choose the destination location among the European NUTS-3 regions. They compute using perfect information and full rationality the expected utility of being in every one of them, according to unemployment rate and population growth. Young individuals give more importance to the average income while elder ones prefer to migrate to a near region in spatial and cultural terms. This utility is multiplied by the population of the region to reach the probability of migration. The individual selects a region randomly, but in line with this computed probability.

Finally, a number of immigrants (non-EU28 individuals) are computed for each region. This number depends on demographics indicators: population, immigrant population already settled and population growth, and economic performance: average income and unemployment rate.

2.4. Calibration of the system

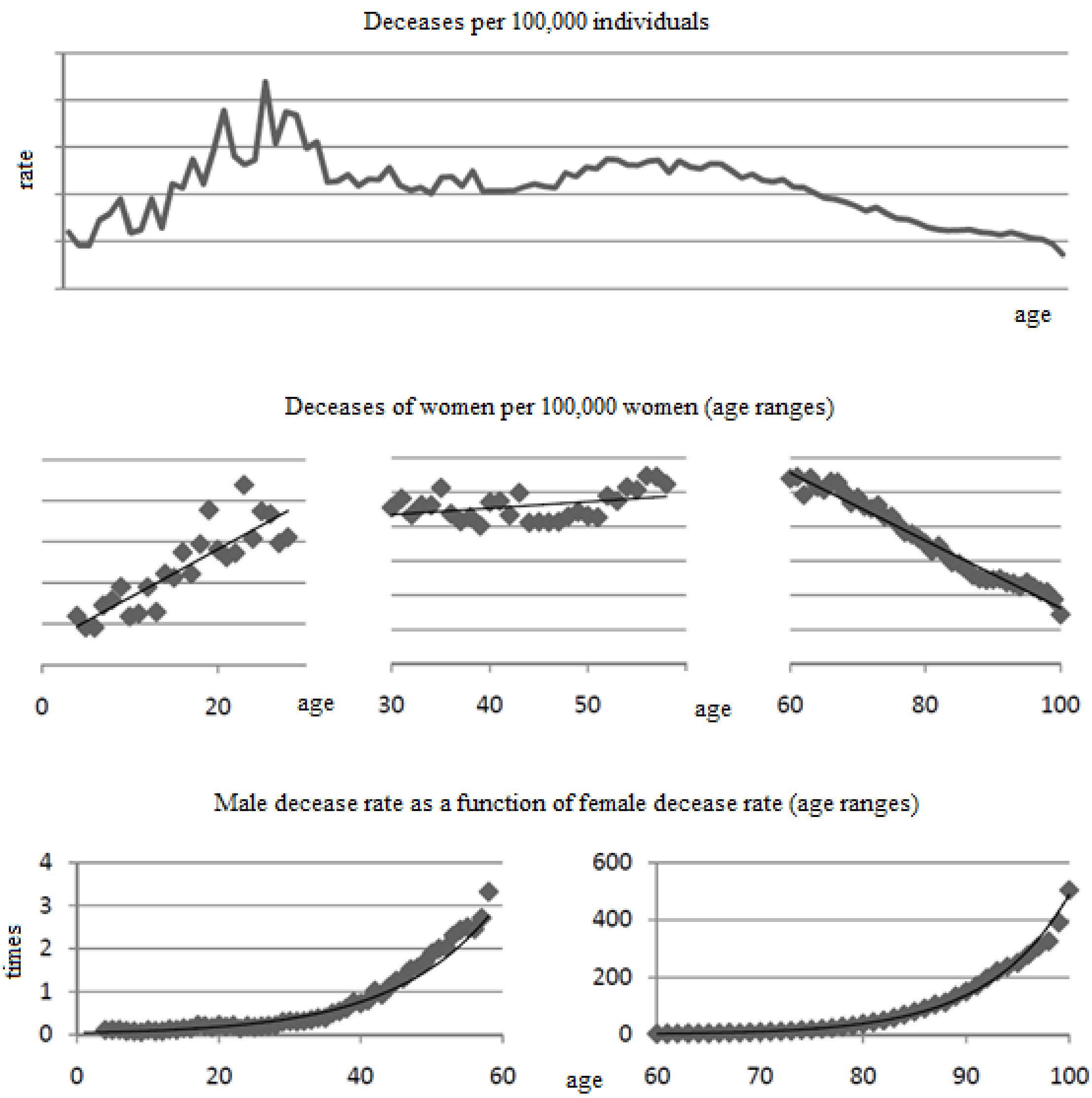
A set of data sources are used to calibrate every coefficient of the system:

- The EU Labour force survey is a household sample survey that provides quarterly results on labour market variables of people aged 15 and over. In 2011 the quarterly labour force survey sample size across the EU was about 1.5 million individuals. This survey has available samples of microdata that make it possible to compute many of the coefficients of the presented model at regional level.
- Education statistics differ across the European countries, but every national Statistics Office includes the number of individuals studying and/or graduated in every education level. These figures are disaggregated at least by gender and age.
- Migration statistics can be obtained through several ways. The most convenient ones to the needs of the MAS are specific migration surveys focusing on the determinants and characteristics of the individuals who changed their location; and municipal registers that yearly compute the population classified by its characteristics (most of the time: age range, gender, marital status and nationality).
- Birth and decease statistics include information about the age and gender of the deceased, and age and marital status and nationality of the newborn's mother.
- Wedding and divorces statistics collect information about the age of the partners. In several countries there are double entry tables combining the ages of the partners.

In order to show how the calibration process takes place it is used the case of the coefficients estimated with the mortality statistics for the region of Grande Lisboa, Portugal (PT171). I start depicting the relation of the male mortality rate into female mortality rate (Fig. 2.4 bottom). The usual behaviour follows three patterns (Figure 2.4 middle). Up to thirty years the mortality rate of the males increases with respect to the one for female individuals. From thirty to approximately sixty years its remains steady and after that age it starts declining. Three linear estimations were carried out:

1) ratio = 0.1171 age + 0.4727 (R^2 = 0.7614)
2) ratio = 2.311 (series average)
3) ratio = -0.0492 age + 5.7332 (R^2 = 0.9704)

Figure 2.4. Calibration of the coefficients of aging and death for PT171



Source: Deceases statistics, population statistics and own elaboration

Once the differential effect of gender has been taken into account, the effect of the age (using women's morbidity rate as the base one) is computed dividing the age range in the age equal to sixty to improve the parameterization adjustment.

1) rate = 0.0433exp(0.0716 age) R^2 = 0.955
2) rate = 0.0015exp(0.1272 age) R^2 = 0.9952

Then, it is possible to include the following coefficients for the aforementioned regions:

1) $\alpha1$=3.291, $\alpha2$ = 3.748, $\alpha3$ =0.0433, $\alpha4$= 0.0716
2) $\alpha5$ =0.0015, $\alpha6$= 0.1272, $\alpha7$ =0.1171, $\alpha8$= 0.4727
3) $\alpha9$= 2.311, $\alpha10$= -0.0492 and $\alpha11$= 5.7332

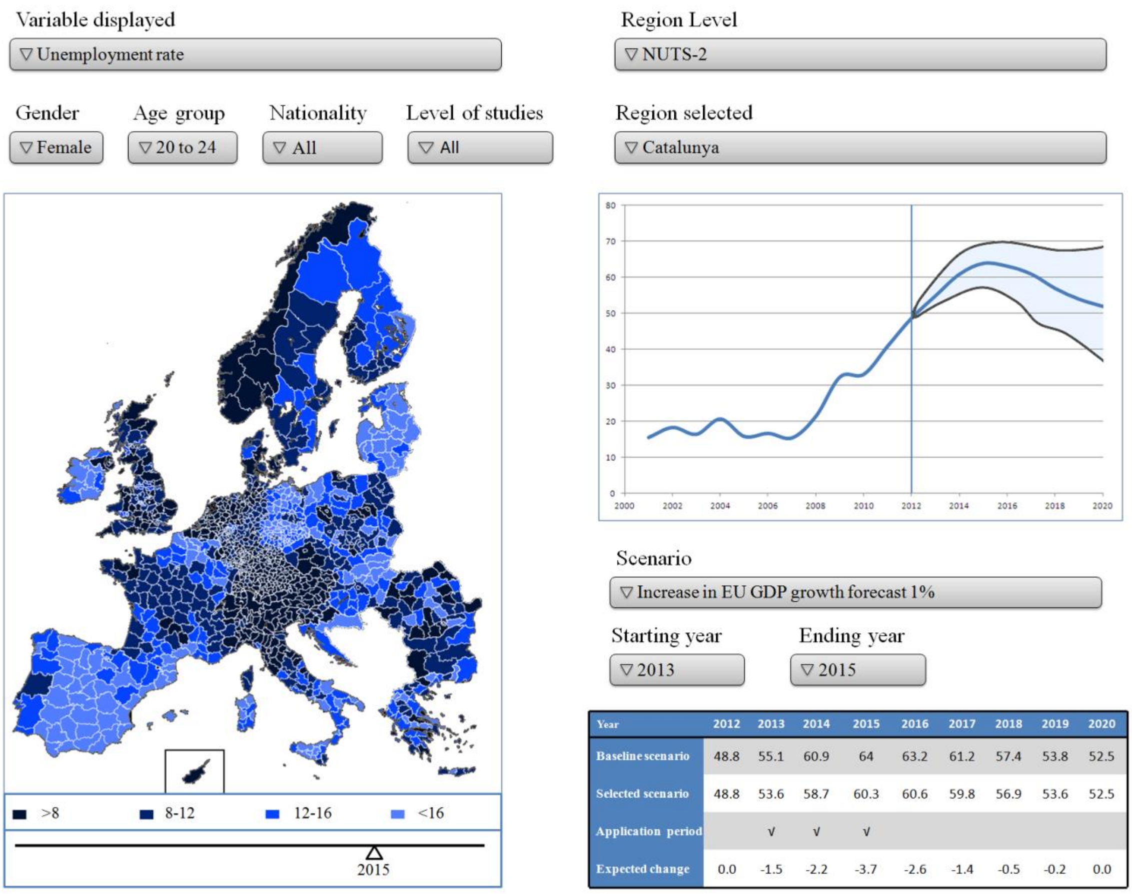
2.5. Visualization and scenarios

The regional variables are computed after each of the iterations. These variables are just the result of agent additions obtained with the proper queries, and in some cases it is also necessary to calculate percentages (unemployment rate) or variation rate (population growth). These endogenous variables are used for visualization purposes. The user selects which ones he wants to display (See figure 2.5):

Variable displayed: population, population growth or unemployment rate.

- Gender: Female, male, both.
- Age group: five years ranges in the interval 15 to 65, 1-15, more than 65, all.
- Nationality: Native (of each country), EU28, noEU28, all. • Level of studies: one to eight, 4-5, 7-8, all.
- Region level: Nuts 1, 2 or 3. Depending on the selected level the possibilities of the following characteristic are different.
- Region selected: In addition to the map presentation the interface has also a graph (time series / population pyramid) of the selected variable.
- Scenario: None, decrease/increase the EU GDP growth forecast by 0.5, 1 or 2 percent.
- Starting & ending year: If any scenario is activated the user can select the starting and ending year and see how the scenario affects to the overall result.

Figure 2.5. User interface of the system



Source: MAS interface, own elaboration

2.6. Final discussion

This chapter presents and discusses a MAS of population dynamics for the European regions at NUTS-3 level. The technique of agent-based modelling has been chosen to allow including the concept of bounded rationality to the model and to combine demographic and economic data. Bounded rationality is implemented through stochastic decisions at individual level that replicate the actual behaviour of the population.

The chapter describes the latest novelties in this area of research, the process of database creation as well as the model development and calibration. The database has been developed on the basis of the population census, and the input data is used then to obtain a model that helps to understand the dynamics of populations and the relations between demography and economics. Each agent in the model passes sequentially through the five implemented modules: ageing and deceases, education, pairing, labour market and migration.

In the first module every agent computes a probability of death and it is determined who deceases. All members of the population increase their age each year. The model also takes into account the possibility of getting widowed or orphaned.

In the second module individuals initiate their education, continue it or finish depending on the levels of studies previously received, age, gender and labour status.

In the third module the pairing process is analysed. Adult agents of one gender decide whether to get married, divorce or have children. In this way the MAS represents more accurately the population dynamics in the long run.

In the following module, labour market characteristics are presented. Individuals have the choice of turning into inactive or starting look for a job. The hiring and firing processes are also defined as well as the labour market demand in case of each of the regions. The module is composed of the three methods: labour supply update, labour suitability update and labour market. The goal of this module has been to present the determinants of labour market movements but without doubts there exist others; to test the methodology of research about the labour market shifts; and their further consequences for the population.

The first method describes the process of getting retired and the results that change depending on the gender differences. This method is assumed to model the differences in individuals' behaviour and to preserve the heterogeneity in the behaviours defined in the system. The second one is used to update the labour demand. The updated evaluation of individual's power in the market has been included to represent faithfully the attractiveness of every individual for the labour demanders. As in widely used models, a higher value increases the chances to be hired and decreases the probability of losing employment. In the third method, the complete labour market is defined by the inclusion of demand and supply sides.

In the last module, the agents compute their satisfaction with their labour status and the probability of housing stress is taken into account as well. Also the decisions of staying or changing the place of living in the same region or for other one is included. At the same time, if individuals decide to change the region, the family also opts for changing their location. There are two main determinants, the persistence of unemployment and the income variation across European regions. Individual decide their destination region according its characteristics and the spatial and cultural distance with their initial location. Immigration it is also taken into account.

For the purpose of model preparation, the one year iteration period was assumed. Parting from the data obtained from the reliable sources, the recalibrated model, jointly with the simulation, has been used to obtain the visualization for Spanish regions.

The novelty of this agent based model is the use of this technique in the studies of complex population dynamics phenomena. The applicability of this technique requires adequate objects definition. In this chapter, the following objects were defined: individuals as the agents that carry out the actions and populate the system, and families that are composed of individuals and have a possibility of migration.

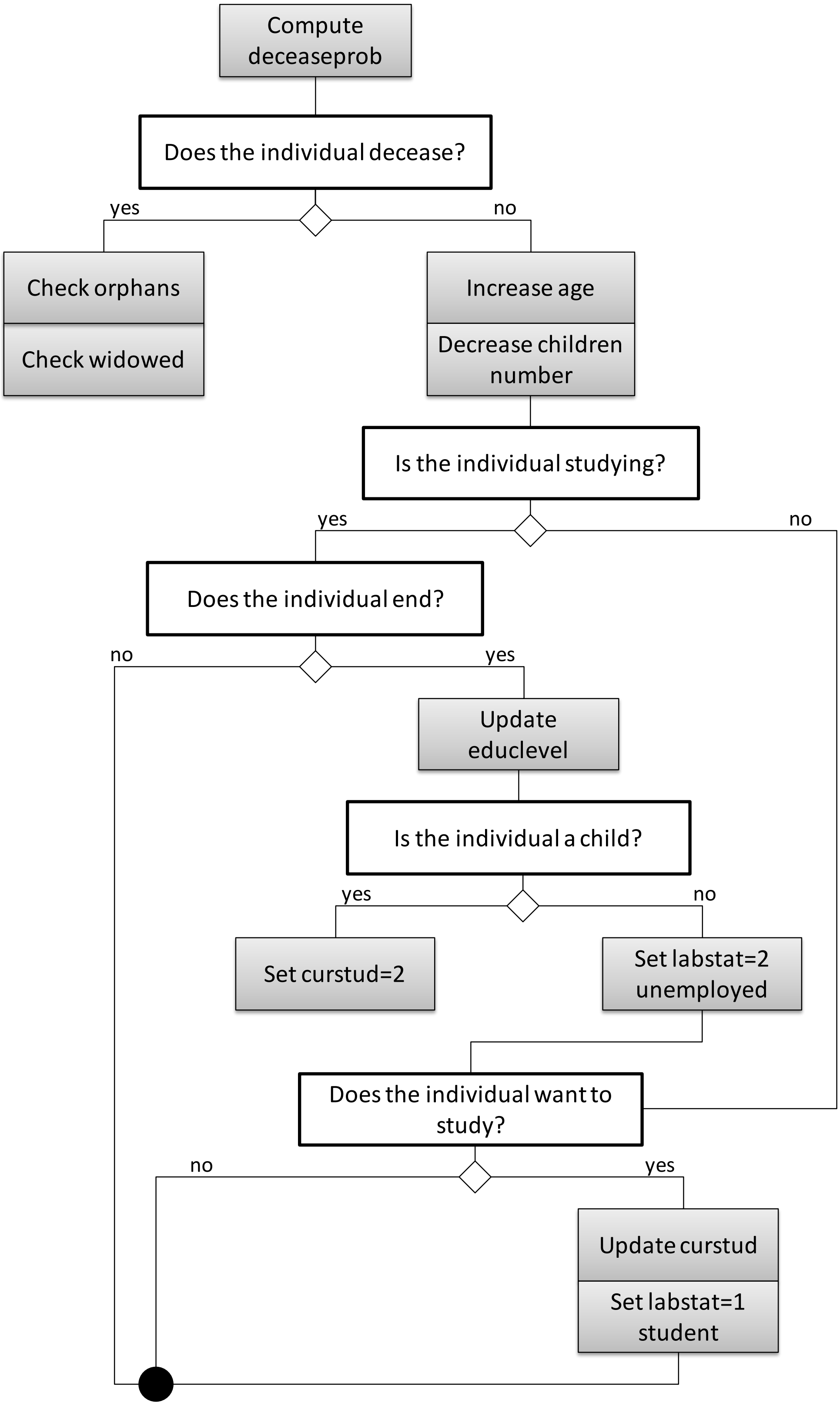
Notwithstanding, not only are objects individuals and families but also specific ones such as regions or space. The regional perspective assumed in the model allows the individuals to be located in the particular territory and then to analyse all consequences of changes in location or position. All regions include variables of aggregate outcomes and macroeconomic forecast. The other important object predefined in this version of population model is space. This object is represented by the matrix that includes the spatial distances between the centroid of every NUTS-3 region and cultural distances.

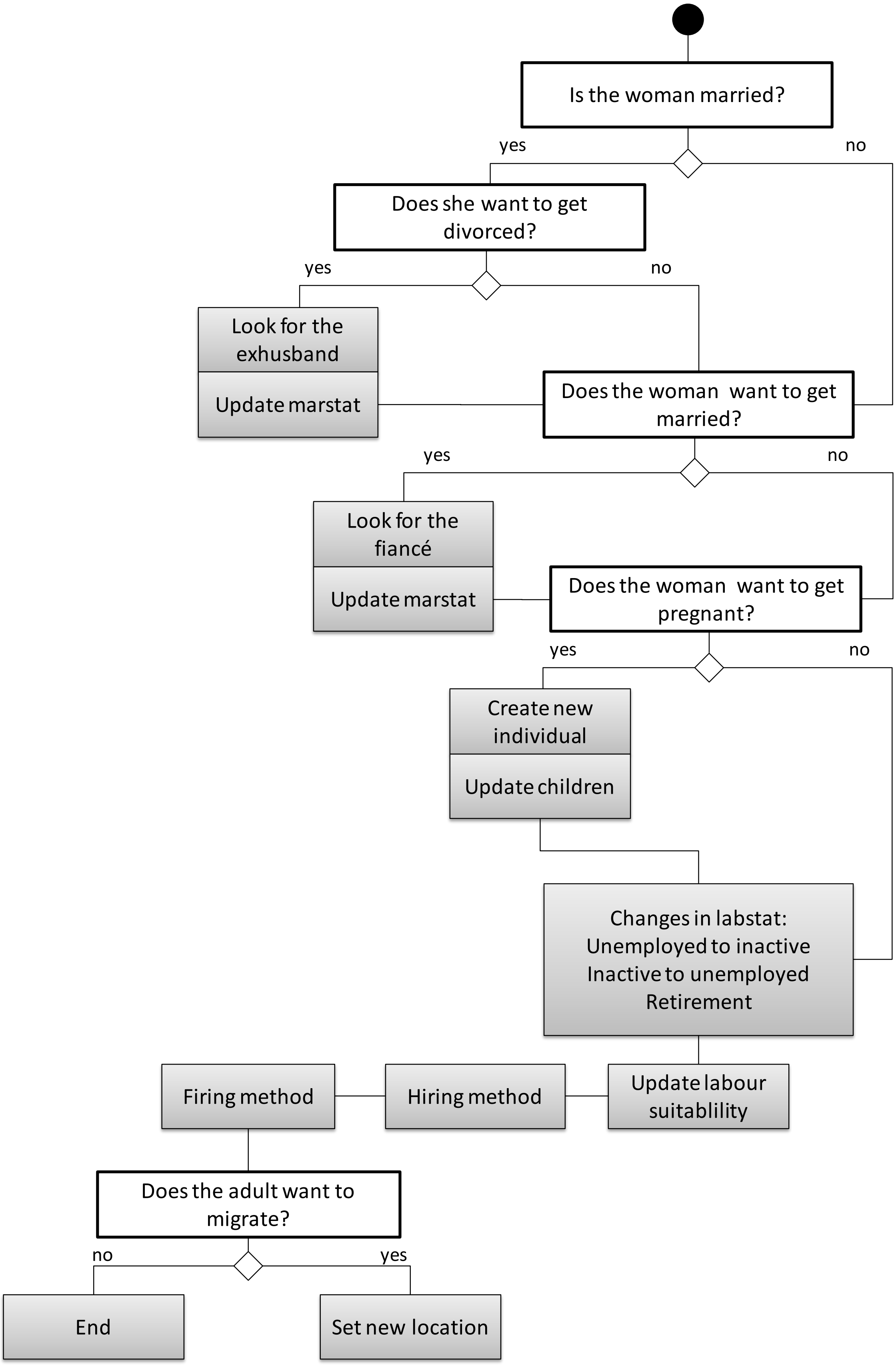
One of the most important features of this MAS is the preservation of the heterogeneity of agents in the model. The heterogeneity was taken into account in the model by including the list of variables that specify the characteristics of the agents. Rising awareness of the benefits of the agent based modelling for population dynamics. Not only is the heterogeneity preserved but also the interactions and changes in individual's behaviours depending on regional location are analysed.

Population dynamics is not only about demography and spatial location. Equally important to carry out the full analysis of population dynamics is the ability to consider the impact of public policies and the position in the labour market of individuals. Those factors have a direct impact on individuals' and families' choices about their futures. The definition of the labour market in the model presented hereby was elaborated assuming that it is composed by two different parts: a real labour supply integrated by individuals who decide if they work for the market or whether they prefer to be inactive, as well as the second one - an artificial labour demand, computed adapting a set of standard macroeconomic rules. Its indirect aim is to contribute to the artificial labour market research mentioned in the introduction as well as to the general research on population dynamics modelling using the agent-based modelling.

This model makes possible to simulate and forecast the population dynamics of European regions in the period 2001-2021 and it explores how agent-based modelling can contribute to integrate demographics and economics to forecast accurately migration patterns as a result of labour and income differences between socioeconomic integrated areas as the European Union.

2.7. Appendix: Activity diagrams and algorithms of the model




Is the woman married?
yes
no
Does she want to get divorced?
yes
no
Look for the exhusband
Update marstat
Does the woman want to get married?
yes
no
Look for the fiancé
Update marstat
Does the woman want to get pregnant?
yes
no
Create new individual
Update children
Changes in labstat:
Unemployed to inactive
Inactive to unemployed
Retirement
Firing method
Hiring method
Update labour suitablility
Does the adult want to migrate?
no
yes
End
Set new location

```
AGEING_1
Double deceaseprob
If age<1 and gender=0 Then deceaseprob= α1
Else If age<1 and gender=1 Then deceaseprob= α2
Else If age<60 and gender=0 Then deceaseprob= α3 exp(α4age)
Else If age>=60 and gender=0 Then deceaseprob= α5 exp(α6age)
Else If age<30 and gender=1 Then deceaseprob= (α7+ α8age) α3 exp(α4age)
Else If 30=<age<60 and gender=1 Then deceaseprob= α9 α3 exp(α4age)
Else If age>=60 and gender=1 Then deceaseprob= ( α10+ α11age) α5 exp(α6age)
End If
```

```
AGEING_2
If deceaseprob<1000*rndm
If age<16
extract a number from N(α12, α13) = agemother
extract a number from N(α14, α15) = agefather
get a woman with age= agemother and chidren>0 and woman.children - -
get a man with age= agefather and chidren>0 and father.children - -
Else If children>0
get an individual with same region, gender and higher age
set ind.children +=children of the deceased
Else if marstat=2
get an individual with same region, different gender and ind.age-age=<|5|
set ind.marstat=3
End If
Individdual --
Else
Age++ and Labperiods ++
End If
```

```
EDUC_1
If curstud>0 and (educ_level=1 and periods_edu=24) | (educ_level=2 and periods_edu=40) |
    (educ_level=3 and curstud=4 and periods_edu=44) | (educ_ level=3 and curstud=6 and
    periods_edu=48) | (educ_level=4 and periods_edu=48) | (educ_level=5 or 6 and
    periods_edu=60) | (educ_level=7 and periods_edu=80)
Then educ_level=curstud, Set curstud=0, labstat=2 and Labperiods=0
Else If curstud>0 Then periods_edu ++ Go to PAIRING_1
End If
```

```
EDUC_2
If educ_level=1 and 5<age<16
Then curstud=2
If educ_level=2 and (age<16 | (unempl=1 and rndm< α16))
Then curstud=3 and labstat =1
```

```
Else if gender=1 and educ_level=3 and ((age<18 and labstat=3 and rndm<α17) | (unempl=1 and
    rndm< α18))
Then curstud=4 and labstat =1
Else if gender=1 and educ_level=3 and ((age<18 and labstat=3 and rndm<α19) | (unempl=1 and
    rndm< α20))
Then curstud=6 and labstat =1
Else if gender=0 and educ_level=3 and ((age<18 and labstat=3 and rndm<α21) | (unempl=1 and
    rndm< α22))
Then curstud=4 and labstat =1
Else if gender=0 and educ_level=3 and ((age<18 and labstat=3 and rndm<α23) | (unempl=1 and
    rndm< α24))
Then curstud=6 and labstat =1
Else if educ_level=4 and ((age<19 and labstat=3 and rndm< α25)
| (unempl=1 and rndm< α26))
Then curstud=5 and labstat =1
Else if educ_level=5 and ((age<20 and labstat=3 and rndm< α27)
| (unempl=1 and rndm< α28))
Then curstud=7 and labstat =1
Else if educ_level=6 and ((age<20 and labstat=3 and rndm< α29)
| (unempl=1 and rndm< α30))
Then curstud=7 and labstat =1
Else if educ_level=7 and age<27 and labstat=3 and rndm< α31
Then curstud=8 and labstat =1
End If


PAIRING_1 //divorce
If gender=1
Go to LABOUR_1
Else If marstat=2 and rndm<∫(age+0.5, age-0.5) N(α32, α33) α34(nac)
Then Marstat=4
Look for an individual with marstat=2, gender=1 and age-wife.age<|5|
Set marstat=4

PARING_2 //marriage
If (marstat=1 and rndm<∫(age+0.5, age-0.5) N(α35, α36) α37(nac)) |(marstat=1 and
    rndm<∫(age+0.5, age-0.5) N(α38, α39) α40(nac)) | (marstat=1 and rndm<∫(age+0.5, age-0.5)
    N(α41, α42) α43(nac))
Then Marstat=2
Look for an individual with marstat!=2, gender=1 and age-wife.age<|5|
Set marstat=2

PARING_3 //parenthood
If marstat!=2 and rndm<∫(age+0.5, age-0.5) N(α44, α45) α46(nac)
Then Children ++
```

Create new individual
If marstat=2 and rndm<∫(age+0.5, age-0.5) N(α47, α48) α49(nac)
Then Children ++
Create new individual
Look for an individual with marstat=2, gender=1 and age-mother.age<|5|
Set children ++
LABOUR_1 //Labour supply update
Set Labstat=4 and Labperiods=0 If any of the following it is true
If labstat=3 and gender=0 and age>55 and rndm<∫(age+0.5, age-0.5) N(α50, α51)
Else if labstat=3 and gender=1 and age>55 and rndm<∫(age+0.5, age-0.5) N(α52, α53)
Else if labstat=2 and labperiods>2 and gender=0 and married!=2 and rnmd<∫(age+0.5, age-0.5) N(α54, α55) (α56+ α57urate)
Else if labstat=2 and labperiods>2 and gender=0 and married=2 and children=0 and rndm<∫ (age+0.5, age-0.5) N(α58, α59) (α60+ α61urate)
Else if labstat=2 and labperiods>2 and gender=0 and married=2 and children>0 and rndm<∫ (age+0.5, age-0.5) N(α58, α59) (α60+ α61urate) α62children
If labstat=2 and labperiods>2 and gender=1 and rndm<∫(age+0.5, age-0.5) N(α63, α64) x(α65+ α66urate)

LABOUR_2 // Update labour suitability
For every individual
If the nationality corresponds to the country that the region is included in Foreigner=0
Else Foreigner = 1
End If; End for
If labstat=2 LABSUIT=Max (((α67age) - (age2)) (1+ α68foreigner) (1+ α69gender) (1+ α70labperiods) (1+ α71educlevel), 0)
If labstat=3 LABSUIT=Max (((α67age) - (age2)) (1+ α72labperiods), 0)
Else LABSUIT=0

LABOUR_3 // Labour market
Get hiredworkers=N
Order all individuals with labstat=2 decreasingly by labsuit
While Hiredworkers>0
If (N/number of unemployed)x(ind.labsuit/Max.labsuit)<rndm
Hiredworkers-- and Ind.labstat=3 and Ind.labperiods=0
Get firedworkers=N
Order all individuals with labstat=3 decreasingly by labsuit
While Firedworkers>0
If (N/number of employed)x(Min.labsuit/ind.labsuit)<rndm
Firedworkers-- and Ind.labstat=2 and Ind.labperiods=0

MIGRATION_1
If labstat=2 and labperiods>2 and marstat=2
Search partner !=gender and |age-partner.age|<5

```
If partner.labstat=2 and partner=labperiods>2 rndm<(α73 +α74urate)2(α75+ α76popgrowth) (α77children)-1 (α78 +α79age)-1Go to MIGRATION_3
Else Go to MIGRATION_2
End If
Else If labstat=2 and labperiods>2 and marstat!=2
If rndm<(α80 +α81urate)2(α82 +α83popgrowth) (α84children)-1(α85 +α86age)-1
Go to MIGRATION_3
Else Go to MIGRATION_2
End If
End If
MIGRATION_2
If labstat!=2 and rndm>(α87 +α88popgrowth) (α89 +α90age)-1(α91 +α92
labperiods)-1(α93children)-1 Go to MIGRATION_3

MIGRATION_3
For each individual
For each region
Probreg=pop (α94 popgrowth) (α95 +α96urate)-1+(α97 +α98age)-1
AVERAGE_INCOME+ (α99 +α100age) (α101 cultural_dist) +(1-α102) spatial_dist)
Order them randomly
If probreg<rndm change location to this region
Else continue
End For; End For

MIGRATION_4
For each region
Nimmi= (α103pop+α104popnoEU28+ α105popgrowth)( α106AVERAGE_INCOME +(1-α107)urate-1)
While Nimmi>0
Get an individual of the region with nacionality=noEU28 and create a new record with the same characteristics except immi.age=N(α108, α109)age
Nimmi --
If immi.marstat=2
Get an individual of the region with nacionality=noEU28, gender!=immi.gender and create a new record with the same characteristics except immi.age=N(α108, α109)age
Nimmi --
End If
End For
```

CHAPTER 3

The effect of firm growth, innovation and firm dynamics on market concentration: an evolutionary simulation model

3.1. Introduction

Firm growth and firm dynamics play a key role in economic analysis (Clementi & Palazzo, 2013). Companies are heterogeneous and the equilibrium in competitive markets is different from the one of perfect competition if there is market power (Tirole, 1988). Furthermore, the dynamic process of creation and closures of firms in the market creates a process of natural selection that improves the sector average efficiency period after period. New firms introduce process and product innovations copied from other sectors or locations, but also genuine innovation (Winter, 1971). Competitive market forces will tend to select those companies that have incorporated these innovations, no matter whether they are new firms or established ones (Lazonick, 2005).[20]

Firm growth and firm dynamics have a major effect on characteristics of the markets because the most successful companies tend to grow at a higher rate than the less successful ones, which have a higher probability to cease their activity. Usual approaches explain the inability of firm dynamics to increase the level of concentration in most of the markets from the existence of different types of entry barriers (Schivardi & Viviano, 2011). In most of the cases new firms would only make a significant decrease in the market concentration, especially in early life stages of an industry, when the chances of product innovation outweigh the disadvantages of being a small size.

The literature on these issues shows that firm dynamics have a positive effect on productivity and competitiveness, although the structure of the market does not appear to be substantially modified. This improvement in productivity is apparently based in innovation provided by new firms (Caggese & Cuñat, 2013). However, the improvement in competitiveness might not only be due to the existence of new innovative companies, but also by the adoption of innovations in these existing firms.

The main objectives of this chapter can be summarized as follows:

- To characterize market concentration in Spain before and after the crisis.
- To determine the realistic behaviours of firm growth and firm dynamics

[20] Innovation can be defined in this context as a change in product or service characteristics that increases the survival of enterprises by promoting a greater adaptation to their environment.

according to the observed data.

- To forecast market concentration under different scenarios with an agent-based simulation that includes firm growth, firm dynamics and innovation of companies.

In the next section we discuss the drawbacks of the game theory approach to this issue and the advantages of simulation. In the third part it is presented the results of market concentration in terms of business concentration. The simulation model is introduced in the fourth section, fifth section is devoted to market concentration forecasting, and finally sixth section show the main results under different scenarios.

3.2. The limitations of game theory for the analysis of firm dynamics and the needing of simulations

Industrial organization works using game theory as the main tool to analyze market dynamics (Belleflamme & Peitz, 2015). Although the results obtained with these techniques have been very successful in a number of scenarios, game theory needs to accept a wide number of restrictive assumptions that can limit the validity of the conclusions obtained (Aiginger & Finsinger, 2013).

Traditionally, research about competition and decisions of companies related to pricing and other factors that affect demand assumed homogeneous companies; it is the very known case of Hotelling model, where the firms only change their price and position (Hotelling, 1931). In the last years the focus has been on aspects such as the interactions with multiple demands (Gaudet & Salant, 2015) or the interrelation between distance, cost and transport scheduling (van der Weijde, Verhoef, & van den Berg, 2014). However, still little attention is paid to the effect of firm growth and the appearance of new companies with a potential different decision-making process.

In addition, static models have a fixed number of companies. This means in practice the omission of markets where there are not infinite entry barriers for new firms, and the closure of the existing ones. This assumption is one of the strongest as firm dynamics is a process that also affects market concentration as new small businesses also have a key role in the market (Ackermann, 2012).

Another important drawback of game theory frameworks is that they are often deterministic. Therefore they do not include stochastic features that have a crucial importance in the process of evolution of firm characteristics. There are sources of randomness that affect the actual demand of companies such as consumer awareness of the company as well as decisions of managers related to location, pricing, firm size and trading hours.

Finally, traditional models have problems in addressing medium term issues, since they consider only a few periods or involve an infinite number of periods in order to find the equilibrium. This makes them inadequate to cope with many problems of companies and governments.

Computer simulation techniques are a preeminent alternative to traditional methods because they are able to tackle the aforementioned problems successfully by a simple and direct approach. In this chapter it is applied an agent-based simulation which demonstrates the usefulness of simulation methods for addressing complex theoretical problems in the field of industrial organization such as analyzing the implications of business growth on market concentration (Pablo-Martí, 1999).

Firm growth has been the subject of an intense debate within the industrial organization. Explanations of business growth come in two types: first, the deterministic ones based on the neoclassical model and the stochastic ones. The deterministic or technological explanation of business growth is closely linked to the idea of optimal size. Firm growth would be the process by which companies try to achieve it.

The explicit introduction of time in a firm growth model is due to Solow (1971). It was a valuable effort to provide a dynamic dimension to the neoclassical theory of business growth by introducing the assumption that the firm faces a problem of inter-temporal maximization of its net present value. The main result from the analysis is the company wants to reach its optimum size as quickly as possible, but the existence of adjustment costs makes this process occur in several periods. From this point of view the diversity of sizes observed in the market is a temporary situation because the companies are in different stages of the adjustment process towards the optimal size.

On the contrary, stochastic explanations are based on the observation of the significant asymmetry in size distribution. Asymmetry does not correspond to the forecast of the determinant theory of firm growth but it is coincident with various theoretical distributions such as lognormal, Yule or Pareto. These theories give less importance to the technological aspects of demand and consider that the evolution of firm size is affected by many explanatory variables which should be treated as random perturbations. Within this group of models the Law of Gibrat (1931) can be highlighted. It is the strong version of the well-known Law of proportional effect that states that business growth can be taken as a purely random phenomenon.

The microfoundations of enterprise performance might exceed the capability of game theoretic frameworks. For instance, “an enterprise might even be good at invention, but it will likely fail to capitalize on its technological accomplishments” (Teece, 2007) and therefore a model that includes interrelations in endogenously and allows random results in a natural way might be highly desirable. While game theory is one of the best approaches to discover optimal behaviours, agent-based simulations might be the most promising option to include an approximation of the actual behaviour.

The goal of the MAS is to predict the level of market concentration under different scenarios. According to the aforementioned disadvantages of other techniques, building a simulation model is presented as the best alternative. In fact, prediction is the first purpose of simulations (Axelrod, 2007). Moreover, agent-based social simulations allow both study combined and complex effects and proximate mechanisms behind the researched phenomenon (Castelfranchi, 2014).

There are an increasing number of applications of agent-based models developed to study market behaviour as the aggregation of individual different behaviour of heterogeneous companies. Zott (2003) proposes a simulation model to search on the interrelation between the dynamic capabilities of companies and their performance within a sector. His model shows how endogenous firm decisions based on imitation and proactive actions can affect firm performance to a great extent. Therefore, the evolution of firms has a strong impact on intra-industry differential performance. This result is only possible under a simulation framework.

In addition to testing conceptual theories, simulation methods can be used for theory development. "Simulation strengths include internal validity and facility with longitudinal, nonlinear, and process phenomena" (Davis et al., 2007). In section four of this chapter the presented simulation model is trained with actual data and therefore it can contribute to developing the theory of the main determinants on firm dynamics. Other papers also aim to contribute to this unresolved issue in industrial organization mixing mathematical and computational models to study the trade-off between flexibility and efficiency (Davis et al., 2009). The previously developed approaches lack the business characteristics included in the model and do not allow companies to evolve them to an extent comparable to the actual ones observed in terms of product differentiation.

In the next section it is presented the measure to determine market power and its performance in the country that was used to train the model. This model is presented in the fourth section and predicts the effects of economic growth on market power on selected industrial sectors.

3.3. Market power in the Spanish manufacturing sector

The Herfindahl index, also known as Herfindahl–Hirschman Index, (Hirschman, 1945; Herfindahl, 1950) is a measure of competition in a market. It is defined as the sum of the squares of the market shares of the N largest firms within the industry. It can range from 1/N to 1. A market with a high number of very small firms would have a Herfindahl index close to zero and market with a monopolistic producer will have a value of one. A high Herfindahl index generally indicates low competition and a high level of market power.

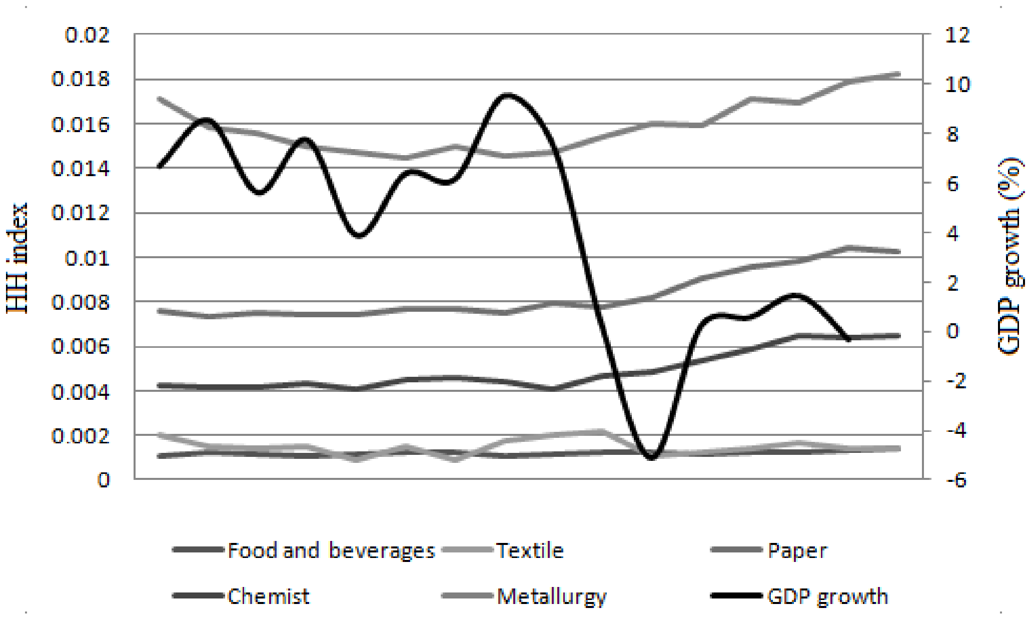
The major advantage of the Herfindahl index with respect to other measures of competition such as the concentration ratio is that it gives more weight to larger firms. In the absence of information about size of the firms in terms of production we use the number of employees as a proxy variable.

In addition to this index, competition can be studied with the number of firms by size and the average number of workers by sector. The change in number of firms by size allows determining if the change in competition is due to a reduction of the number of small enterprises or the increase in the number of big corporations.

The average number of workers by sector is an additional measure to know the level of competition. A high average number of workers implies that current established companies are relatively bigger than new ones. Big companies face lower average costs than small ones. Therefore, an increase in the average number of workers has a negative effect on the likelihood of success for new companies that face higher costs than the established ones.

There is available information for all the firms of Spain in five manufacturing sectors: food and beverages, textile, paper, chemist and metallurgy.

Figure 3.1. Evolution of Herfindahl index



Source: Own elaboration with data from DIRCE (INE) and Bank of Spain

As it can be seen in figure 3.1 all the five sectors studied are highly competitive or low concentrated.[21] Food and beverages and textile sectors are composed of small firms to a

[21] A Herfindahl–Hirschman index below 0.01 indicates a highly competitive index. A HHI index below 0.15 indicates a low concentrated index.

great extent. In addition, the crisis did not have impact on the degree of business competition. Paper, chemist and metallurgy sectors have a higher degree of concentration. These three sectors increased their concentration after the crisis emergence. Nevertheless, it is remarkable that the three sectors were the ones with a higher concentration and their Herfindahl index started to increase in 2008, at the beginning of the economic downturn as shown with the great downturn of GDP. This suggests the smallest firms had a higher closure rate than the bigger ones in these sectors.

The Herfindahl index suggests a decrease in market competition. Table 3.1 shows the decline in the number of firms in the last six years of economic and financial crisis in the Spanish economy, in all sectors and size groups. In the two sectors that did not increase the business concentration (foods and beverages and textile) the smallest firms were the ones that experienced the smallest closure rate while the biggest were more affected by the crisis. This behaviour suggests a reduction in the number of employees that lead to a change in the stratum of employees in a number of firms. Thus, a high number of small firms could have closed while most of the medium firms became small businesses as a result of downsizing.

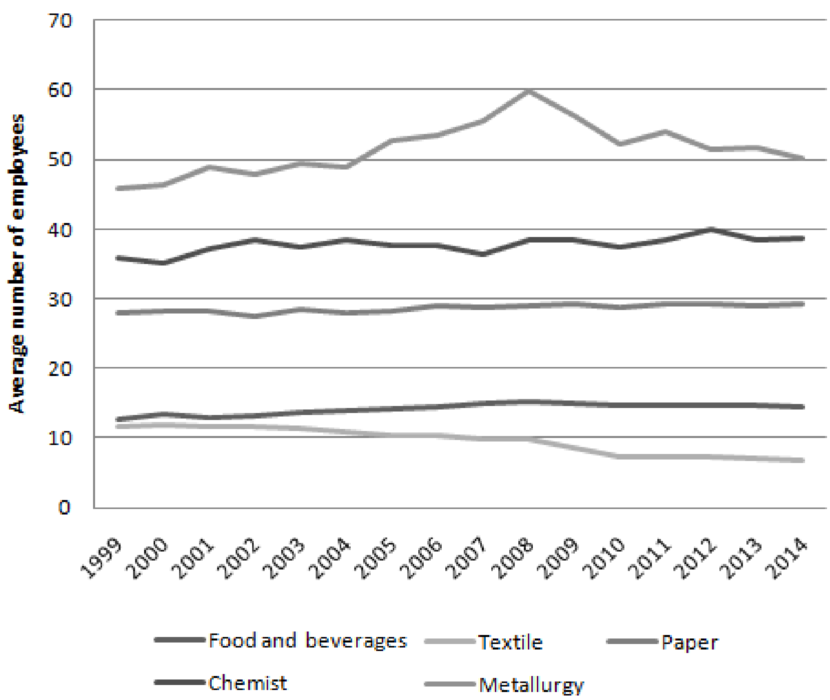
Table 3.1. Effect of the crisis on firms by size

Sector	**Year**	**Firms with less than 10 employees**		**Firms between 10 and 99 employes**		**Firms between 100 and 499 employees**		**Firms with more than 500 employees**	
		Number	Change (%)	Number	Change (%)	Number	Change	Number	Change (%)
Foods and beverage	2014	22779	-8.07	5,100	-11.57	408	-16.73	56	-20.00
	2008	24779		5,767		490		70	
Textile	2014	5,229	-26.85	773	-48.05	37	-52.56	0	-100.00
	2008	7,148		1,488		78		2	
Paper	2018	1,171	-12.35	522	-25.85	87	-5.43	11	-8.33
	2008	1,336		704		92		12	
Chemical	2014	2,461	-12.98	986	-26.64	137	-40.95	22	-48.84
	2008	2,828		1,344		232		43	
Metallurgy	2014	815	-8.22	394	-32.19	85	-21.30	22	-4.35
	2008	888		581		108		23	

Source: Own elaboration with data from DIRCE (INE)

In addition, the three sectors that increased their business concentration had a different behaviour by a number of workers. The most affected in paper manufacturing was the group between 10 and 99 employees while bigger firms show a better performance in the recent years. The chemical companies least affected by the crisis were the smallest ones and the biggest firms decreased their number by half. Nevertheless the Herfindahl index increased in this sector because the average number of employees in each group grew. Finally, in metallurgy the most affected firms were the groups between 10 and 99 employees while the biggest and the smallest firms showed a better performance. The overall outcome resulted in an increase of concentration and a more polarized size distribution.

Figure 3.2. Evolution of the average number of employees



Source: Own elaboration with data from DIRCE (INE)

Another indicator of market competition is the average size of firms in the sector. Figure 3.2 shows that the sectors with the lowest concentration are also the ones with the smallest average number of workers. The number of firms in each group by size and the Herfindahl index could remain constant and the average size of firm grows at the same time. In this way, new firms would face more difficulties to survive as their competitors have higher scale economies.

This indicator with data available from 1999 does not show any generalized effect of the economic crisis since 2008 onwards. The average size of firms in textile manufacturing has been decreasing since the beginning of the studied period. Foods and beverages, paper and chemical industries have had a constant average size. Finally, metallurgy had increased the average size until the starting point of the crisis. In the last six years this sector has switched this trend and again has an average of fifty workers per firm.

3.4. A simulation model for firm growth and firm dynamics

The hereby presented model is composed of four modules: Demand, innovation, firm growth and firm dynamics. The second to fourth modules could be removed and the model would still work but with less accurate results:

- In the first module the total demand in each sector is computed according to the growth of the economy in the last year. Then this demand is divided into the firms depending on their alignment with the preference of consumers in a set of variables and price.
- In the module of innovation each firm decides whether to devote resources to innovate. This expense makes it possible to change product characteristics (product innovation) and increase the production with the same amount of resources (process innovation) in a faster way than firms that do not spend on innovation.
- In the firm growth module firms decide if they increase, reduce or maintain their workforce depending on their turnover. According to the observed behaviour the decision is divided into four possibilities: no change, firm growth, small reduction and severe reduction.
- Firm dynamics is the result of the creation of new businesses and closing some of the already established. The companies that close are the ones with no activity in the previous period simulated. The number and characteristics of new companies

are determined stochastically taking into account the features of the competitors and economic growth.

This design strategy allows a sufficient degree of modularity. According to 'Best Practices in Programming Agent-Based Models in Economics and Finance' (Vermeir & Bersini, 2015) this modularity in agents' strategies increases robustness and it is one of the desirable characteristics of an agent-based model along with simplicity and clarity. Therefore the model does not have modules and agents that are not directly related with the research objectives.

The hereby presented simulation aims to discover how changes in demand due to different economic growth levels affect innovation patterns and firm growth. The calibration is performed adjusting the coefficients (denoted with Greek letters) in order to minimize the difference between simulated and actual results in the period 1999-2013. The equations of each module follow traditional economic theory in industrial organization and economics of innovation. However, this model is the first attempt to combine all of them in an agent-based simulation.

3.4.1. Demand

In the first part, the total demand (Dem) in each of the five studied sectors (s) is computed for the period t. It depends on the economic growth in the last period, as it changes the output level and increases the total demand of the following period. The demand of each sector behaves in a different way with respect to the change in the Gross Domestic Product (GDP). Hence, food and beverages and textile demand are inelastic to GDP ($\beta_1<1$) while the other three sectors decrease their demand at a higher rate than the decrease of total production in the previous year.

$$\Delta Dem_{t,s} = \beta_{1,s}\Delta GDP + \varepsilon \qquad [3.1]$$

In addition a stochastic perturbation (ε) is included. It represents external shocks that impact on exports and imports and it follows a normal distribution centred in zero and with standard deviation equal to minimize the error between the demand computed for 1999-2013 and the actual one. In the five sectors the standard variation ranges from 0.023 to 0.036 β_1.

Consumers calculate the values of a series of variables (PQ_i) related to Product Quality of different companies in each sector (s) and period (t). This process tries to reproduce the actual decision-making and it is based on the work of Chie and Chen (2014) when a similar rationale was applied to the case of the mobile phones market.

With these variables PQ_i the impact of economic growth on innovation and survival rate of the firms more adapted to market desired characteristics it is shown. These i variables are in the range zero to one. The vector of variables aims to cover all different characteristics of the product quality. They follow a random walk conditioned to the changes in the k last periods in order to have dynamics ($\delta<1$) conditioned to the past values. Then these variables tend to move in the same direction for a number of periods, representing different time evolving trends. For example, if one characteristic is the sugar content of drinks, the optimum value varies over time from a high level several decades ago to a low one nowadays. Companies that would like to produce goods with a high quality, as perceived by customers, would have to adapt their production to the evolving demands of consumers.

$$PQ_{i,t,s} = Max(1, Min\left(0, N(0,\sigma)\sum_1^k \delta_k PQ_{i,t-k,s}\right)) \qquad [3.2]$$

The more different the product of the firm ($PQ_{i,t,f}$) from consumer preferences at sector level ($PQ_{i,t,s}$) the less amount of product is demanded, the remaining characteristics equal. Price (P) is also taken into account to determine the quantity demanded from each firm (f), which is never higher than the production in the previous period. That is, a company would face a high demand if its production is cheaper than the average price of the sector ($P_{t,s}$) and it is perceived as closed to the consumer's definition of quality.

$$Dem_{f,t} = Max\left(1, \left(\frac{Dem_{t,s}}{Prod_{t-1,s}}\right)\sum_1^i\left(\alpha_i\left|PQ_{i,t,s} - PQ_{i,t,f}\right|\right) + \alpha_{i+1}\left(P_{t,f} - P_{t,s}\right)\right)Prod_{f,t-1} \quad [3.3]$$

This equation-based approach allows including most of the relevant theoretical factors of Economics while also they include an agent-based approach. As recently highlighted in Baptista et al. (2014) "more attention should be given to the explicit modelling of consumer behaviour with agent-based modelling" even if the model is company-centered. This approach combines the advantages of equational-based models with agent-based modelling that copes with the heterogeneity of consumers that have different ideas on what quality is and how important is quality with respect to price.

3.4.2. Innovation

Innovation plays a key role to decide the success of companies as aforementioned in section 1. Schumpeterian ideas about how creative destruction reconfigures sectors during time are especially important during periods of crisis (Archibugi, Filippetti & Frenz, 2013). Every company decides whether to devote resources to product and process innovation with two different objectives: Product innovation allows them to change the values of product quality (PQ_i) faster in order to adapt to the market variable requirements. Process innovation makes it possible to produce a higher quantity with the same number of employees. In this way, it also includes investment in equipment, not explicitly defined in the model. However, only a small portion of companies decide to allocate resources on innovation.

Firms (f) only devote resources to process innovation ($InnoProc_{f,t}$), that includes investment in equipment, when their previous demand ($Dem_{f,t-1}$) approximately equal to their production ($Prod_{f,t-1}$). The money allocated to process innovation is conditioned by the ratio demand-production and it is higher when the potential demand exceeded the production (β_2>0). In addition sector plays a role and some sectors invest more money in innovation than others (e.g. $\beta_{3,s}$ is higher in chemical than in metallurgy).

$$InnoProc_{f,t} = Max\left(0, \beta_2 \left(Dem_{f,t-1} - \beta_{3,s}\right)/Prod_{f,t-1}\right) \quad [3.4]$$

The main determinant of product innovation ($InnoProd_{f,t}$) is the inadequacy to market requirements in the variables that reflect optimum product characteristics (difference between $PQ_{i,t,s}$ and $PQ_{i,t,f}$). Nevertheless firms only devote a positive amount to this respect if they have profits, no matter how outdated their product is. This is an assumption only valid for traditional industries, and should be updated in the case of services and high technology industries where the production rapidly adapts to consumer requirements.

$$InnoProd_{f,t} = Max\left(0, \left(Profit_{f,t-1}/WF_{f,t-1}\right)\beta_4 \sum_1^i |PQ_{i,t,s} - PQ_{i,t,f}|\right) \qquad [3.5]$$

Finally, firms determine the price change (ΔP) for the period depending on the excess of capacity in the previous period (t-1). Firms with a lower ratio demand over production will decrease the price while firms with a higher demand for their products would increase the price to maximize the profits.

$$\Delta P_{t,f} = \beta_4\left(Dem_{t,s} - \beta_{6,s}\right)/Prod_{t-1,s} \qquad [3.6]$$

Initial values for β coefficients in this module were obtained using the microdata of the panel database PITEC that includes more than 12,000 companies and also previous analysis of innovation in Spain with this database (Sandulli et al., 2012; Segarra, & Teruel, 2014; De Marchi, 2012).

3.4.3. Firm growth

Previous attempts of agents-based models that were focused on growth (Mandel et al., 2010; Chang, 2010) emphasized on firm dynamics: firm entries and exits, but neglected the key role of the size change of current established firms. In the model, firm exits are just the deletion of companies with no employees and the next module is focus on entries. However, according to the industrial organization theory the change in firm size is the main process that affects sector concentration and therefore the level of competition.

In this part of the model, firms (f) decide their workforce for the next period (t) taking into account their revenue per employee. There are four situations: If ($Dem_{f,t}/Prod_{f,t-1}$)<γ_1 the firm carries out a large downsizing (equation 7). If this ratio is higher than γ_1 and lower than γ_2 the firm undertakes a small reduction of staff (equation 8). Firms only increase their workforce when the previously presented ratio is higher than γ_3 and in addition the firm profits ($Profit_{f,t-1}$) are positive (equation 9). A high portion of the companies have the ratio demand over production between γ_2 and γ_3 and they do not change their size.

This approach allows focusing on the behaviour of the two kinds of companies that are more affected during periods of crisis. On the one hand, those companies at risk of closing down face the need to perform great adaptations (equation 7) and they reduce their workforce until the return to making profits or they are forced to cease their activity. On the other hand, the firms that might still have profits, but need to take actions to converge to the average profits of the sector and not falling behind, also adjust their workforce to some extent (equation 8).

Additionally, this division in four sets allows for separating different behaviours and keeping the model as simple as possible without incorporating non-linear equations. Others simulations proved this approach and achieved a high degree of accuracy. In addition, the computational time is lower than making use of alternative approaches based on the economic theory (Pablo-Martí et al., 2013& 2014)

$$\Delta WF_{f,t} = Min\left(\rho_1 \left(Dem_{f,t} - \beta_{7,s}\right)/WF_{f,t-1}, -1\right) \quad [3.7]$$

$$\Delta WF_{f,t} = Min\left(\rho_2 \left(Dem_{f,t} - \beta_{8,s}\right)/WF_{f,t-1}, -\mu WF_{f,t-1}, -1\right) \quad [3.8]$$

$$\Delta WF_{f,t} = Max\left(\rho_3 \left(Dem_{f,t} - \beta_{9,s}\right)/WF_{f,t-1}, 1\right) \quad [3.9]$$

The coefficients ρ and β are calibrated in order to adjust the number of employees with the actual ones for the period 1999-2013. In equation 8, after calibration of coefficients, in some cases the change in the workforce was positive, so the alternative $-\mu WF_{f,t-1}$ was included, where μ is a set in one percent according to evidence in downsizing: practices in manufacturing firms (McCune, 1988).

Once firms have determined their new size they produce according to the production function in equation 10. This function does not try to represent the actual production, which is adjusted to the expected demand, but the maximum possible production with the current firm resources. Due to the lack of data I avoid making use of inputs and supply chain specifications. With this simplification the analysis is broaden to a great variety of situations with lack of detailed data about firms and increase the usability of the model for a higher number of countries.

The production ($Prod_{f,t}$) of each firm f in the period t is a function of the workforce ($WF_{f,t}$) to the power of α. This coefficient depends on the sector and is below the unit in all the cases. Therefore the production function is concave as usually described in the economic theory.

$$Prod_{f,t} = LP_{f,t} \cdot WF_{f,t}^{\alpha} \quad [3.10]$$

In addition, the limit of the production function is modified by $LP_{f,t}$. This variable represents the labor productivity and it is different for each firm at any period and it is computed in equation 11. The change of this coefficient depends on the depreciation (d) and the rate of technological change of the sector (θ_s) multiplied by the amount invested in process innovation. Consequently, process innovation has a direct impact on the firm as it increases the potential production of each employee.

$$\Delta LP_{f,t} = \theta_s InnoProc_{f,t} - d \quad [3.11]$$

The characteristics of the production ($PQ_{i,t,f}$) are modified according to the amount of money spent in product innovation. They approximate to the optimal value but they have a random error (ε) that represents the mistaken decisions of the board of directors, firm owners and the product development department. It follows a normal distribution centred in zero and the standard distribution is set as a five percent of the average $\Delta PQ_{i,t,f}$. Additional information would be needed to calibrate accurately the distribution of variation. However, it is convenient to include arbitrary behaviours of companies that make their production more or less appealing to the potential consumers, as these decisions affect companies' profits to a great extent.

$$\Delta PQ_{i,t,f} = sign\left(PQ_{i,t,s} - PQ_{i,t,f}\right) Max\left(\left|PQ_{i,t,s} - PQ_{i,t,f}\right|, \beta_{10} InnoProd_{f,t}\right) + \varepsilon \quad [3.12]$$

Finally, profits before taxes are computed as the difference between revenues and costs. Taxes (τ) are decreased to this amount in order to calculate the net profits. There are two kinds of costs: On the one hand, fixed costs depend on the size of the firm and they are calculated as a function of the number of employees ($WF_{f,t-1}$) and they are different in each sector ($\beta_{12,s}$). On the other hand, variable costs ($\beta_{11,s}$) depend on the effective production and involve inputs. The effective production is taken as the demand ($Dem_{f,t}$). In the long run these two quantities match as the stock changes represent a small portion of the sum of production in all the periods. In equation 13 revenues are computed as the demand multiplied by the price per unit sold ($P_{f,t}$) less the variable costs, and fixed costs are decrease to this amount.

$$Profit_{f,t=}\left(Dem_{f,t}\left(P_{f,t}-\beta_{11,s}\right)-\beta_{12,s}WF_{f,t-1}\right)(1-\tau) \quad [3.13]$$

3.4.4. Firm dynamics

In this section of the model new companies decide their characteristics. This process is performed in order to replicate the empirical evidence found in the database DIRCE that collects censual data of Spanish firms by sector and number of employees.

First, firms with no employees are deleted from the database. Then the number of new firms in each sector and their characteristics are determined by the total profit of the sector and the level of economic growth. There is evidence in the Global Entrepreneurship Monitor (Xavier et al., 2012) about two kinds of entrepreneurs: the ones who create a company because they are aware of an opportunity and the ones who have a business idea. The entrepreneurs of the first kind usually start several new businesses in different sectors that present a high profit. These entrepreneurs are the ones that initiate real estate companies during the real estate bubble, internet-cafes in the late nineties and are willing to create a company in any business sector that has higher profits than average. On the other hand, entrepreneurs with a business idea start a lower number of companies in sectors that do not have high profits. In both cases the possibility to start a company depend on the GDP growth, as institutions in the financial sector are less willing to fund new businesses during economic crisis (García-Tabuenca et al., 2015).

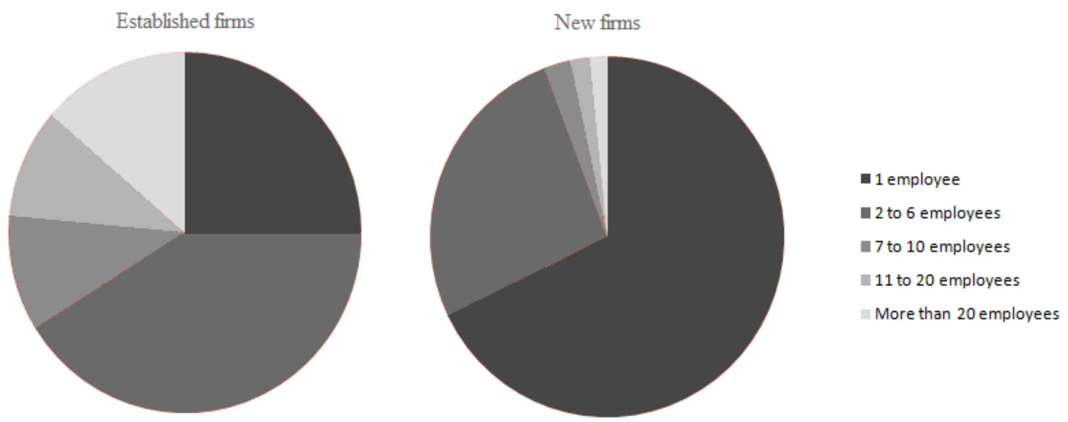
The number new firms (N) in each sector (s) are as aforementioned conditioned by the sector profits and the GDP growth. The main determinants of the three coefficients in equation 14 are related to the maturity of the sector and the size of the sector. New activities tend to have a higher rate of new firms than traditional activities regardless of profit level and GDP growth. Then, innovative sectors present higher values for coefficients β_{12} and β_{13} than long-established activities. New companies are created in novel sectors even if the established ones operate with losses and there is recession. However, traditional sectors have a higher size and then they have a higher number of new firms, although these firms represent a lower portion of the sector firms with respect to newer sectors. Then, a portion of entrepreneurs create new businesses in traditional activities especially during economic downturns and β_{14} is lower in sectors that behave comparatively better during crisis.

$$N_{s,t} = \beta_{12,s} Profit_{s,t} + \beta_{13,s}(\Delta GDP - \beta_{14,s}) \qquad [3.14]$$

The equation 14, once it is calibrated, reports a positive number of firms created in each period for every sector even during crisis, no matter if sectors have low negative profit. However, it may be the case that the number of new firms is lower than the firms that are closing down in the same period and the observed result is a net reduction in the number of firms.

Every firm has nine time-evolving variables: Demand is computed in equation 3, process innovation in equation 4, product innovation in equation 5, price in equation 6, workforce in equations 7 to 9, production in equation 10, labor productivity in equation 11, the string of product quality in equation 12 and profits in equation 13. Every new firm needs an initial value for all these variables. In addition firms have sector that cannot change during the periods of the simulation.

Figure 3.3. Size distribution in established and new firms



Source: Own elaboration with data from DIRCE (INE)

The initial size decision is one of the main elections of new firms. New firms tend to initialize their activity with a small size because they do not know their actual chances of success and they prefer to minimize the sunk costs in the event of failure (Evans, 1987; Dunne et al., 1989). As it can be seen in the following figure, the size of new firms is lower than the average size of current firms in the sector.

For the election of the initial number of employees, I make use of the data of new companies created during the period 1999-2013. No companies have an initial number of employees lower than one, and the probability of high number of employees tends to zero much more rapidly than in the case of established companies: 68% of new firms have only one employee, 27% have between two and six employees, only 5% of new firms have more than six employees (see figure 3). The data follows a chi-square distribution with one degree of liberty.

$$WF_{f,t} = int(\chi_{i,k}) \qquad [3.15]$$

The remaining decisions in the model are the initial price and product characteristics. These variables are set as the average in the sector with a deviation (Geroski, Mata & Portugal, 2010). In the absence of information about the actual distribution of price and product characteristics for new companies, normal distribution is adopted with a standard deviation equal to five percent of the average value.

$$P_{f,t} = N\left(P_{s,t}, \varepsilon\right) ; \; PQ_{i.f,t} = N\left(PQ_{i,s,t}, \varepsilon\right) \qquad [3.16]$$

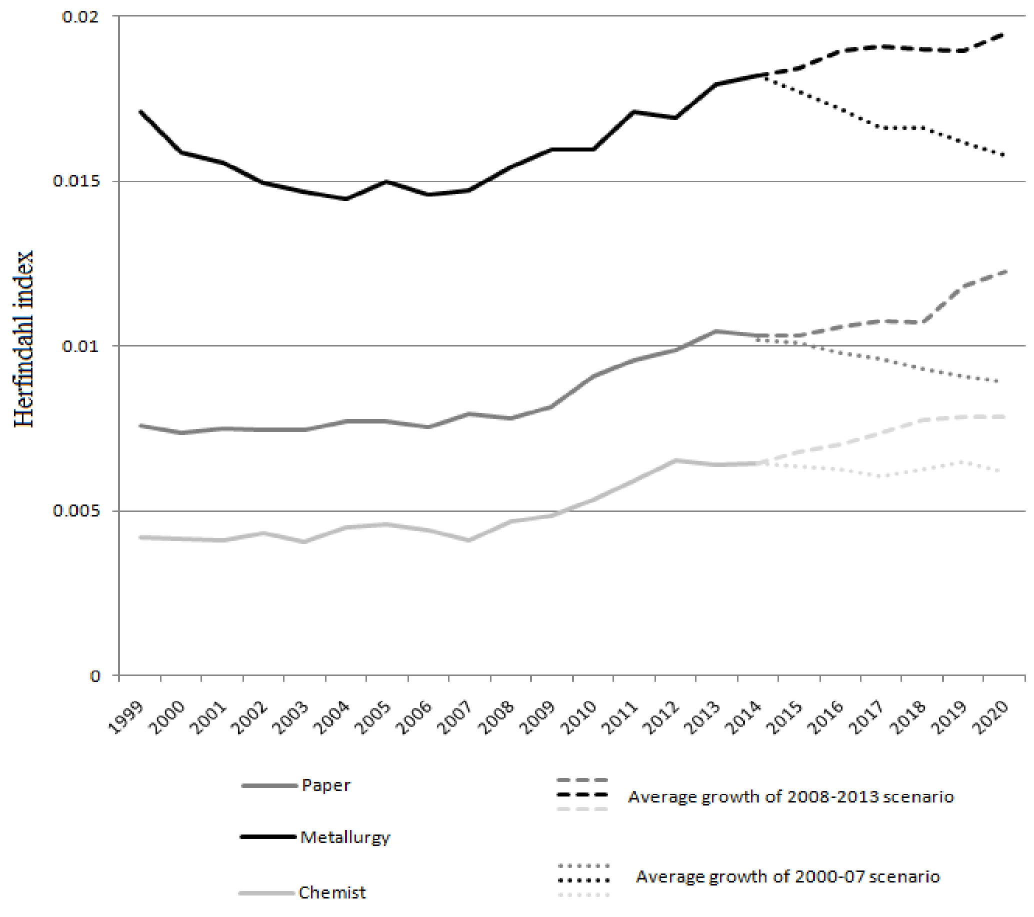
The heterogeneity with respect to the optimum values of price and product characteristic that maximize the survival rate of the new firms is due to environment restrictions such as financial constraints and agents' bounded rationality. The lack of perfect prediction of the impact of business decisions and the inability to compute the optimum behaviour are increasingly used in Economics. Agent-based models allow including bounded rationality preserving the appropriate level heterogeneity (Ebenhoh & Pahl-Wostl, 2007).

The labour productivity of new firms (f) is higher than the sector average (s), reflecting the new machinery and better practices of new companies with respect to the average of established firms ($\beta_{15} > 1$). There is a stochastic perturbation in order to include some disruptive new companies with significant higher productivity than the sector average as well as other new companies with lower productivity, reflecting the lack of know-how.

$$LP_{f,t} = \beta_{15} LP_{s,t} + \varepsilon \quad [3.17]$$

Finally, the initial value of production is set according to equation 10. The initial values for process innovation, product innovation, profits and demand are set to zero.

Figure 3.4. Herfindahl index under different scenarios



Source: Own elaboration with data from DIRCE (INE) and model outcomes

3.5. Market concentration forecast. The effect of economic growth

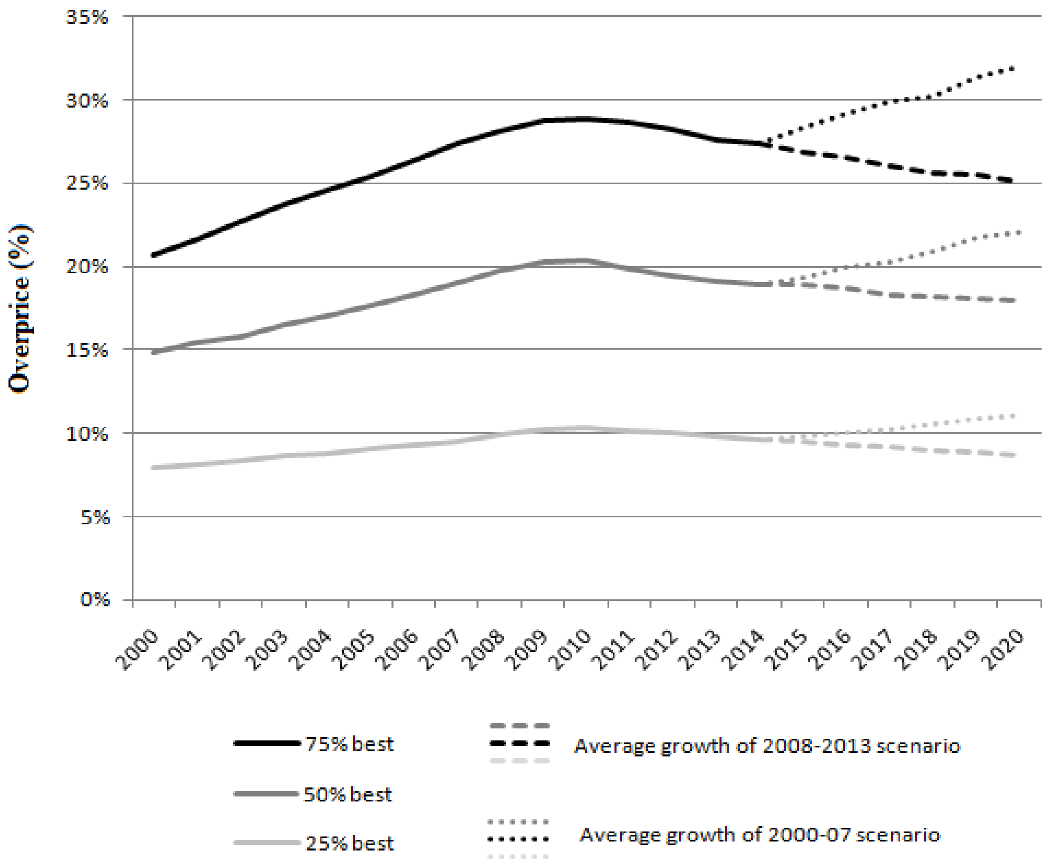
Calibration is one of the most important phases in agent-based modelling (Ballot, Mandel & Vignes, 2014). The coefficients of the model are calibrated with data for the five sectors in Spain in the period 1999-2013 using macroeconomic and microdata of firms available in Eurostat. The model shows the effect of economic growth on market concentration both with the Herfindahl Index in the following years up to 2020.

The only variable in the model that is not endogenously decided is GDP growth. It modifies demand as explained in equation 1. The rest of variables are endogenous and scenarios are set exploring different values for exogenous variables. Two scenarios of GDP growth are set. In the first scenario the average growth corresponds to the one of the period 2008-2013 (-2%) and in the second scenario is the average growth of the period 2000-2007, before the starting point of the current crisis (3%). In the first scenario the economy would continue in crisis and in the second one it would return to the growth path prior to the recession.

The forecast on market concentration is denoted by the dotted lines in figure 4 is the one if the crisis would end. The concentration would remain stable with high GDP growth in four out of the five sectors, excluding metallurgy that decreases its concentration level with economic growth. In the continuation of the crisis scenario, denoted by dashed lines, the concentration index tends to increase, especially in metallurgy and chemical, the two sectors with the highest concentration of the five industries studied.

The findings of the crisis scenario are in line with the concentration increase found during the crisis period (2008-2013). Sectors become more concentrated because small and medium companies behave comparatively worse than big corporations and there are a fewer number of new companies, that tend to be of small size. However, it is worth pointing out that the concentration in the simulations would only decrease in one of the five industrial sectors under study if the GDP would grow as it did before the crisis. Therefore, it is observed a tendency to increase market concentration in industrial sectors.

Figure 3.5. Prices higher than the lowest in the sector (simulation for 2000-2020)



Source: Own elaboration

Another key variable that changes in a consistent different way in both scenarios is price. If prices are ordered from lowest to highest in every period, overprice can be studied graphically in figure 5 depicting the prices that are in quartiles 1, 2 and 3 of the ordination. During the economic expansion (2000-2007) overprice increased in the three quartiles as an increasing portion of companies decided to raise prices in response to a higher demand. After the crisis emergence (2008-2013) prices went back to the best one of the sector at a smaller rate than the previous increase.

Nevertheless this MAS does not intend to model nor forecast the actual level of prices and this behaviour, then only can be analysed at qualitative level. However, this trend is maintained in the two scenarios: in the one of crisis extension (dashed lines) companies charging prices higher to the minimum one would continue diminishing the difference to the optimum. By contrast, in the scenario of a new economic expansion (dotted lines) the price variability would continue increasing as it did before the crisis.

3.6. Conclusions

The chapter focuses on changes in market competition taking into account the change in the size of the companies as well as firm entries and exits. An agent-based simulation is developed that includes changing requirements from consumers and the effects of different kinds of innovation.

The chapter presents the strengths and weaknesses of a broadly used methodology for studying the dynamics of companies and an emerging one: Game theory approach and simulation respectively. Also, it is analysed market power and its variation during growth and recession periods using a bottom-up approach.

The aim of the agent-based model described in section 4 is to include the components of corporate restructuring evaluated in terms of changes in pricing and non-pricing decisions such as human resources, innovation and product characteristics. The behaviours of companies under different situations are also taken into account. For example, there are two groups of transformation companies explicitly modelled:

- Those that are in a difficult economic situation, at risk of closing down and facing the need to perform great adaptations (equation 7)
- Economically healthy and taking actions to converge to the average profits of the sector (equation 8)

The main findings of the application to five traditional industrial sectors in the period 2000-2013 and the forecast of 2014-2020 are:

- The average number of employees in each sector is stable and there are great changes in the dispersion of the number of employees at sector level during crisis with respect to economic growth.
- Market power increases during crisis as a result of a lower number of new firms as well as the higher probability of small firms closing down in comparison with medium and big enterprises. Small companies are on average less adapted to

expectations of consumers in terms of price and product characteristics. This new finding is not incorporated in most of the models and it supplements alternative findings such as lack of funding and higher costs.

- Firms evolve during crisis periods converging to optimal attributes in terms of lowest price and desired product characteristics and there is a divergence during economic growth (Figure 5). In this way, crisis fosters companies to adopt a more efficient behaviour in order to survive while during economic expansion the increasing demand diminishes the Schumpeterian mechanism of creative destruction. In the case of Spain, after eight years of continued growth, companies that did not adapt optimally to consumers requirements faced a demand high enough to not trigger corporation restructuring.

Finally, this chapter aims to enhance the role played by the contemporary modelling multi-stakeholder in Social Sciences. Future research needs to be done on the dynamics and structure of supply chain and the incorporation of high technology sectors, where firms have other determinants that modify their decisions. Decision rules about mergers and acquisitions will also be included to study in detail their effects on concentration. The effects of mergers and acquisitions on market power were studied from an evolutionary viewpoint recently (Viegas et al., 2014) only using one market isolated from the rest of the economy. The model hereby presented would allow not only seeing this effect but also the interrelations between industrial sectors with the supply chain relations to be included in the demand module.

CHAPTER 4

A multiagent system of insurance and protection decisions

4.1 Introduction

Over the last few years research about economics of security has built interdisciplinary links and produced valuable common viewpoints from unexpected sectors. Security is becoming a multidisciplinary research topic that crosses academic borders (Anderson and Moore, 2009). The role of incentives is becoming as important as the correct design of computing systems in order to be trusted by potential clients. Multiagent systems is one of the tools that allows simulating and predicting the behaviour of agents and the system as a whole. It is possible to introduce the appropriate decision-making process and incentives scheme to produce a result and to explore the process dynamics.

To improve the security level of systems, engineers and economists use many different mathematical techniques, including concepts extracted from game theory and microeconomic theory. Multiagent systems have not yet been applied to this research area. However, most of the conclusions in previous works highlight the importance of agents' behaviour. For example, the level of security of a system depends on the effort of designers and users to keep the system safe.

In the same way, testing results depend on the effort of all the testers to provide valuable results. Therefore, the vulnerabilities of systems are closely-tied to activities conducted by the less-caring developer of the system (Anderson and Moore, 2006). Nonetheless, the authors suggest that under some circumstances, the vulnerabilities can also depend on activities performed by the individuals who usually perform relatively well during development of a system. In their opinion, the role of incentives on human behaviors should be studied from an economic perspective. However, effects of decisions on the rest of the users should also be taken into account, as it is assumed in a game theoretic approach (Grossklags et al., 2008a).

When a user invests in protection, other users also benefit from this action as it reduces their possibility of being attacked. Then, investments have a positive externality on the rest of the members of the system. At the same time, a selfish user does not have to have taken into account this externality when it chooses its security investment level. The security result of the network as a whole is suboptimal most of the cases (Jiang, 2009;

Kunreuther & Heal, 2003). According to Anderson (2001), in addition to the technical part, information security involves power relations, trade barriers and market segmentation with product differentiation. Asymmetric information and moral hazard are the two main concepts in this regard. As soon as regulators eliminate bad incentives, other organizations will be prone to create new ones. The model presented in this chapter deals with information transmission and management, and how agents decide to invest in their protection.

This chapter is structured as follows. The second part is focused on the interdependency of users in systems and the effects of externalities and, then, adequate incentive schemes to mitigate free-riding behaviour are discussed, which incentives are convenient to promote the exchange of information about security and, finally, the role of uncertainty in security decisions is studied. The third section claims that a game theory approach might not be sufficient to cover these aspects and propose a multiagent system as a novel approach to information systems security. Finally, it is concluded that an agent-based model is capable of including in its design the most important aspects related with security and offers a framework to simulate the behaviour of users under different scenarios.

4.2. Role of incentives in information systems

In this section the main findings of economics of security applied works are reviewed and it is discussed how these issues may be included within a game theory framework and how they can be incorporated into an agent-based model.

4.2.1. Users interdependency

Clearly, security investments by one user might reduce the dissemination of virus and other malware. One of the characteristics of network systems is the positive externality of security investments. Network security depends not only on the sum of individual security investments of the members but also on the interdependencies among them. Indeed, when a user does not care about the protection of its system the dissemination of malware becomes easier. On the contrary, when a user invests in its protection, then other users are positively affected as their infection risk is reduced. Consequently, any investment undertaken by the users has a positive externality for other users (Jiang et al., 2011).

Users do not take into account this externality when they choose their investment level and, therefore, security of the whole system will be suboptimal. In a non-cooperative environment, each user tends to choose her investment level in order to minimize both the risk and the cost of security (Jiang et al., 2008). This behaviour often ends with a degradation of security in the system as a whole. An agent-based model can introduce different levels of rationality and memory, allowing internalizing this positive externality.

Security does not come for free and tolerating a certain insecurity level may be economically rational. From an economic perspective, the key question is whether the considered costs and benefits by users are aligned with social costs and benefits. Rational security decisions from a user perspective might have costs for other involved agents or even for all the society. Incentives tend to have the correct directionality, but are not sufficiently strong to prevent a suboptimal provision of security. However, there are feedback cycles such as reputation effects that return part of the cost to the agent who originates it (van Eeten and Bauer, 2008).

It can also be studied how incentives change the decision between the investment in a public good (protection) or a private good (insurance) due to factors such as network size, attack types, probability of loss and its magnitude, and technology costs. Despite users affirming to have interest on their security, most of them do not implement any measure to prevent attacks even if security technology is available and reasonably affordable. It is necessary to know the main determinants for this lack of correspondence between users' actions and intentions in order to deal with this incoherent behaviour (Grossklags et al., 2008b).

These authors argue the lack of investment emerges as a dynamical equilibrium in which the optimal decision is not to invest for users facing the lowest potential losses. An increasing number of agents reduce or cease their security investment due to this behaviour. In order to reverse this, it is necessary to have a central planner who decides about the optimal security level and establishes incentives to contract protection. Once a sufficient portion of individuals is protected, protection is preferred over insurance because risk level has decreased. Through this behaviour the strong correlation between the decisions of agents can be valued.

Grossklags et al. (2010), he same authors, doubt the effectiveness of conventional game theory modelling, since in their behavioural experiment they discovered important differences between the observed behaviour and the predicted one: users do no tend to converge to one of the potential equilibriums and they have difficulties isolating the impact of each parameter from the impact of the decisions of other users. Thus, they do not follow correctly the economic incentives and the inherent complexities make it impossible to achieve good security results. Alternative approaches include fuzzy rationality to information security risk analysis (Henriques de Gusmao et al., 2016) Agent-based models take into account the interdependencies among users as they include all relevant interrelations between them and incorporate in the system how each agent react to the state of other agents. In fact, agents should have the following characteristics according to Wooldridge and Jennings (1995):

- Autonomy to interact in an independent way
- Social skills to interact among themselves
- Reactivity to perceive changes in the environment
- Proactivity to make decisions on their own.

The degree of rationality is described by the probability of making the optimal decision, taking into account the available information. The behaviour of these agents can be modelled in such a way that their simulated behaviour corresponds to the one observed in empirical tests. Under a low rationality scheme no equilibrium would be reached in the short run.

4.2.2. Incentives to reduce free-riding behaviour

According to Schechter (2004), security level can be expressed as a measure of the cost of resources used to penetrate in the system. Thus, it is fundamentally an economic measure. Users would like to predict accurately the risk of security systems and the effectiveness of different strategies to reduce such risks. From a business perspective, security is an investment that should be measured in saved money as a result of the reduction in expected losses due to successful attacks. Therefore, concepts such as return on investment (ROI) or internal rate of return (IRR) could be used to evaluate security investments and risk reduction. However, these measures are calculated ex-post based on historical data and forecasting future values is problematic. The frequency and severity of incidents changes over time, and is not likely to remain stationary as security is conditioned by the advancement of technology. Nevertheless, there are already some examples of models aimed at assessing the investment in security technologies (see e.g. Cavusoglu et al., 2004) that can be implemented in business applications.

Despite significant spending on security, companies developing software reward teams that send the software products on time. However, few firms offer employees the incentives needed to provide a sufficient degree of security. A large portion of vulnerabilities are discovered after selling the product. The developer who wrote the code and the tester that verified it may not be working on the same team, or even in the same company. A rational customer should pay more to developers but specify that the tests will be paid by the company in charge of the development of the program. Then, they will have incentives to achieve a high quality level of the initial delivered product.

Grossklags et al. (2008c) present a model of security decision-making that describes a controlled situation in which differences between the observed and predicted behaviour of the user are revealed: users show willingness to experiment with parameters, rarely converge to a fixed behaviour and face difficulties isolating the impact of individual influence the decisions of other users. Despite high economic incentives, the complexity prevents obtaining good results in realistic tests. Agent-based models allow including different levels of rationality in users of security systems. In this respect, they surpass the possibilities of game theory. The role of incentives can be tested with agent-based models under several rationality levels, understood as the decision of agents to change the parameters in a proactive way and different decision-making schemes.

Free-riding behaviour is observed in the provision of security when there are network effects. In fact, this behaviour is at all times observed when there are public goods involved. The overall level of security in a network has the characteristics of public goods: non-rivalry and non-exclusion. Under the voluntary provision of public goods appears strategic behaviour: individuals contribute less than the social optimal amount as they expect others will do the work. In general, positive network effects reduce the motivation for free-riding. A larger network makes most likely to take advantage of it without contributing. Therefore, mechanisms to significantly improve contribution levels are needed. (Kim and Shapiro, 2005). However, SMEs are unwilling to spend more in protecting information (Henson and Garfield, 2015) and specific legislation would be needed.

August and Tunca (2006), in the field of computer security software, developed a model by proves that in a network environment, where the maintenance of security of the software affects each user and, in consequence, the value of the software to other users, incentives can be a useful tool both from the private point of view as well as to maximize the welfare of the network as a whole. Particularly, four policies for computer security, both proprietary and free software are compared: protected users, mandatory protection, protection refund and tax on use. In the case of proprietary software, the seller internalizes the effect of any policy on users.

However, with free software it is not the case. In terms of economic theory, one could say that the role of the 'social planner' is more direct and critical. Therefore, for proprietary software, the preferred solutions are the protection of customers and the repayment of that spending. For freeware tax, the higher social welfare is reached with tax on use and spending refund options. Mandatory protection is suboptimal in both cases.

It is possible to conclude that incentives play a key role in security decisions to reduce free-riding. This behaviour is one of the main sources of lack of security in networks and the spread of malware among users. Agent-based models allow displaying policy scenarios with different measures and simulate the impact they have on heterogeneous users according to their characteristics and the response of the remaining users. In this way, these models are able to extend the work of August and Tunca (2006) and combine it with the limited rationality findings by Grossklags et al. (2008c).

4.2.3. Incentives to exchange information about security

Key factors to improve computer security include the collection, analysis and exchange of related information. Some companies are willing to cooperate in computer security issues and share best practices and relevant information system vulnerabilities. There are several positive aspects about sharing security information: the benefit of mutual exchange of information on security flaws can be divided into specific benefits for the company and an external benefit at the industry level. This private benefit includes both preventing new security breaches in the future (e.g., identification and repair vulnerabilities in their information system security), as well as an increase in demand due to the reputation for high security.

The results of Gal-Or and Ghose (2005) point out the benefits of information sharing in this regard. First there is a direct effect, which increases the demand; second there is a strategic effect, which reduces price competition. These incentives become stronger for bigger firms with a higher degree of competition. It is important to notice that the nature of the cost function of security technology plays a main role in determining whether the dissemination of information is beneficial or harmful and therefore if firms need external incentives for companies to join network and share information on security. Research by these authors suggests that if firms differ in their ability to use shared information, more efficient firms have greater incentives to underinvest in security technology and fewer incentives to share information, compared with less efficient companies. This opens up the possibility of an indirect free rider behaviour of the most efficient firms of the network.

The impact of incentives on information sharing about attacks is studied by Schechter and Smith (2003). They argue that organizations must take into account that someone seeking to enter in their system without permission will evaluate the potential financial gains in a global context. The hacker decides whether it is profitable to find and exploit a vulnerability based on all organizations that use that information system and their behavior once they find a potential vulnerability. Thus, it is possible to reduce dramatically the expected revenue of an attack if the vulnerability is disclosed and, subsequently, the probability of its reparation in similar systems increases.

Custom-made security tools are the most effective product as they do not generate subsequent revenues once any vulnerability is known. However, these tools are also the most expensive. Therefore it may be economically convenient to use a standard tool and maintain a reputation of information sharing. As a result, organizations that share information with other people constitute less attractive targets. Similarly, having a reputation for detecting unauthorized entries into the system is beneficial, as it will deter individuals who may have planned to launch attacks.

4.2.4. Uncertainty and security

The difference in profitability resulting from changes in the security options giving the level of available information is critical to make the optimal security decision. In this regard the concept of price uncertainty emerges as the measure that quantifies the maximum discrepancy in estimates between two information levels. Studying how the decisions and benefits of an agent change depending on the information she knows it is possible to determine how she reacts to the cost of security measures, the magnitude of potential losses, the initial security budget or endowment and the number of agents involved in the system.

Theoretical works have found that interdependencies among security settings of other agents do not allow them to accurately appreciate the impact of their decisions on other individuals and they take suboptimal decisions on the protection level. The greater the number of agents, the information becomes less important and agents can act better under high uncertainty levels. This is positive because a larger information network becomes more expensive to maintain (Grossklags et al., 2010).

Meanwhile, Gordon and Loeb (2002) show that for a broad class of probability functions for security breaches, the optimal amount to spend on information about security is an increasing function of the vulnerability of such information. Their analysis also shows that for a second broad class of probability functions the optimal amount to spend on this information does not always increase with the vulnerability level. For this second class, the optimal amount increases at first, but ultimately decreases with information vulnerability. The rational choice would consist of assigning most of the security budget to information that falls in the middle range of vulnerability to security breaches. Some information may be so difficult to protect that is preferred to spend a moderate amount of resources in an insufficient defence.

An agent-based modelling framework appears more appropriate to carry out this analysis as the effect of uncertainty can be tested under a set of hypothesis over the number of agents, the rationality level and the network effect, among other characteristics. Game theory constrains the scope of the analysis especially in terms of dynamics. With an agent-based model it is possible to show how the equilibrium (if any) is reached and how likely is to get to it under different scenarios. However, it might achieve optimal decision making if uncertainty is constrained (Zhang, Zheng and Shroff, 2015).

After this review of the most important findings in relation with economics of security, a multiagent system that includes the main concerns previously discussed will be presented.

4.3. An multiagent system for economics of security

A good model should provide information of the aggregate outcome of a system but it is not enough. It should also be capable to explain the behaviour of the components of that system and their interactions that determine such aggregate outcome. Understanding a socioeconomic system requires more information and knowledge than understanding only the behaviour of the organizations that compose the system. The overall effect of modelling interactions shows that the results can be more than the simple sum of its parts. In mathematical terms, additivity should be rejected and the focus should be on the concept of 'emergence' (that is known from statistical physics)[22].

Agent-based modelling is the appropriate technique for studying systems with plenty of decision makers acting at the same time as they use methods that assume the complex interactions between agents. They describe emergence properties of simpler interactions

[22] In systems theory emergence is the process where larger entities or aggregate behaviour appear due to interactions among smaller or simpler entities. The larger entities or aggregate has properties or characteristics the simpler entities such as agents do not have available.

between them that cannot be deduced only from the aggregation of properties of all individuals (agents). For instance, the system vulnerability cannot be treated as a sum of vulnerabilities of its components (agents). It is an interaction between them that may increase the system vulnerability. The common example from other field of study is the interaction in financial markets. The systemic risk is more than simple the sum of vulnerabilities of each of banks composing the financial system. The same idea can be easily applied to IT systems.

The other advantage of agent-based modelling is that it allows for introduction of elements that are not easily modelled using other mathematical tools. For instance, if interactions are analyzed, it should be remembered that agents make decisions based on their experience, but they also constantly adapt to new scenarios (they learn or imitate) and they usually expect some outcome in advance. These elements are hardly implemented in game-theoretic framework. As first explained in Epstein and Axtell (1996), and later re-elaborated in Tesfatsion (2007), Agent-Based Computational Economics is the calculation of economic processes modelled as dynamic systems of interacting agents. The term "agent" refers to a collection of individualized data and the methods of representation of an entity defined in the virtual system. These agents may be living beings, social groups, institutions or companies.

This is a relatively unexplored approach that tends to study economic systems holistically. Once the behavioural rules are specified, all subsequent events are driven by the interactions between agents. These interactions are determined dynamically by the internal structures of behaviour, the information level, beliefs, motivations and data processing methods. A crucial issue of multi-agent systems is that there are no prior conditions that could limit interactions between agents and that usually they are subjective. Strictly speaking, there is no fine-tuning as defined in statistical physics. The agents in these models will tend to be completely free to act and interact with their virtual realities as it is done in the real world by its counterparts.

In order to create a multiagent system to facilitate the understanding of the economic process that underlies information security, three criteria have to be followed:

- Firstly, the model should include the correct classification of agents (they should be correctly identified using empirical methods and then replicate/simulate them).

- Secondly, the model scale should be appropriate for studying the phenomenon (choose how many agents and interactions should be studied)

- Finally, the specification of the model should be subject to the empirical and technical validation to ensure that final and intermediate causal mechanisms are modelled adequately (the model truly represent the empirical findings, especially causal effects).

Before developing a model, phenomena under study should be classified into categories. It facilitates the detection and correction of errors while allowing for the comparison of results. The classification should be performed paying attention to the empirical findings highlighting its importance in the process of theorizing and modeling. Before proceeding to the classification, the researcher should know the answer to the following issues:

- Types of needs which are relevant for understanding the economic phenomena of interest.
- Types of goods and services that satisfy such needs.
- Possible presence of agents that supervise the design and operation of organizations and, where applicable, their motivations.

Agent-based modelling as such does not offer answers to these questions as it is rather an approach than a theory. But without doubts, these models provide a systematic way of incorporating any classification that is relevant in the study of any particular socioeconomic phenomena such as the motivations and actions related to information systems security.

Multiagent systems are typically implemented using object-oriented programming to develop the skills of each agent and their relations. Hence, they allow the start of the modelling process and lately introduce further required changes in order to match the model to the studied economic reality.

The choice of the number of agents is critical for their impact on the economic activity, in contrast to game theory-based models, where once the number exceeds several players it does not have a main impact in most of the cases. On one hand, in agent-based models when market participants are sufficiently numerous, there is perfect competition among buyers and sellers. On the other, when a market consists of a seller and many buyers, agents face the highest degree of market power. There are plenty of intermediate positions, where the rivalry between buyers and sellers leads to imperfect competition.

In addition to a realistic degree of rationality during the decision-making process, imitation might also be fundamental in these models. Even if imitation is usually considered a second-class learning in conventional theories, it is the essential part of evolutionary models. This mimetic capacity is much higher in human beings than in the case of other animals. By contrast, the individual learning, especially merely making use of the personal experience, is the most widespread mechanism in social sciences, though it is usually slow and inefficient.

In order to define the imitation patterns correctly, the networks of agents should be developed adequately. The establishment of such networks between agents is highly subjective and no sufficient empirical support has been given, since there are no statistics and studies sufficiently developed to provide information about the way individuals convey information and how they respond to economic issues at a given moment of time, such as job searching, investment and consumption. In the case of information security there is not sufficient available information about these networks and it would be unreasonable to assume any specification.

4.3.1. Model specification

In this section an agent-based model is developed to study the main determinants of the security decisions by network users is described. Heterogeneous individuals decide the optimal amount of two goods: Insurance and Protection in each period as in Grossklags et al. (2008b). Moreover, in this case a game theory framework that was ineffective to predict the actual behaviour of individuals in an experiment is avoided (Grossklags et al., 2010).

The model parameters are:

- In the system there are N agents related among them. They represent users of information systems like organizations.
- Every agent receives an endowment M. If her security is breached she has a loss L. The endowment and the loss are real positive numbers that represent utility and also can be measured in monetary units.
- Each agent chooses her protection level R_i and her insurance level S_i. Those can be zero or one. If they are one it means the information system of the agent is protected or she has active insurance. If R_i is zero the agent is not protected in that period and if S_i is zero the agent will not be refunded in case of a successful attack.
- Protection has a fixed cost C_R that represents the fee paid for a security system offered by a specific provider. It is a real positive and protection is only selected when C_R is lower than L.
- Insurance has a variable cost C_S. At the beginning it has the same cost of protection. Later, it is equal to the portion of agents successfully attacked in the previous period multiplied by the endowment (fair insurance). If the population of agents is inadequately protected C_S will tend to increase.
- The initial probability P_{A0} of being attacked is defined exogenously. It follows a random uniform distribution centred in P_A with different result for each agent and each period.

In the second and subsequent cycles this probability of being attacked depends on the number of new attacks and successful attacks in the past. A high successful attack rate in the past promotes more attacks, because the incentives of attackers increase. If the individual is protected she cannot be successfully attacked. Otherwise, the formula $P_{Ai}=P_{A0}(1-R_i)(1+C_s)$ remains relevant.

Hence, a higher rate of individuals protected would decrease the probability of being attacked. The effect of insurance is the opposite, since those agents would have no incentive to protect their systems and attacks would have a higher rate of success.

The utility of agents in every period is computed as $U_i=M-P_{Ai}L(S_i-1)-R_iC_R-S_iC_s$

The endowment M is decreased by the loss L only if the agent did not contracted insurance and in any case she pays the fees of protection and insurance.

The following decision rules are taken into account:

1. Utility maximizing agents
2. Agents invert their previous decision if their utility was below the average utility
3. Agents imitate the decision of one agent with a higher utility

These decision rules are simplifications of actual decision-making processes observed in reality. They represent different rationality levels. In the first case, the agents have the maximum level of rationality, but it does not include a direct communication mechanism among agents, as they only take into account their result and the aggregate situation in order to decide their option for the next period. In the second case, the rationality level is lower, since they only act if they obtain a result under the average, that is, if they appreciate that they are in worse situation with respect to other modeled organizations. Additionally, imitation mechanisms can be taken into account.

4.3.2. Utility maximizing agents

Utility maximizing agents are the ones who develop a strategy aimed to maximize their expected utility. Agents try to maximize their expected utility. The optimization problem they face is the following: Max (R_i, S_i) $E(U_i)= E(M-P_{Ai}L(S_i-1)-R_iC_R-S_iC_s)$ They try to face the lowest risk of attack and at the same time recover the loss if the attack has taken place.

They compare the value of that maximization problem in all the four possibilities: no/yes insurance and no/yes protection and they choose the best one. In case of a tie in the expected value, they select one strategy randomly as there is not an optimal one.

1. $R_i=1, S_i=1$ $M-C_R-C_s$
2. $R_i=0, S_i=1$ $M-C_s$
3. $R_i=1, S_i=0$ $M-C_R$
4. $R_i=0, S_i=0$ $M- P_{A0}(1+C_s)L$

According to the assumptions, the first strategy is dominated by the second and the third ones under any parameter configuration. Agents decide if they contract protection, insurance or nothing according to the price of both, the exogenous probability of attack and the potential loss.

Thus, the decision-making process is the following:

If $min(C_S,C_R)>|P_{A0}(1+C_s)L\ |$ *then* $R_i=0, S_i=0$

Else if $C_S=C_R$ *then* $Ri=0$ *with* $p=1/2$ *and* $S_i=1-R_i$

Else if $C_S<C_R$ *then* $R_i=0, S_i=1$

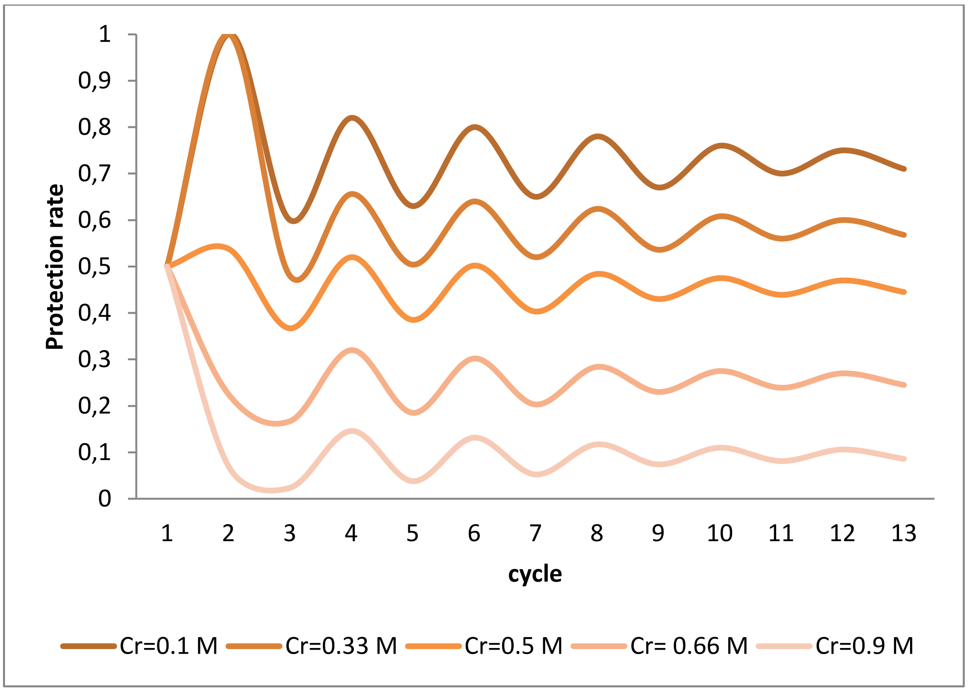
Else $R_i=1, S_i=0$

For each decision-making rule, the effect of different protection cost and the amount lost in the event of successful attack are considered. It is worth noticing that the differences of results between two runs diminishes when the number of agents increases, but the average result of Montecarlo Simulations presented in the chapter is not affected in all the scenarios.

Effect of protection cost

When the protection cost is null, the utility is always equal to the endowment because every agent chooses to protect herself and there are no successful attacks. However, in the case of a scenario without successful attacks the fee of the fair insurance would be zero and the agents would randomly choose between protection and insurance. This random selection would lead to a situation with a positive rate of successful attacks. In the next period, the insurance fee would have a positive value and every agent would get protection. Then, a sawtooth pattern would be observed. If the price of insurance is updated taking into account the attacks over the last k periods the equilibrium (as defined in agent-based modelling by properties extracted from statistical physics) is complete protection and no successful attacks.

Figure 4.1. **Effect of protection cost on protection rate**



Source: Own elaboration

When the protection cost, expressed as a portion of the total endowment M increases, the average protection rate decreases and the successful rate of attacks will increase as well. Then, the average utility of the agents also decrease and the average utility in the equilibrium is equal to $E(U_{i,t}) \leq M\text{-}C_R$. In some runs, the average utility is exactly the difference between the endowments but there are exogenous factors that can reduce it due to a high rate of successful attacks and the subsequent increase in insurance fees.

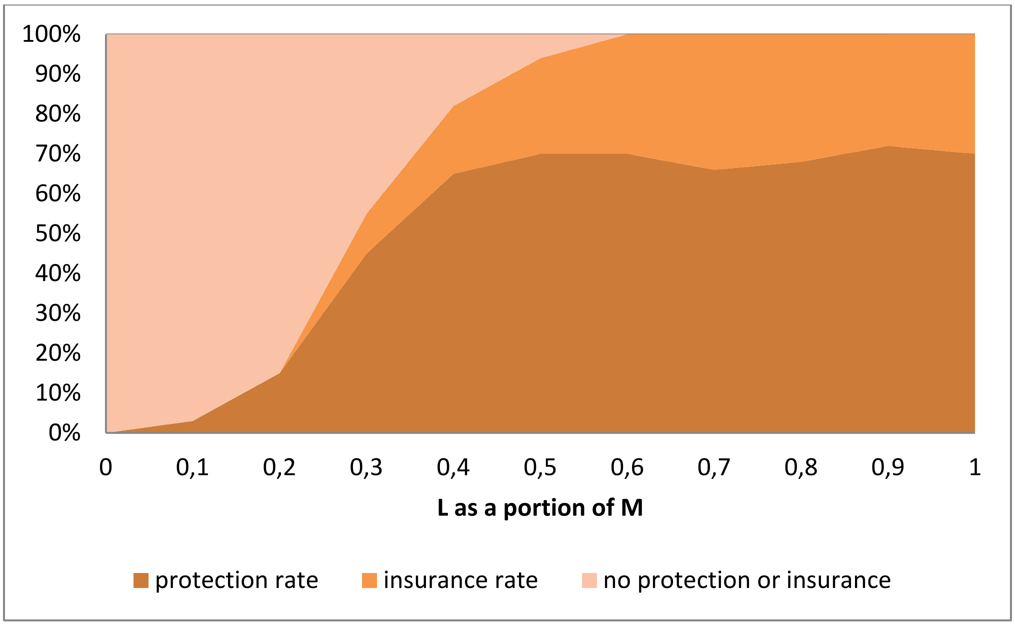
In the event of a cost of protection equal to the endowment, the agents do not protect because the cost of insurance is always lower except in the case every agent is attacked. This effect can be seen in Figure 1 where the protection rate in the equilibrium for C_R=0.5 is lower than ½. This protection rate is lower than expected for every value of C_R.

Effect of the amount lost

The amount lost affects only the decision to invest in insurance or protection when the price of protection is high. If the loss is small enough the optimal behaviour is the one where the agent is neither insured nor protected. It is rational because there is less to lose and the cost of insurance (protected side) does not offset the fee:

If $min(CS,CR)>|P_{A0}(1+C_s)L\ |$ *then* $R_i=0$, $S_i=0$

Figure 4.2. **Effect of the amount loss on agents' behaviour**



Source: Own elaboration

As it can be seen in Figure 4.2, the rates of protection and insurance vary as a result of the change in the amount lost, if the rest of parameters remain constant.

The values for protection and insurance rates would also depend on other parameters such as C_R. The exact shape of Figure 4.2 would change but it is observed that if L is a small portion of M the optimum strategy is neither get insurance nor protection to maximize the expected utility. Otherwise, agents decide between protection and insurance independently of the value of L.

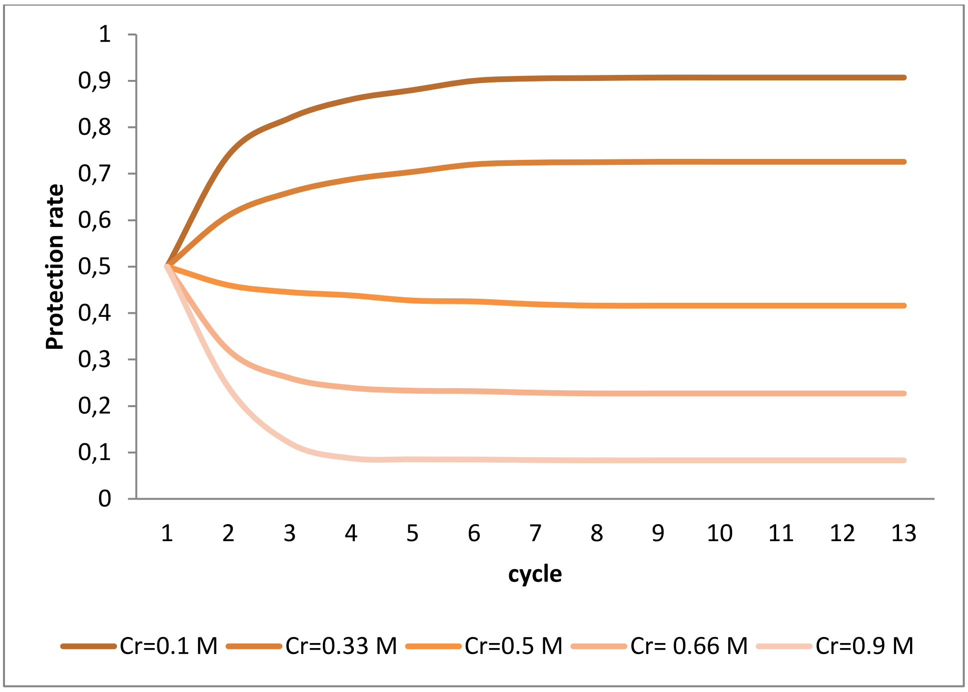
4.3.3 Agents invert their previous decision if their utility was under the mean

In this case, agents have less information. They do not know the previous successful attack rate and they only know the average utility of the agents and their last utility. In that case, if utility is below the mean of the population's utility, they change their decision from protection to insurance or from insurance to protection. With a small probability, they can also decide not to get insurance or protection in order to achieve the equilibrium with that decision with any value of parameters.

Effect of protection cost

Under this assumption for the decision-making process the equilibrium is reached in a smaller number of periods in comparison to the previous scenario.

Figure 4.3. **Effect of protection cost on protection rate**



Source: Own elaboration

In this case the different cost of protection leads to different protection rates. Higher protection cost reduces the protection rate, increases the rate of successful attacks. Therefore the average utility of agents is reduced.

Effect of the amount lost

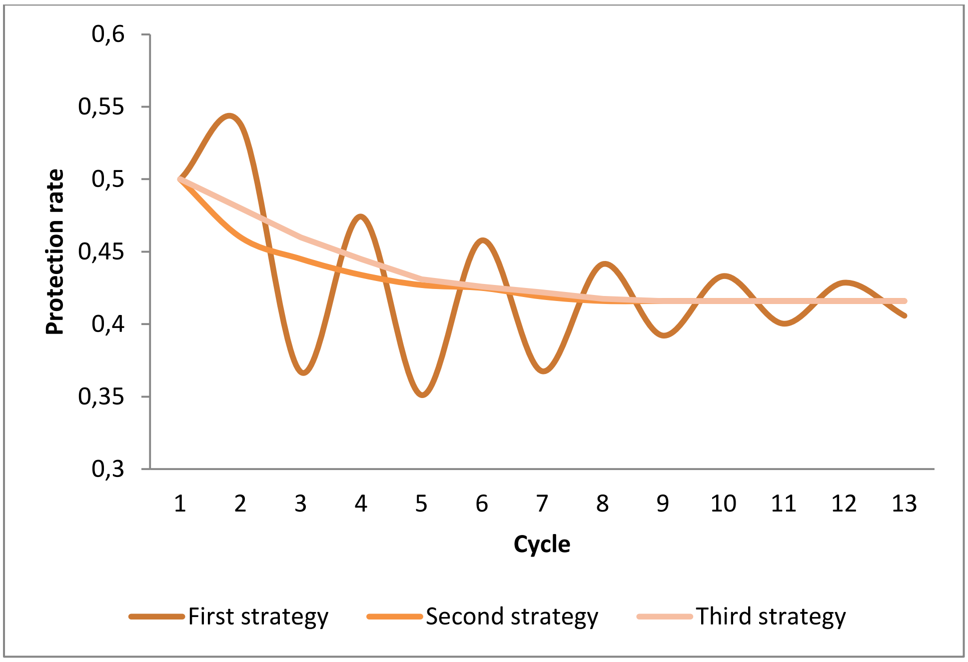
The utility is $U_i=M-P_{Ai}L(S_i-1)-R_iC_R-S_iC_{s.}$ and the agents only take into account the loss value when there is no insurance ($S_i=0$). If agents are protected $P_{Ai}=0$ and their utility should be higher than the mean if a portion of agents are successfully attacked.

If L is low and C_R or C_S are sufficiently high, the optimal decision is not to acquire insurance or protection (see Figure 4.2). This equilibrium is achieved faster than with the decision-making considered in Section 4.3.2.

4.3.4. Agents imitate the decision of an agent that has a higher utility

In this case, agents communicate in pairs and if their utility reached in the previous period was different the agent with the lowest one adopt the behaviour of the agent with a higher utility. This is equivalent to the second decision-making process in the long run. However, the equilibrium is reached more slowly as the agents have less information about the optimal decision.

Figure 4.4. Effect of the implemented strategy on protection rate



Source: Own elaboration

Figure 4.4 shows the equilibrium reached for C_R=0.5 is the same for the three approaches. However, in the third case, with direct connections of pairs of agents the equilibrium appears later than in the case of agents knowing the average of utility of the population.

The same happens for the case of changes in the amount lost in the case of a successful attack. In the tenth period, the equilibrium reached with the second and third strategies is the same for any value of L (see table 1). However, up to the fifth cycle, the rate of each strategy is more distant from the equilibrium reached by the third strategy. For example, if the potential loss is one half of the endowment the rate of protection in the equilibrium is approximately 0.74; in the second strategy when every agent compares her performance with the average this rate is 0.73 on average using Monte Carlo simulations. If the strategy is imitating the behaviour of only one contacted agent with a better performance, this rate would be just 0.63.

This shows that the amount of shared information allows agents getting to the best state for all the population in a lower time. The different number of cycles to arrive to an equilibrium depends on the shared information in each cycle and how agents process that information. Both information sharing and its internalization in the decision-making process are vital to obtain the best state of the system state and the maximum expected utility for every agent. Nevertheless, optimization is sensitive to external parameters such as the amount lost (L) in case of attach.

Table 4.1. **Evolution of the behaviour in strategies 2 and 3 with different L**

		t=5			t=10		
		R=S=0	R=1	S=1	R=S=0	R=1	S=1
strategy 2	L=0.2 M	0.79	0.2	0.01	0.81	0.18	0.01
	L=0.5 M	0.05	0.73	0.22	0.06	0.74	0.21
	L=0.8 M	0	0.72	0.28	0	0.72	0.28
strategy 3	L=0.2 M	0.61	0.29	0.1	0.81	0.18	0.01
	L=0.5 M	0.15	0.63	0.22	0.06	0.74	0.21
	L=0.8 M	0.09	0.6	0.31	0	0.72	0.28

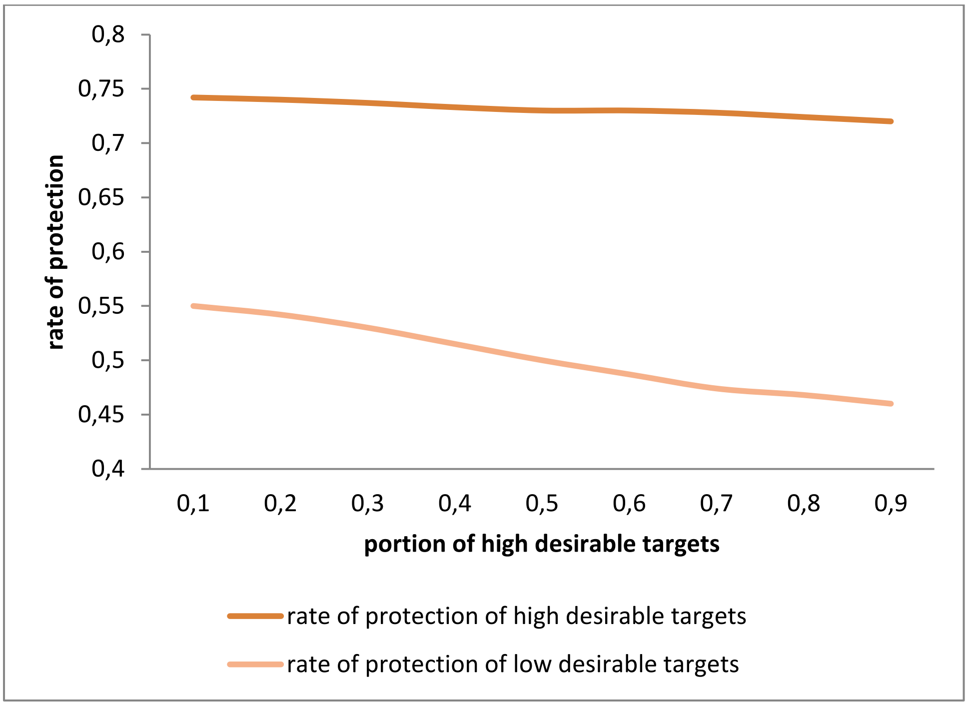
Source: Own elaboration

4.3.5. Effect of agents with different interest as targets

In reality, the agents differ, they are heterogeneous (not homogenous as it is assumed in many models in different fields of study such as economics, finance or even information system security). A natural extension of the agent-based approach presented above, would be to model the heterogeneity of agents. In this way, a portion of agents could have a higher value as targets for attackers due to their more valuable information. This happens in the case of banks and other firms that use sensitive information of their clients and their vulnerabilities can report a higher amount of money in comparison with average objectives.

These agents have a higher potential loss (L) with respect to their endowment (M) and therefore they will be asked for a higher insurance fee. Then, they have more incentives to develop protection systems to ensure they are not going to suffer successful attacks.

Figure 4.5. **Effect of the portion of desirable targets on protection rate (t=10)**



Source: Own elaboration

If an increasingly high portion of agents represent a more desirable target to attackers (L=0.8 M) while the rest are regular objectives (L=0.4 M) the average rate of protection of all the population decrease. This is a result of a positive externality of the distinct motivations of both subpopulations. The most concerned for security issues decide to protect as it is the optimal decision for them. Once they are protected, the rest of agents have lower incentives to acquire protection mechanisms as the rate of attacks is reduced (Figure 5). With this result, one weak point of this model is observed: Attacker motivations do not respond to behavioural changes of defenders. In a real situation attackers would take into account the higher rate of protection of the most desirable defenders and would change their activities towards the less valuable targets that would be less protected.

4.4. Final discussion

The analysis of incentives appears to be one of the most promising approaches to know the determinants of suboptimal provision of protection systems. The free-rider behaviour of network users makes them less secure than desirable. The lack of protection is only rational from an individual point of view but it is not optimal for the entire population.

Standard game theory models provide key tools for the understanding of the behaviour of agents but they do not achieve the results of behavioural experiments. Agent-based models are able to represent the decision-making process of security systems users and allow for testing different assumptions and observe how incentives and decisions change.

The simple agent-based model presented allows observing the role of incentives with a low number of parameters in order to keep the model as simple as possible to represent the issues of suboptimal provision of protection and free-riding behaviour linked to positive externalities from one subpopulation to the rest of agents.

In the future the agent-based model should include information about actual agents that interact among themselves as it happens in real networks in IT systems. Then quantitative results will be added to the qualitative results extracted from this model, the first one of its kind focused on the economics of security in IT systems.

CHAPTER 5

Consequences of trading hours deregulation. A spatio-temporal object-oriented data system for Madrid region

This chapter study the effects of the deregulation of trading hours in Madrid during 2012 using a complex approach that incorporates individuals and establishments located in Madrid region taking into account their activity and position with a one hour time lapse.

The first section study the change in the regulation and the main literature on this field.. The second one describes previous advances in the development of databases to study the behaviour of consumers from a spatial perspective. Then, it is necessary taking into account the temporal component of consumption with the proposal to combine the best aspects of this new technique and for developing an object-oriented data system suitable to evaluate the consequences of the deregulation. In the third section the initialization mechanism and the calculation of the demand of individuals are explained. Additionally, it is discussed the mechanism of behavioural change of establishments. Finally, there are presented the preliminary results and further work needed to extend the scope of this novel approach.

5.1. Trading hours in Spain and Madrid deregulation

The process of a higher liberalization in trading hours began in Madrid in 2008, with the Law 1/2008 of June 26, of Trade Modernization (Modernización de Comercio), with an extension of the opening hours for retail establishments. With the Law 2/2012 of June 12, Commercial Revitalization Act (Ley de Dinamización Comercial), full freedom of trading hours was given to retail establishments.[23]

The regional government in office, right political oriented, emphasizes that this Law would be an important support to the employment and the trade competitiveness. They affirm freedom in trading hours benefits consumers, as they have a greater schedule that allows them to a better distribution their purchasing activities and schedule them more conveniently.

On the contrary, the opposition parties argue that it encourages irrational and impulsive consumption and it destroys part of the small and medium business which cannot keep

[23] Until the adoption of this law 2/2012, freedom in trading hours only was applicable to establishments over 300 square meters. Nevertheless, it must be highlighted that despite the deregulation was not complete it was more flexible than other Spanish regions policies.

up with the large commercial facilities, thereby reducing the total volume of sales and the employment level in this sector. In addition, an extended trading service is associated with a higher cost, which, at least in part, would impact the prices of products, and therefore it would damage consumers.

The consumer opinion on this issue in Spain is strongly polarized. The official Centre for Sociological Research (CIS) asked about this issue in October 2011 (Díaz de Rada, 2012). 53.1% of the sample was in favour of opening hours regulated by general legislation and not be freely set by retail establishments. In addition, 51.8% of respondents preferred a limited time rather than a complete deregulation. Only 2.1% stated that the limitation of opening hours was affecting their shopping. Meanwhile, 61.3% of consumers did not perceive any disadvantage in current trading hours.

The large interregional variability in Spain of trading hours and holiday closings as well as the criteria that define a shop as a big establishment can be seen in the study by De los Llanos and Mora-Sanguinetti (2012), They concluded that more regulation generates a lower employment in the sector and a higher inflation level. Stricter regulations aimed at protecting traditional establishments have served in the food industry to encourage supermarkets to the detriment of large stores.

A noteworthy article on commercial distribution in Spain by Juan Velarde (2012) confirmed these findings on employment and inflation and therefore encouraged to harmonize regulation and re-define it to combine the objectives of employment, prices and protection to the smaller establishments. It also offers a revealing conclusion of retail sector quoting Paul Samuelson: This sector should not be studied with the tools designed for perfect competition but as a market of inefficient and specialized producers with a low initial capital that allow them to stay in the market until they lose their stock of capital. They do not compete in prices and they try to charge the highest possible prices and share the market.

From my point of view it is necessary taking into account in this analysis the importance of spatial dimension, which allow the specialization of otherwise identical establishments. Distance to consumers makes only a small part of them accessible and then the importance of local markets emerges as a key characteristic of retail sector. Moreover, it has been added the time variable which limits competition when only a portion of the stores is opened.

Aranda & Santos (2012) showed how less regulation is correlated with increased time spent on buying. The two most deregulated autonomous communities, Madrid and Valencia are on average where more time is devoted to buying activities. It also differentiated consumers into five groups according to their characteristics and shows that the group most strongly in favour of deregulation is characterized by people between 26 and 45, with a high percentage of single individual and most of them currently employed.

Fernandez (2012) is in favour of the complete deregulation of trading hours. He argues with examples from other countries that the effect of this measure is not linear or unidirectional depending on the size of the business establishments. The effect depends on the type of business, location and the relationship they have with large retail chains, which on occasions are the providers of small shops. Besides the effect that deregulation would have on competition in electronic commerce, it should also be taken into account how that policy might reduce it.

In addition to the regulation of trading hours, entry barriers to retail chains also pose a major barrier to competition. Despite the adoption of EU services directive and its legal adaptation in Spain, entry barriers have not been eliminated as in several regions the second opening license of an establishment of the same company depends on the size of the establishment. This policy limits the number of new supermarkets, department stores and large size clothing stores. The policy change of 2012 in Madrid also eliminates such measures and this new regulation could reduce by 3% the number of small establishments, while medium and large size could increase up to a 30% in the simulation conducted in Asensio (2013) for Spanish regions.

5.2. Databases for trading activities. The design of a spatio-temporal database for Madrid region

Trading studies have increased its potential in recent years with the construction of spatial databases that have individuals and establishments. This new approach relies on the increasing popularity of object-relational databases, an extension of the traditional relational data bases, which are provided with features of object-oriented programming.

The information of an object oriented database is represented by objects to facilitate system performance. When the characteristics of a database are combined with those of an object-oriented programming language, the result is an object oriented system capable of working in coordination with a complex model. This system makes the objects of the database appear as objects in one or more programming languages such as Java or C++.

To study trade patterns an object-relational database is not solely enough and it is necessary to add the spatial dimension by using a geographic information system. These systems are an organized integration of hardware, software, and geographic data designed to collect, store, modify and analyze geographically referenced information in order to solve planning and management. Until just over a decade ago, these systems were not truly dynamic but now the time variable can be introduced to achieve the following objectives:

- Modifying the position of objects at any given time. For example changes in the location of commercial establishments, movement of people and goods.
- Changing the characteristics of objects recursively. For example changing the activity of an establishment, crop rotation, etc.
- Adding and removing objects. For example urban and industrial development, business closures, new transportation routes.

Even nowadays, most of the geographical information systems use multiple static databases to represent time, which is used as an attribute of some specific events (Kavouras, 2003). The spatial change in most of the cases is not modelled correctly because it is necessary to build spatiotemporal databases that collect events occurring at any time and altering not only the characteristics of objects such as their position but also the relationships between them. In this case, a commercial establishment that is closed at one moment will not generate consumer demand. In addition, consumers will not be attracted to areas where no shops are open at that time, which in turn will lead to the stores not having incentives to remain open.

The combination of agent-based models and object-oriented spatiotemporal databases began to be used a decade ago to study the movement of people and goods as well as to characterize non-work activities, where they occur and how long they allocated to each of them (Buliung, 2001). This approach started to be used in business applications and had offered increasingly good results as the costs had decreased exponentially and the accuracy of data has increased over the last decade.

Empirical scientific papers show labour and leisure activities are usually located in urban centres and there is a correlation between employment and commercial activities both in urban centres and outlying areas. This concentration results in a high traffic congestion and increased journey times that have to be taken into account in the calculation of time uses. In addition, individuals have different movement behaviours depending on their age, employment and family classification.

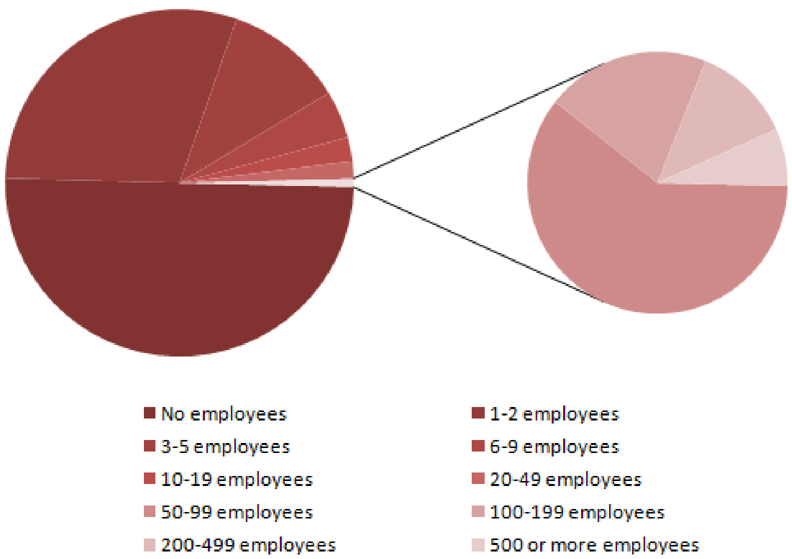
The study of transport demand has begun to use object-relational databases to study the determinants of the behaviour of travellers and freight companies. Then, temporal component is as important as space. To carry out the analysis it is important to use geospatial data transport networks such as roads and public transport modes. The addition of transport surveys is necessary to model properly the trip timing, reason and mode. This makes it possible to reach an ordered set of activities and trips over time (Frihida et al., 2002), and the use of transport networks allows calculating the speed of each trip (Buliung & Kanaroglou, 2004). Then, it is possible to know where each individual will be located and what they are doing at any time making use of a statistical spatio-temporal behavioural model.

Besides the importance of transport and workplace, there is a set of consumer characteristics such as income, gender and age that modify their consumption habits. In addition, businesses have a number of characteristics that modify their potential demand including location, trading hours, price and quality of products (Schenk et al., 2007). This set of information makes it possible to model how purchases made by each individual will be distributed among the establishments through a probabilistic microeconomic model.

In order to build the database to allow carrying out the proposed analysis it is necessary to combine the available information from the a set of databases.

SABI (Sistema de Análisis de Balances Ibéricos) is the Iberian counterpart of Amadeus European database. SABI is a database that contains general information and annual accounts of over 1.25 million Spanish companies. The software lets companies search using any combination of criteria, such as the regional and sector level. This database includes all medium and large companies but only contains a small part of the microSMEs with fewer than 10 employees and it does not contain information about self-employed individuals.

Figure 5.1. Local units of sector 47 in Madrid region



Source: Own elaboration and DIRCE

The distribution of local units in the region of Madrid for the sector of retail trade[24] it is presented as an example (see Figure 5.1). Half of the local units (34715 out of 69351) have no employees besides the self-employed entrepreneur and more than a quarter have only one or two employees. However, SABI data shows a distribution of firms with most of the companies with a size of 20 to 50 employees and it does not include firms with fewer than three employees.

The Central Directory of Companies (***DIRCE*** in the Spanish acronym) combines all Spanish companies and their local units in the country into a single information system. It is updated once a year, generating a new information system every 1 January. With this information the National Statistics Institute publishes the results for enterprises and local units, classified into regions, main economic activity and level of employees. The CCD generates information associated with entrances, closures and current number of firms classified according to their economic sector, legal status and workforce level.

This example suggests that for building a geo-located database of all establishments it must first be taken into account the coordinates of all the companies in SABI ordered by sector and then it is possible to create the ones of the remaining establishments according to this information.

The procedure follows selecting the coordinate of one establishment available in SABI and correcting it according to a bivariate normal function centred on the current coordinates and a standard deviation such that the spatial concentration is preserved when it is measured with a grid of a mile square with a kernel function applied following the methodology described in Pablo-Martí and Muñoz-Yebra (2009) and extended in Pablo-Martí and Arauzo-Carod (2012) to avoid the problem spatial aggregation known by the acronym MAUP (Modifiable Areal Unit Problem).

The merger of these two databases permit to obtain the income and the expenditure of local units according to the data of enterprises taking into account the sector and the number of workers, the first two objects of the database are built (see Figure 5.2): WORKPLACE includes all establishments, their coordinates and opening hours, the

[24] Retail trade for motor vehicles and motorcycles is not included. Then, it corresponds to sector number 47 in the current industry classification (CNAE 2009).

latter depending on the sector. For the sector 47 it is also created the object SHOP that takes into account the income and expenses. Retailers are divided into subsectors according to the classification of Table 5.1.

Table 5.1. Number of retail establishments by subsectors

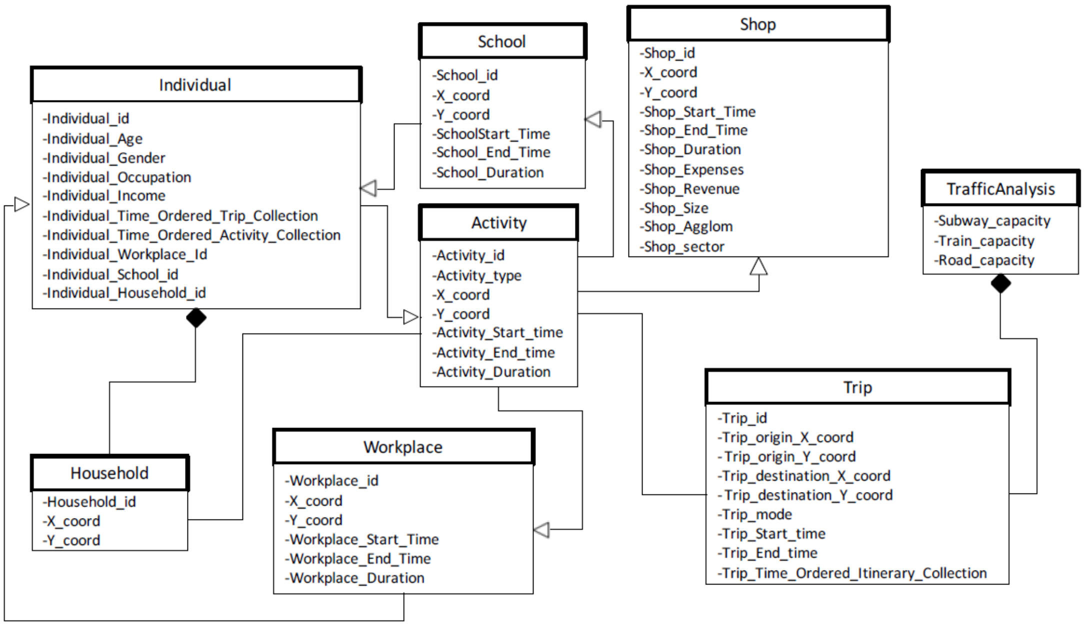
Classification	CNAE 2009 sector	No. of establishments
Food products, beverages and tobacco	4.72	12,653
Automotive fuel	47.3	520
Equipment for ICT	47.4	921
Other household items	47.5	15,763
Cultural and recreational goods	47.6	5,118
Other items or non-specialized stores	47.1, 47.7, 47.8, 47.9	34,376

Source: Own elaboration and DIRCE

The ***Population Census*** of 2011 was conducted in all countries of the European Union and in the Spanish case it provides microdata for 10% of the citizens in each region of Spain. It includes information on gender, age, occupation and zip code, company sector of those employed, and the zip code of the workplace and the educational centre. Each household is geo-referenced and each observation can be replicated ten times to get to the entire population. It is performed the same statistical spatial process to arrive to a household distribution with the same spatial concentration as the 10% sample. The location of each school is created by replicating a location of a random home in that zip code and adding many schools in each area as the integer part of the result of dividing the number of students in the area between five hundred. Thus, with census data is possible to build INDIVIDUAL, SCHOOL and HOUSEHOLD objects (see Figure 5.2).

The ***Wage Structure Survey*** is carried out in the European Union with common criteria methodology and content with a four-year lapse, in order to obtain comparable results on the structure and distribution of earnings between the States of the EU. This survey investigates the distribution of wages based on a variety of variables such as gender, occupation, sector of activity, age, or size of the company. The latest information available is for the year 2010 and it is used to assign the variable wage (Individual_Income) using a statistical matching process between the agents of object Individual and the observations of this survey. This entry is corrected to take into account capital gains so that the net national income data after tax is computed for Madrid region in 2012. The process allows arriving to the total value for this variable and keeping the statistical properties of the original distribution, with the same value of the Gini index for individual income.

Figure 5.2. Database design



Source: Author elaboration

Once the information about wages is aggregated to individuals, this object can be joined by other statistical matching process so each worker is assigned her appropriate workplace minimizing the spatial and sector errors and taking into account the sum of wages in every establishment which minimizes the difference with the expenses reflected in SABI. In the absence of a centralized statistical source of the places where public employees perform their work, they are distributed randomly in the workplaces of the private sector.

Finally, the ***Household Mobility Survey*** is carried out in Madrid region and surrounding provinces and it is the main statistical source for the transport behaviour in this region. It has an eight-year lapse but the survey of 2012 is still under progress at the latest data available are the ones for 2004. This data show the mobility patterns of Madrid citizens. They include the transport mode and the reason for travelling.

This data makes it possible to know for every occasion (work, study, purchase and others) what are the trip patterns according to the distance to the household; the mode of transport; the characteristics of the traveller; and, the places of departure and destination. With this information it is possible to achieve the necessary equations that calculate each trip probabilistically.

Spatial information of road networks, rail and metro (TraficAnalysis) is added to the MAS. Roads have a maximum capacity and a function is assigned to compute the potential speed based on their level of use at any time.

5.3. Agents behavioural rules

Once all individuals are located in their households, workplaces and education facilities, as well as all the locations where it is possible to shop, the next stage is to populate the ACTIVITY object with the activities carried out by individuals:

- Staying at home
- Staying in the workplace
- Staying in the education facility
- Shopping

Therefore each activity takes place at the location of an object in one of four classes: Household, Workplace, School or Shop. The simulation model should establish what times of day and day of the week each activity is performed. To move between each of the locations the TRIP class is created. This object in the case of a trip by road uses the information contained in the geographic information system as well as the rest of the trips that are taking place to calculate the speed and duration of the trip.

Activities that do not involve the Shop object as a point of origin or destination are performed according to the schedules of the School or Workplace objects. To calculate activities involving trade, and given the lack of data in this regard from the Household Mobility Survey, which focuses primarily on commuting to the workplace, a simulation model is used that relies on the work of Schenk, Löffler and Rauh (2007) and combines it with the multinomial logit developed by Arentze and Timmermans (2001)[25]. The analysis also takes into account the findings of Teller & Reutterer (2008) that show empirically the importance of agglomerations in shopping behaviour.

The utility obtained by each individual agent in the case she purchases at each store is defined as follows:

$$U_{i,s} = Min\left(\frac{1}{d_{h,s}}, \frac{1}{d_{w,s}}, Min\left(\frac{1}{d_{sc,s}}\right)_h\right)\sum_f P_{i,f}\, A_{s,f}$$

Where the utility (U) at individual level (i) in the case she purchases in a shop (s) is the minimum distance between the shop and home (h), workplace (w), or the minimum distance from the school (sh) of any household member. This minimum distance is multiplied by the sum of the individual's preferences (P_i) in each factor (f) taking into account the characteristics of the property (As). The factors I include are the agglomeration of establishments nearby, the size of the premises and opening times:

- Agglomeration is the amount of shops within a 150 meter radius as in (Schenk et al., 2007) but considering only those that are open in that time slot. After computing this variable, it is divided in each period by the maximum value for resize it to a maximum value of one unit. It is estimated that all individuals have a preference for establishments located in agglomerations according to the majority of empirical findings, but mostly younger people and those with more time to spend shopping.
- Shop size is expressed as the square root of the number of workers. Young individuals have greater preference for larger ones whereas older individuals

[25] This article is focused on the behaviour of the multipurpose movements to shopping areas as well as the requirements imposed by information included in the database.

prefer smaller establishments.

- The utility is computed at each time slot and is zero when the establishment is closed.

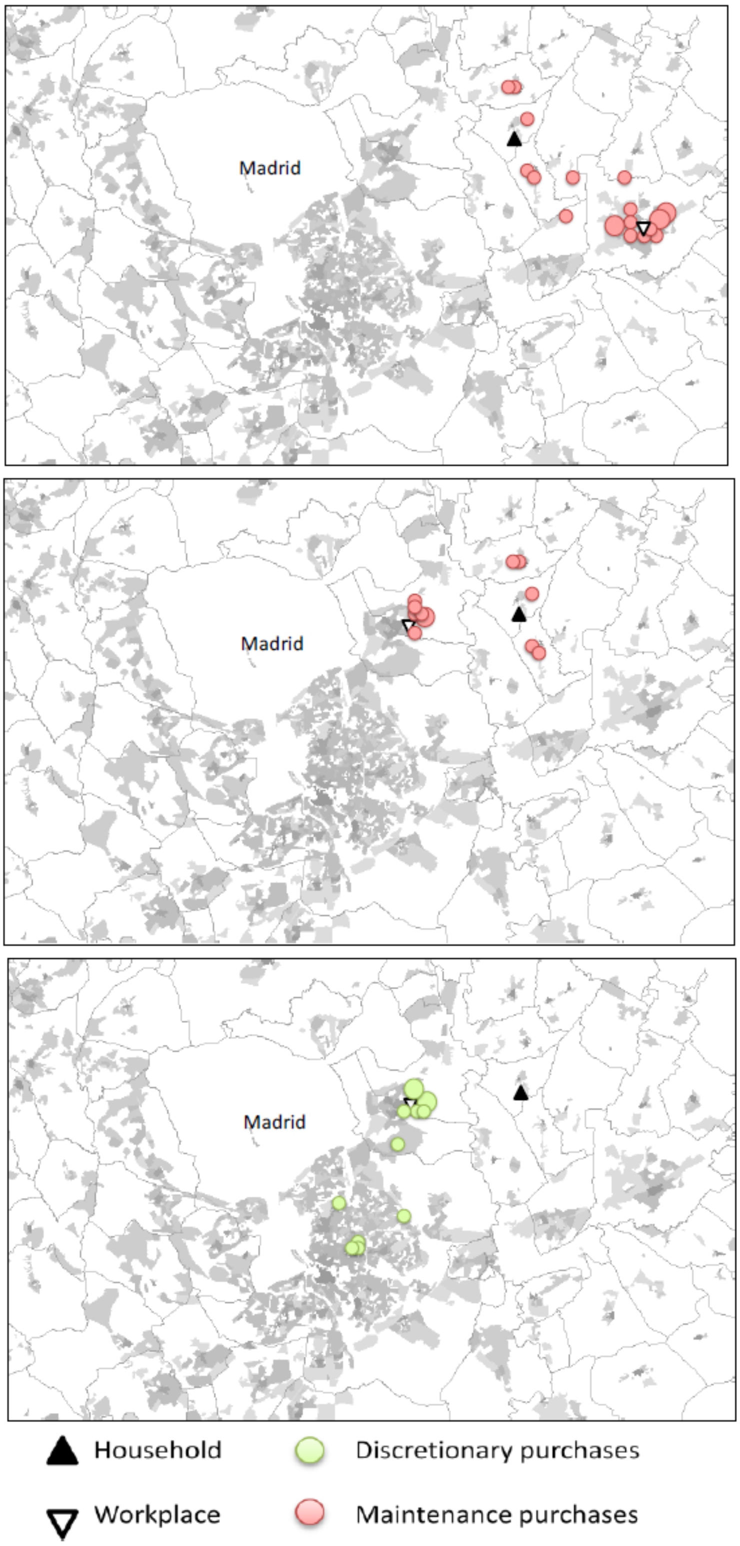
Once each individual computes the time devoted to work and to study she estimates where she goes shopping. The expense is divided into a part for maintenance cost and other for discretionary consumption. In the part of the maintenance individuals buy the products sold in stores included in subsectors 47.2, 47.3 and 47.5. Discretionary items include goods sold in the remaining subsectors of retail sector. The share of income that goes to each purchase is obtained from the statistical matching of the observations in the database with microdata from the ***Consumer Expenditure Survey*** carried out in 2012.

Individuals randomly purchase between one and three time in sector 47.2, a purchase of 47.3 sectors and 47.5 randomly and a discretionary purchase. Individuals decide in what shop to buy in each case according to the utility via a decision rule assigning p=U(0,1) where the unit is the most useful shop and the correspondent portion is assigned to other shops so that imperfect rationality is included in the preferences. These purchases are made outside working hours or school and never between midnight and six in the morning to avoid having to assign different probabilities to each time zone without data to support the choice. However, both days of the weekend are taken into account as possible times to purchase.

Thus, the importance of individual preferences $P_{i,f}$ is re-scaled to achieve a deviation of no more than a 5% in estimated sales compared to those reported in the balance sheets of companies in any establishment when the average of a high number of iterations is observed. However, due to the random nature of the purchases computation important variations arise between the calculated demand and the reported one for some establishments.

Figure 5.3 shows the patterns of maintenance and discretionary consumption of an individual and the importance of the workplace to choose the shops she buys to. As I have developed a probabilistic model only those establishments where the individual consumed more than 5% of the times are depicted in order to avoid including a large number of shops that are represented on the maps.

Figure 5.3. Main shops in which an individual purchases



Source: Own elaboration

To observe the effect of trading hours deregulation two processes are implemented: business demographics and decisions about trading hours.

First, for each monthly period of four weeks, shops with the lowest ratio of revenue/expenses per employee close and then the same number of establishments opens located less than 500 meters away from an existing establishment. The number of employees is taken from a random establishment currently open and thus the dynamics shows how the size of establishments evolves after a number of periods.

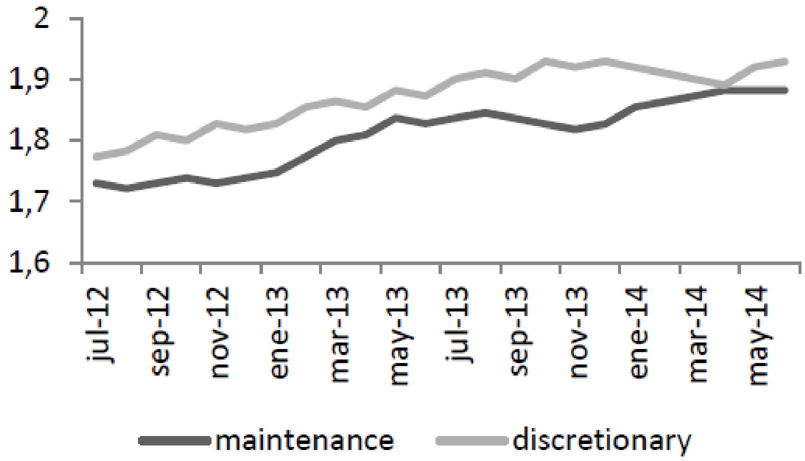
Second, each shop decides its trading hours for the next period so that more time allows its demand to be greater because it allows for some time slots to report a high utility as others shops would remain closed and thus transfer the purchases to those shops that extended their opening hours. Establishments can increase or decrease their opening and closing times by one hour. They also decide whether to open on Saturdays and Sundays proactively in a small percentage of cases or imitate the behaviour of competitors (Wooldridge & Jennings, 1995) in most of the occasions.

5.4. Main results

In the next figures are shown three of the most relevant results of the implementation of the trading hours deregulation. In the first case it is studied the effect of the deregulation on the size of the establishment through the average number of employees. Secondly, it is shown the impact on the average opening hours per week. Finally, it is simulated how many establishments would decide to open on Sundays. The results shown correspond to the average of one hundred simulations.

Figure 5.4 shows with the deregulation the average number of employees increases from 1.75 to 1.9 in the first two years after the law was applied. This dynamics is also found if the law was not passed, but the increase tends to be higher in simulations with the scenario of deregulation implemented. Retail establishments for discretionary consumption have a higher average number of employees and maintain this different in the period under study.

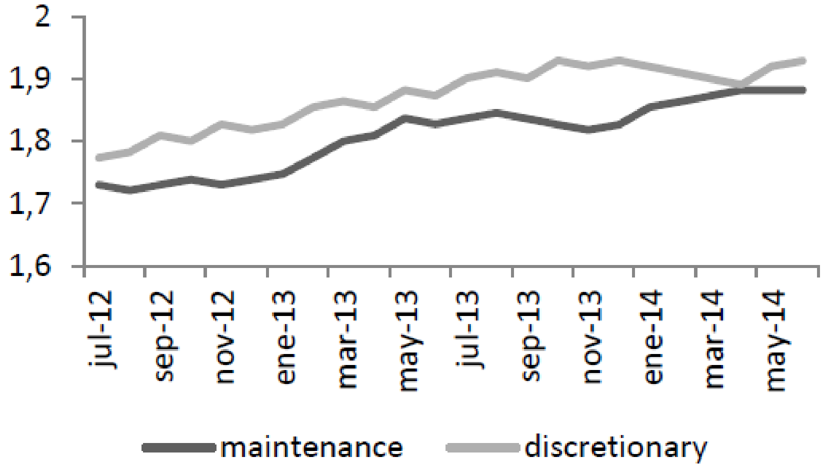
Figure 5.4. Effect of the deregulation on the average number of employees



Source: Own elaboration

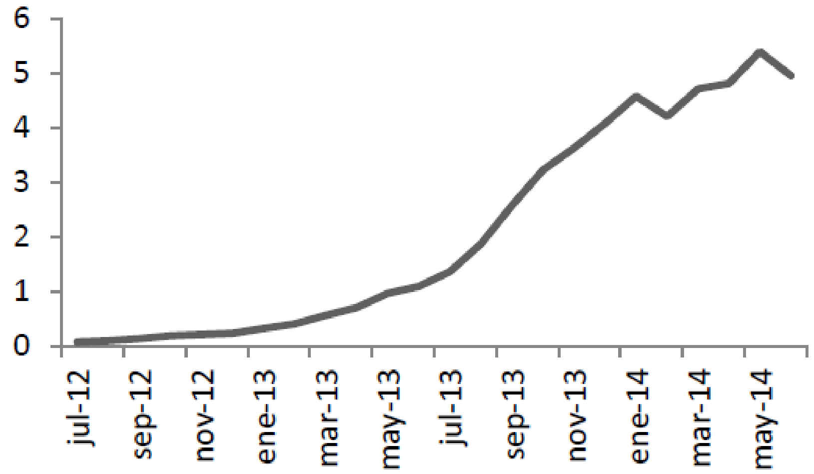
Figure 5.5 presents the effect of the deregulation on the average opening hours per week. Since the first period of the simulation the number of hours increases about 1.75 on average. Retail establishments of discretionary consumption tend to increase more the opening hours in response to the deregulation. After the first decision-making process some of the establishments decide to go back to the previous time schedule, others maintain the increase in opening hours and finally other shops imitate the ones who react to the law in the first place and also open more hours per week. The aggregate result is a slightly increase in the average opening hours per week and this variable stabilizes after one year in levels 2% higher than before the deregulation.

Figure 5.5. Effect of the deregulation on the extra opening hours per week (%)



Source: Own elaboration

Figure 5.6. Effect of the deregulation on the establishments that open on Sundays (%)



Source: Own elaboration

Finally, figure 5.6 presents the percentage of establishments that open on Sundays. According to the developed MAS the number stabilised in 5%. These establishments tend to be the bigger ones, both in maintenance and discretionary consumption. In addition they are located in the downtown of Madrid and other towns of the region, but also in establishments located in shopping centres. However, the simulations presented a high level of randomness in this decision.

CHAPTER 6

An Agent-Based Model to Forecast the Inflation rate in the Eurozone

6.1. Introduction

Agent-based models are increasingly applied to macroeconomic forecasting. However, they require high amounts of information about agents in terms of microdata and behavioural rules and they have only recently been used successfully thanks to several European Commission funded projects such as Eurace (Cincotti, Raberto & Teglio, 2012) and Mosips (Mancha et al., 2012; Pablo-Martí et al., 2013; 2014).

The aim of these initiatives is to improve the state of the art and they have created massive agent-based models with millions of heterogeneous agents of different classes. The frameworks of these projects are the ones used as building blocks to develop the multiagent system presented in this chapter.

The chapter starts with a background section that presents previous related research in the ABM field. The main flaw of past approaches is the lack of agents' variety with respect to classes (individuals, corporations, financial institutions, etc) and behavioural rules variety. Behavioural rules are a key point to the validity of an agent-based model and in Economics they tend to be selected from Microeconomic Theory. However, since 1970s Macroeconomic Theory incorporates microfoundations. These models have important drawbacks and should be replaced for agent-based models because their assumptions made them unable to predict disruptive changes in the economic growth such as the recent financial crisis (Colander et al., 2008) and also because the equational models have important flaws that can be solved with algorithmic models (Velupillai & Zambelli, 2015). However, they allow explaining macroeconomic variables such as Gross Domestic Product (GDP) growth. Therefore, they can be included in the set of behavioural rules of agents. In the chapter it is explained how it can be performed in an agent-based model aimed to forecast the inflation rate.

The third sections presents three macroeconomic theories that aim to explain inflation performance. They are the quantitative theory of money, the Phillips curve and the aggregated supply and demand as well as the money market equilibrium. In general, they forecast low inflation except in the case of the quantitative theory of money that allows forecasting a high inflation rate in the long run under some assumptions about the behaviour of agents.

After the theoretical review of the main possibilities to explain inflation, the multiagent system is presented following Dahlem guidelines[26] in the fourth section. It makes it possible to arrive to a set of scenarios that incorporate different assumptions to the model in order to assess how they would impact the inflation rate in the Eurozone for the next five years.

The agent-based model is built in order to cope with the main objective of the chapter: to study how behavioural rules impact the level of prices. The most important transactions in the economy are modelled and their results modify the subsequent states and behaviour of agents in the economy. The central bank is explicitly modelled and its decisions are also taken into account. Imports and exports are also incorporated as they modify the performance of companies.

In addition to the actual inflation rate simulated in every period in the baseline scenario that aims to replicate the actual economy, it is possible to introduce shocks in the model in order to measure how agents react to different scenarios. Three conditions are tested: inflationary tensions, deflationary tensions and an oil shock. The agent-based model developed allows the implementation of counterfactual analysis useful for monetary policy.

6.2. Background

Agent-based simulations incorporate individuals and organizations such as companies modelled as agents. They have a set of behavioural rules and interact among themselves. They are heterogeneous and react in different ways to their environment. In addition they can be proactive and take decisions having into account future outcomes they predicted (Klügl & Bazzan, 2012).

[26] Agent-based models and MAS still do not have a widely spread convention to be presented. Among the alternatives it is selected the one agreed during the 100th Dahlem Conference because it covers most of the needs of these models developed in Social Sciences.

Case studies in agent-based simulations have several requirements. For example, they need to couple social and economic models, include micro-level decision making, allow social interaction and adaptive decision-making behaviour (Hare & Deadman, 2004). Agent-based simulations have important advantages to represent economic systems, especially during crisis. However they have also some drawbacks such as the difficulties for their empirical validation and the correct specification of behavioral rules and interrelations still not sufficiently studied.

Complex agent-based macroeconomics is a research field recent that has benefited from the increasing availability of big amounts of available data, new systems that allow computing at a higher speed and the inadequacy of the previous paradigm, dynamic stochastic general equilibrium (DSGE) models. These DSGE models are still used, but the number of researchers in the new approach is increasing thanks to the potential it has.

However, texts such as ‘Complex agent-based macroeconomics: a manifesto for a new paradigm’ (Gatti, Gaffeo & Gallegati, 2010) are still recent. In this manifesto there is an explicit reference to macroeconomic variables such as inflation rate: “Levels and growth rates of aggregate output and employment, inflation […] and all other typical macroeconomic occurrences and laws […] are social phenomena that must be explained at a different level than microeconomic units”. In addition, the authors claim monetary policy can be studied with more accuracy with agent-based models because the supervision of financial institutions is neglected in DSGE models, which assume efficient markets and perfect arbitrage.

Ashraf, Gershman, & Howitt (2013) use an agent-based model to show how inflation impact on the GDP growth by unsettling the mechanisms underlying transactions. In this model it is possible to measure the effect of a rise or a decrease in inflation rate on the economic growth. This model shows the 2% target of the majority of central banks in developed economies is too narrow because deviations of a ±1.5% do not have a clear effect on employment and GDP levels. This model minimizes the cost of high inflation rate and suggests the responsible organizations of monetary policy can relax their targets.

Another agent-based model (Salle, Senegas & Yildizoglu, 2013) intends to quantify the effect of stating an inflation target. The advantages of agent-based modelling on this regards are heterogeneity, bounded rationality and the capacity to learn from the past. Thus, if the monetary authority wants to achieve credibility, the inflation rate has to experience low volatility around the specified target. Agents can modify the expectations about inflation and the degree of confidence in the central bank target accomplishment. If the target it is not reached, an increasing portion of individuals and companies stop including the target inflation in their forecasting mechanism after several periods. In this way, it would be better to have a less explicit and more adaptable inflation rate target.

The study of business cycles is also one of the methodological approaches that can be improved with the use of agent-based models. For example, stabilization policies are usually studied with DSGE models, but agent-based models are increasingly used, as they offer more reliable results. For example, the last fiscal stabilization policies can be included in a model (Harting, 2015) that estimates its impact on companies' behaviour allowing categorizing them by technological level. The impact on every agent is aggregated and this policy changes the aggregated output level in the long run. Another agent-based model that also explicitly focuses on the business cycle intends to measure the impact of consumers' confidence on economic activity (Závacká, 2015). The relations among heterogeneous consumers and producers who anticipate the level of growth of prices and GDP generates business cycle without complicated equational-based models that assume equilibrium takes place in every period.

Finally, there are three noteworthy recent examples of agent-based models aimed to study monetary policy. First, Giri, Riccetti, Russo and Gallegati (2016) research the effect of how interest rate changes produce bubble burst. Therefore, monetary policy decisions can be considered the starting point of the crisis events of 1929 and 2008. In addition, once the recession has started, interest rate is also a key factor in the extension and magnitude of the crisis and zero interest rate can effectively prevent an additional slowdown. Secondly, Gualdi, Tarzia, Zamponi and Bouchaud, (2015) develop an agent-based model that takes into account the potential non-linear effects of monetary policy.

A small change may have little or no effect while an overreaction can produce economic instability. The conclusion of this model contrasts with the smooth effect usually detected with the alternative DSGE models. Lastly, Popoyan, Napoletano and Roventini (2016) study the interactions of macroprudential regulations with other monetary policies. In this agent-based model financial sector is correctly represented and allow the implementation of counterfactual analysis on Basel III new regulations.

The main drawbacks of previous agent-based approaches to the study of inflation are the excessive simplification of the economy and the exclusion in the analysis of current macroeconomic theories widespread in the economic agents. A truly meaningful agent-based approach may include public sector, a key agent that determines taxes, subsidies and represents between one third and a half of the GDP of advances economies. In addition, the central bank is a reactive agent that modifies key variables such as the interest rate and the money supply. Moreover, external sector can be modelled as a whole and interact with companies, which export and import a high portion of their production. Changes in international prices such as the price of oil have a significant impact on GDP growth and inflation and should be taken into account too.

On the other hand, there are macroeconomic theories that explain inflation and previous macroeconomic agent-based models did not include them in their analysis. However, these approaches have microfoundations and have proven useful to forecast inflation. Thus, they can become part of the behavioural rules as agents have them into account when forecasting inflation.

6.3. Macroeconomic theories to explain inflation

In the recent years the average inflation rate in the Euro Area has decreased to a level close to zero, according to Harmonized Index of Consumer Prices (HICP). This rate is below the two percent goal of the ECB. Inflation is defined as the rate (%) at which the general price level of goods and services change, causing purchasing power to fall when it is positive.

To explain inflation or the change of the price level in an economy, different theories exists. We well pay attention to the following:: Quantitative theory of the money, the

Phillips curve, and aggregate supply and aggregate demand and the equilibrium in the money market.

A forecast of the future development of the inflation rate will be provided thanks to an agent-based model that includes the behaviours of the central banker, public sector, individuals, companies and financial institutions. These agents do not only have behavioural rules based on microeconomic theory findings, as it is usual, but they incorporate elements of the three macroeconomic theories most used to explain inflation rates.

6.3.1. Quantitative Theory of Money

The Quantitative Theory of the Money is based on the well-known Equation:

$$M * V = P * T \qquad [6.1]$$

M represents the money supply, V the velocity of the money, P the price level, and T the volume of transactions in the economy.

The money supply can be defined in different liquidity levels in economy. The most liquid form is M0 followed by M1, M2, and M3, each wider definition expands, or includes, the definition of one level below. The money supply is managed by the central bank of the currency area, in the Eurozone by the ECB.

After the international financial crisis of 2007-2008, the ECB started an expansionary monetary policy as it can be seen on the quantitative easing program and the low interest rates monetary policy. Thanks to banking regulations, the ECB can control money supply. The main source of uncertainty emerges from the entities that are not directly by the regulations, the so called ‘shadow bank system’, which includes institutions which are not regulated, for instance asset managers. However, the lack of data makes it extremely difficult to investigate the influences of the ‘shadow bank system’ on inflation. Therefore, only the action of the ECB is observed, and thanks to the ECB, M increases. The growth rate of M3 and M1 has been above GDP in the Eurozone since the financial crisis but the inflation rate decreased.

The velocity of the money has not had a great variation in recent years. The left side of equation 1 has increased in the post crises years, according to ECB policy aim and the actual data offered by the institution. The volume of transactions can be indicated by the GDP. The GDP has grown in the Eurozone since its existence, with the exception of crisis years 2008 and 2009, when there was relatively small a negative growth. Therefore, the first variable of the right side of equation 1 increases. M grew faster than T (GDP). With the assumption of a constant V, it can be followed, that inflation rate should be positive and in the long run at the level of:

$$\frac{1+\Delta \mathrm{M}}{1+\Delta \mathrm{T}\ (\mathrm{GDP})} - 1 = \Delta \mathrm{P}\ (\text{inflation rate}) \qquad [6.2]$$

The left side of equation 2 in the post crisis years was above the inflation rate. Therefore, inflation will rise in the Future according to the quantitative theory of money. This theory has wide evidence in countries of Europe and other advanced economies such as United States, were a recent research (Lanne, Luoto, & Nyberg, 2014) showed that it worked in the long run for the period 1947Q1-2013Q4, including the period of the Great Moderation.

The quantitative theory of money implies that the central bankers are aware of the potential increase of inflation and will implement corrective measures if the objective is exceeded. Agents such as public sector and companies only have recently benefited from the quantitative easing policy in measurable ways, according to their size and indebtedness. The behaviour of these agents is modified. Namely, public sector can increase the expenditure level thanks to the reduction of debt service. In addition, big companies can increase their level of debt following the findings of Modigliani-Miller theorem under realistic assumptions (Joyce, Miles, Scott & Vayanos, 2012). The additional funding can be used to acquire other companies, to reduce the number of their shares in the stock market and to increase their investment only if the expected return on investment is high enough to compensate the uncertainty of a new recession. Consumers are affected by the alteration of the behaviour of companies and public sector as they are public and private employees and they invest their money in government bonds and the stock market. In addition, individuals modify their expectations due to the overall effect of the monetary policy on GDP growth.

6.3.2. The Phillips Curve

The Phillips curve describes the empirical relation observed between unemployment and inflation rate that happens in most of the developed countries during most of the periods analyzed.[27] In summary, it states that inflation tends to reduce unemployment, at least in the short run. Hence, when the ECB increases the money supply and supports inflation, unemployment will be reduced, and vice versa. The economy can buy employment decreasing the value of the currency (price level). This relation is called the sacrifice ratio, and is described as follows:

$$\text{Sacrifice relation} = \frac{\text{points above 'natural' unemployment}}{\text{reduction of inflation rate}} \qquad [6.3]$$

The points above 'natural' unemployment are the difference between the computed normal unemployment rate and the current unemployment rate. The so-called natural unemployment rate is a construct developed during 1970s. In that time it was necessary to adjust the macroeconomic models, because they didn't explain the situation in a sufficient correct way. This rate is explained, for instance, with the time that is needed to find a new job, or the required time to find a sufficient qualified occupation. The reduction of the inflation rate is the difference between the previous and the actual inflation rate. The sacrifice relation is the relation for this both variables and is different for every economy. However, in equation 3, when it is assumed that the sacrifice relation is constant, high unemployment could be compensated with a reduction of inflation, producing a low inflation rate.

Focusing on the development of the unemployment in the Eurozone and comparing the unemployment post and pre financial crisis, it can be seen, that unemployment is much higher post financial crisis. When it is assumed that the average long term

[27] Phillips curve analysis started with the author's seminal paper of 1958. It was later improved and continued by Samuelson and Solow (1960). Its importance grew and it even represented a kind of dogma for its convenience to justify expansionary economic policies during the decade of 1970s in the developed countries.

unemployment is the natural unemployment, the Eurozone has unemployment above this natural unemployment. With a constant sacrifice relation, the inflation rate is lower than pre financial crisis and as long as unemployment is high, inflation rate will be low.

The Phillips curve shows that a reduction of the unemployment level will be interpreted as a signal of following inflation increase by the central bankers, which will modify their policy accordingly even if employment is not a target variable of ECB. In addition, changes in the unemployment rate modify the behaviour of public sector via social expenses and other automatic stabilizers. Changes in monetary and fiscal policies modify subsequently the behaviour of companies and individuals as their income level and expectations are modified.

One of the most important models that explain the inflation with the Phillips curve is Gordon's (1990) triangle model. Inflation in the following period depends on a constant μ, the current inflation π_t, the unemployment rate $u_t$and supply shock variables z_t. $\alpha^G(L)$ is the step-function restriction so the coefficients are equal within the groups of lags 1-4, 5-8, etc., and also that the coefficients sum to one because a unit root is imposed.

$$\pi_{t+1} = \mu + \alpha^G(L)\pi_t + \beta(L)u_t + \gamma(L)z_t \qquad [6.4]$$

6.3.3. Aggregate Supply and Aggregate Demand and the Money Market

The Aggregate Demand and Aggregate Supply describe the relation between the price level (P) and the output of an economy (GDP). The Money Market equilibrium describes the relationship between the Quantity of Money (M) and the interest rate (i). According to this framework the level of prices raises when money markets and goods and services markets interact and at least one of the following conditions takes place:

- Increase in the Money Supply
- Decrease in the Aggregate Demand for Money
- Increase in the Aggregate Demand in goods and services
- Decrease in the Aggregate Supply in goods and services

Inflation can be driven from two sides: cost-push inflation and demand-pull inflation. On the one hand, the cost-push inflation takes place when the costs of production factors increase. In this way changes in the amount of raw materials such as oil, modify the equilibrium of supply and demand and affect the inflation through the agents of company class. In addition, in the last years Europe had many structural changes affecting labour factor in order to make it flexible the economy and boost the competitiveness, or in other words, reduce costs. When total income of the economy between the income distribution among capital income and income from work during the last 20 years is compared, it can be seen the change is in favour of capital income. Therefore, it can be followed, that the cost moderation supported the low inflation rates in the last years.

Capital effect can be seen from two perspectives, the asset in the statement of financial positions of the firms or as the financial perspective as equity and liability in of the statements of financial positions. Taking a look on the asset side, firms' balance sheets have a higher value than in the past and that means the costs for machinery are higher. This can be explained by inflation itself, because inflation drives prices of assets and machinery and by the new accounting methods. On the asset side, it is necessary to highlight that technologic development makes assets more efficient, therefore this yields in lower production costs. Analyzing equity and liability, that is, looking at the costs for an asset from the financial perspective, companies pay interests for its financing, interests for liabilities and costs of equity. Money supply expanded and costs of capital lowered. According to data for recent years, expansive monetary policy has lowered costs of capital, and therefore, the costs of holding assets. Accordingly, raw materials and, in particular commodities, price increases positively affect inflation and vice versa. However, it may depend more on supply and demand of the products defined by agents of the class companies.

S&P GCSI Index makes it possible to conclude that after its crash commodities did not recover to the previous prices. In particular oil prices recovered partially for a few years, but thanks to the increase of supply and the consumption reduction, decreased again. Therefore, raw materials, and particularly commodities and the oil, are cheaper post crisis and low inflation rates are supported as a result. To sum up, as described in this section, the indicators of cost-push inflation support most likely, with exception of the capital component, a low inflation rate as they modify the behaviour of companies.

The demand-pull inflation occurs when aggregated demand increases. Aggregated demand is categorized by individuals, companies and public sector. According to individuals, thanks to high unemployment and the structural changes, aggregated demand decreases. What may increase aggregated demand is the low level of interest according to monetary policy on savings. For this class of agents, there is an incentive to spending instead of saving. Moreover, thanks to the low inflation rate, capital income (from equity and debt) should be lower post crisis. According to the public sector, when the data of pre and post crisis government spending is compared, government spending is on a lower level in most countries. Therefore, the demand of this agent supports a low level of inflation too. Finally, according to companies, when the public sector spends less and individuals spend less because of high unemployment, they should have less revenue and therefore also less expenses, in particular, in service-based industries, where costs can be reduced at a very fast rate. Even if the external sector demands more exports due to an increase in competitiveness, the behaviour of agents of the class companies also contribute to arrive and maintain low inflation rate.

As aforementioned the supply of the money increased according to the expansionary monetary policy. This supports higher inflation rate and indeed it has been performed aiming an effect on the real economy. Therefore, interest rate decreased and if the demand for money is on a low level, this would support higher inflation rates additionally.

It is possible to consider that the aggregate supply and aggregate demand support mainly a low inflation rate and the money market supports a higher inflation rate. As it can be seen on the HICP, the inflation rate is close to zero in the recent years in the Eurozone. An increase of domestic demand could stimulate the economy and increase inflation in the middle run. In the short term, a correct model would preliminary forecast a low inflation rate.

6.4. An agent-based Model to forecast the inflation rate

The agent-based model is constructed incorporating building-blocks of three finished public-funded projects:

- Eurace project (FP6 European Commission) incorporated parallel computing for macroeconomic policy design that included real economy, financial economy and the public sector (Deissenberg, Van der Hoog & Dawid, 2010;Cincotti, Raberto & Teglio, 2012).
- Mosips project (FP7 European Commission) focused on the real economy and public sector and expanded the behavioral rules of individuals and companies (Mancha et al., 2012; Pablo-Martí et al., 2013; 2014).
- ‘Analysis of systemic risk after the crisis emergence’ (NCN Poland) focused on macroprudential policy and interbank markets (Kaszowska & Santos, 2014).[28]

Java repast is used for its implementation because it allows parallel execution of different task by agents of the same class and different classes and reduces the computation time in comparison with other alternatives. The model includes the approximate number of agents in any category: around 340 million individual, 22 millions companies and nearly 7500 financial institutions. Public administrations, central bank and rest of the world have only one agent in each category.

The next depiction of the model follows Dahlem guidelines for describing economic agent-based models (Wolf et al., 2011) that were developed after a high number of researchers agreed it was necessary to have a standard to describe complex agent-based models following an approach that allows the readers to understand the main insight of the model focusing in the underlying concepts rather than in an enumeration of algorithms or equations.

[28] This work supplements the aspects not included in Eurace and Mosips projects related to interbank markets.

6.4.1. Rationale

The object under consideration is the economy of a monetary union, in this case the Eurozone. They are considered goods and services markets, interbank markets and the government bonds primary market. The economy depicted is an open economy that transacts goods and services with the rest of the world modelled as a whole.

The intended usages of the model are forecasting of inflation and secondly, monetary policy analysis. The time period used is a quarter and agents explicitly forecast inflation for the next period taking into account their situation. It is possible to introduce different policies of the ECB and check their effectiveness. In addition ECB behaves reacting to changing conditions of the economy and especially the financial institutions performance.

The model predicts also the levels of unemployment, interest rate, GDP growth, public deficit and public debt, sector growth rate, non-performing loan rate of financial institutions, exports and imports. Nevertheless these variables are not the focus of the model, the algorithms that affect them have structural validity but the process of calibration did not intend to mimic the actual behavioural rules accurately and therefore their values in the model cannot be taken as actual forecasts.

6.4.2. Agents and other entities

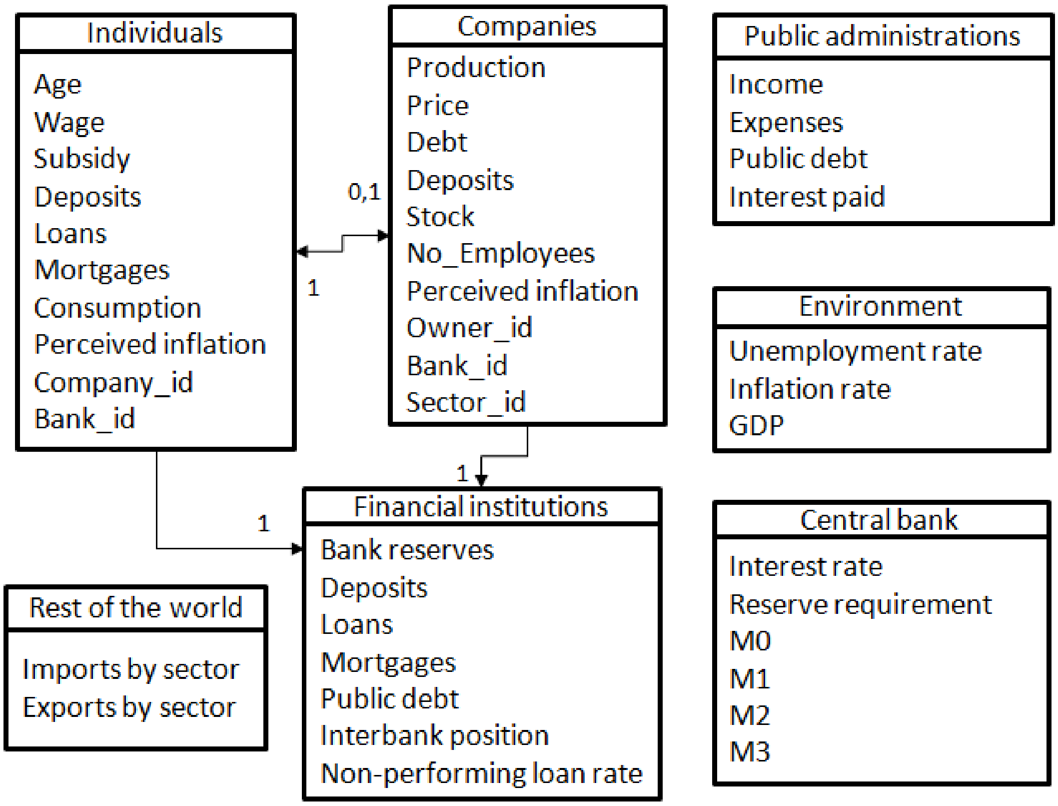
There are three kinds of heterogeneous agents in the model: individuals, companies and financial institutions. In addition there is a central bank, public administrations and external sector. The database is built from 2008Q1 in order to use the first periods of one quarter of duration to perform the calibration and validation according to the double mechanism described in Klügl (2008). First the model was tested for the internal structural validity, the relations between variables and agents. In a second stage, the behavioural validity is tested trough the results both at aggregate and agent level. This approach allows obtaining a model that is valid from a technical point of view and arrives in all the simulations to plausible results. In addition, the first periods are used for calibration and at the same time the behaviours of agents are checked in order to obtain a valid model.

Individuals have age expressed in years, which they increase by 0.25 each period. Individuals can be employed or not. If they are employed they have a company identifier code associated to them and a positive wage. If they are not employed they can have a positive subsidy received from the public administrations that represent retirement and disability subsidies and wages of public employees. Individuals can have positive amounts of deposits, loans and mortgages in a financial institution included in the model, denoted by bank identifier code. Individuals devote their income to taxes, savings and consumption in companies of different sectors that are included in the database. According to the paid prices they have their own perceived inflation in each period.

Companies produce goods or services each period and assign a price for them. Companies also have deposits and loans with a financial institution of the model, designated as bank identifier code. Their production is added to the inventory, that it is updated after each period, decreasing with the amount that was sold. Companies need employees to produce, and they hire and fire them every period. They also have an owner who is an individual in the database, identified by an individual identifier code. He or she receives profits if the company is successful and incorporates them to their deposits. Companies can be in one of the different sectors. They face demand from consumers, public administrations, other companies and the rest of the world (exports). They need raw inputs for produce their goods and services from other companies in the data base (supply chain) and the rest of the world (imports).They pay profit taxes and the Value Added Tax (VAT) to the public administrations. They also compute their perceived inflation rate in each period according to the change in the prices of inputs.

The last class of agent, financial institutions, holds deposits, loans and mortgages of individuals and companies of the data base. They pay interest for deposits and receive income from loans and mortgages if companies and individuals can afford to pay them back. If agents with debts are bankrupted their liabilities are included. They also buy government bonds issued by public administrations and receive interest for it. They hold bank reserves in the ECB and participate in the interbank markets where they lend their excess of liquidity to other financial institutions with liquidity shortage.

Figure 6.1. Agents and variables in the model



Source: Own elaboration

There are three other entities: public administrations, central bank and the rest of the world:

- Public administrations collect taxes, purchase goods and services from companies, pay interest rate for the previously issued government bonds and give subsidies to individuals. During periods of low economic growth they tend to have deficit and they need to issue large amounts of public debt.
- The central bank fixes the interest rate and controls different definitions of money supply. They can offer or remove liquidity in the interbank market in order to manage inflation and guarantee the financial institutions are sufficiently stable.
- The rest of the world represents the rest of the economies. It only interchanges goods and services with companies and there are not financial relations.

6.4.3. Boundaries

There are several boundaries in the agent-based model described in this chapter. Most of the models tend to hide them but it is highly advisable to state them in order to help identifying potential sources of inaccuracies and get rid of these boundaries in subsequent versions of the model or approaches by other researches. The first three ones are partially included in Pablo-Martí, Santos and Kaszowska (2015) that focuses on individuals. Most of the boundaries are solved in Mosips and Eurace projects, but in this case they would make the model too complex for the intended aims with no expected significant change. Others are highlighted in the section 6.6 and would need to be included in the system for achieving an optimum forecast:

- The number of individuals and companies is fixed: there are not newborns, deceases, emigration, immigration and firm dynamics (bankruptcies, mergers and acquisitions, new firms). This boundary only has limited impact in the short term but it prevents using the model for long-term forecasting.
- Individuals decide about saving and consumption grouped in households but this class of agents is not included in the model to reduce the complexity. Households would require specific activities such as marriages and divorce, real estate investment and mobility to represent the real population dynamics.
- Individuals do not have gender, educational level, nationality, location and other characteristics that help to determine their labor status and financial position.
- Companies do not have product specialization, research and development activities and their financial statement and balance sheet is oversimplified.
- There is no stock market, cooperatives, and financial institutions are the only source of company funding excluding other mechanisms such as public loans, grants and business angels.
- Sectors are simplified into five: agriculture, manufactures, construction, business services and personal services.
- The real estate sector is not modelled and real estate bubbles and bursts are not included in the system. Thus, demand for the construction sector experience a lower variation than in the actual economy.

- Companies only can hire and fire employees taking into account age and labour status. Companies and individuals have no location and this model cannot study different economic growth in European regions, which would lead to different inflation rates.
- When companies have difficulties repaying their loans, they reduce production but there are not bankruptcies and potential sector crises can be minimized.
- Financial institutions always decide to participate in interbank markets and concede loans and mortgages if companies and individuals are sufficiently solvent. This behaviour does not take into account situations of systemic risk when banks prefer to leave the excess of liquidity in the central bank rather than lend it to other financial institutions. In those conditions, loans to individuals and companies are also significantly reduced as it happened during the some years after the recent financial crisis in Europe.
- Public administrations do not react to changes in the ruling parties and adapt their expense level to the political business cycle.
- All public administrations are aggregated into one agent and there are not different risk premiums for government bonds issued by countries incorporated in the monetary union.
- Central bank is independent and only cares about inflation rate and solvency of financial institutions under its supervision. In the case of ECB there are other objectives that are not strictly economic and are related to political concerns about Euro stability.
- The rest of the world acts in the baseline scenario as a steady object that only changes exports and imports in relation to changes in prices in the Eurozone, as a proxy of productivity. However, in the actual economy there are different countries that can adjust their local currency to be more competitive, they can experience diverse levels of growth and there are also trade agreements that might change imports and exports between commercial areas.

6.4.4. Relations

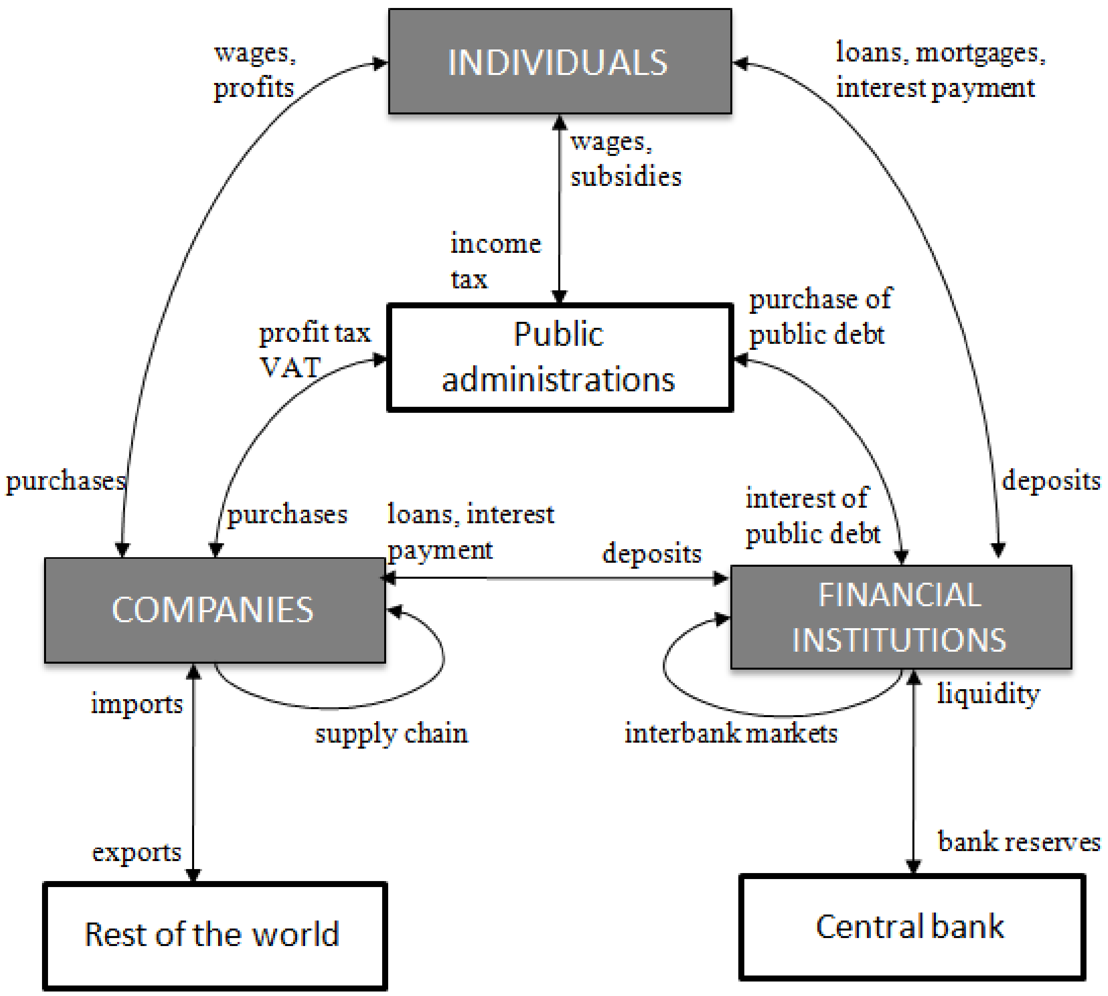
All the individuals of Eurozone are included in the system according to data from Eurostat. They have deposits in one and only one financial institution, where they also might have loans and mortgages. They can have a wage from one and only one company or a subsidy. In any case they pay income taxes. They decide which portion of their net income it is spent and which part it is saved according to the life-cycle model of consumption and saving (Browning & Crossley, 2001) taking into account their net assets, their labour status and their age. By having deposits, individuals receive an interest rate from financial institutions (i) lower than the official one and they have an expectation about the inflation rate for next period (π_{t+1}^{agent}). The real return (r) as it appears in Fisher equation [5] can be different for every agent and it also conditions their decision about saving in addition to the standard life-cycle model.

$$i = r + \pi_{t+1}^{agent} \qquad [6.5]$$

Companies are modelled according to information obtained from the database Amadeus that compiles information from public records. Companies need raw materials to produce and they are bought from other companies in the rest of the world (imports) and from other companies in the Eurozone according to the available data in the input-output matrix. In the case of positive profits they pay profit tax to public administrations and the owner of the company increases the deposits.

Financial institutions hold the deposits of individuals and companies paying an interest rate inferior to the one stated by the central bank and borrow money to these agents, asking for an interest rate superior to the official one. They need to deposit the statutory reserve requirement in the central bank and they buy the public debt issued by the public administrations. Finally, they participate in the interbank markets where the banks with liquidity surplus lend money to other with liquidity deficit, asking an interest rate slightly superior than the official one set by the central bank.

Figure 6.2. Agents and transactions in the model



Source: Own elaboration

6.4.5. Activities

Individuals' consumption is computed using fuzzy rationality. They consume a higher portion in companies with the lowest prices and those with a higher number of employees, as bigger companies have higher visibility and consumers only purchase in companies they have heard about before. Companies also decide their suppliers taking into account price and firm size. Bigger companies have access to more information as they have employees only devoted to select suppliers and they decide with a higher degree of rationality.

If individuals' income is below a threshold and their amount of deposits is zero they are not able to repay loans (short term) and mortgages (long-term). Financial institutions ask for different interest rate depending on the duration of the loan, as the interest rate in mortgages is lower because they have an asset in return. When a mortgage is not paid the bank cancels one half and the other half is included as a non-performing loan. This process represents an eviction. In the case of unpaid loans there are fully included as non-performing loans.

When company's inventories are below a threshold they increase their number of employees and they decrease their staff when the inventory of products is higher than another threshold. They prefer to hire and fire young employees but they also apply fuzzy rationality and sometimes choose a person near the retirement age.

Companies (c) update wages and public sector updates subsidies using the inflation on the last year (four periods). Companies decrease the amount spent in raw materials, wages and interest rate of the loans to deposits. They update their price in every period according to three determinants: previous experienced inflation (π_t^c), relative amount of inventory and inflation expectation (π_{t+1}^c):

$$\Delta \mathrm{price}_{c,t} = \pi_t^c \left(\frac{\alpha\ \mathrm{stock}_c/\mathrm{production}_c}{\sum \mathrm{stock}_c/\mathrm{production}_c}\right) \pi_{t+1}^c \qquad [6.6]$$

In this way, the most successful companies tend to increase their prices more than the ones with a higher amount of inventories (α>0) but this change depends also on their costs and their expectations about inflation that can differ to a great extend.

Individuals and companies forecast inflation according to the three theories presented in the previous sections, assigning to them different weights θ:

$$\pi_{t+1}^{agent} = \theta_1 \left(\frac{1+\Delta M_t}{1+\Delta GDP_t} - 1\right) + \theta_2\left(\mu + \alpha^G(L)\pi_t + \beta(L)u_t + \gamma(L)z_t\right) + \theta_3 \left(\rho e_d \frac{\Delta GDP_t}{\Delta \sum \mathrm{stock}_t}\right) \quad [6.7]$$

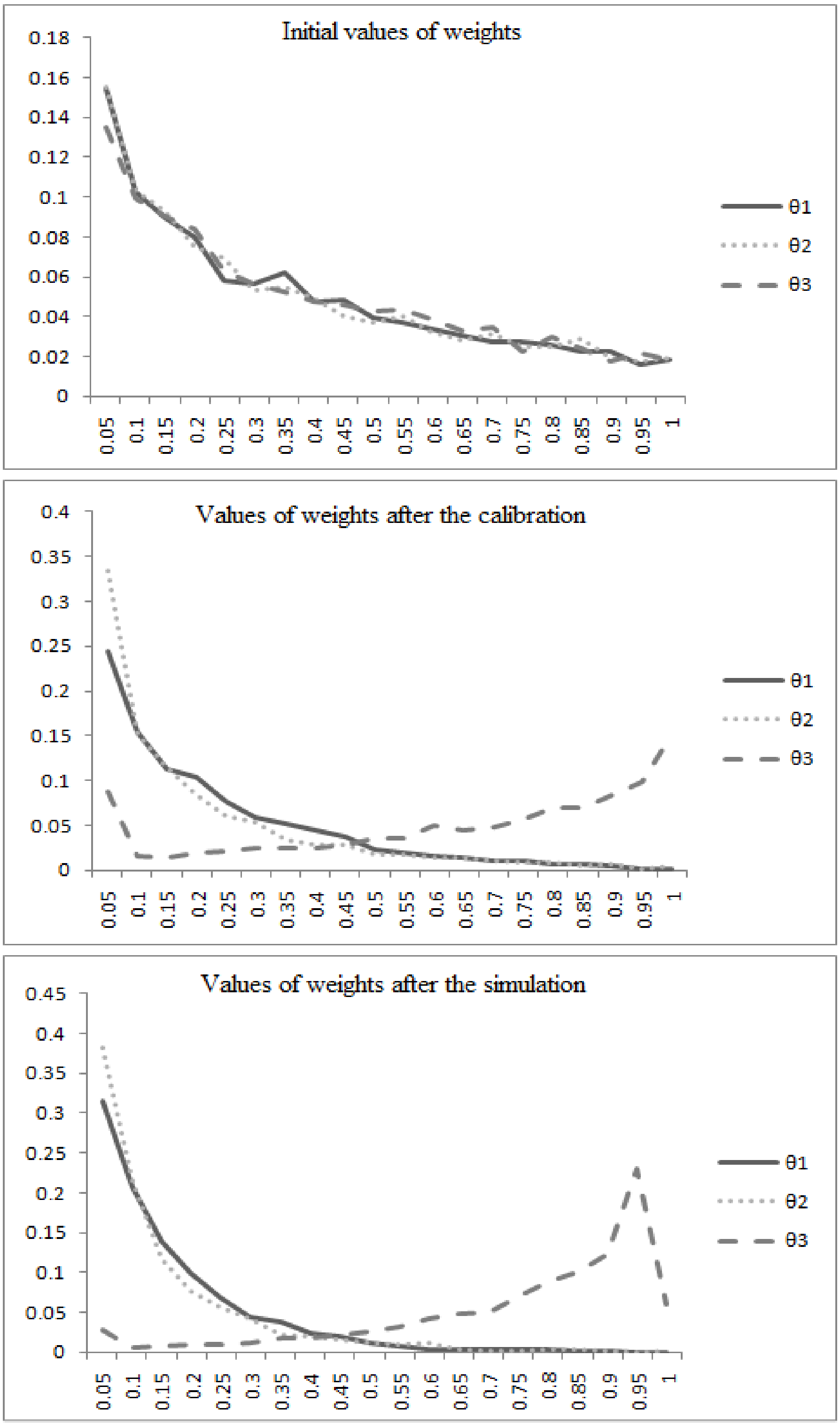
The coefficients θ are the percentage of likelihood they assign to each theory. They are between zero and one and sum the unit. The first one is related to the Quantitative Theory of Money (eq. 2), the second one to the Phillips curve (eq.4) and the third one to the Aggregated Supply and Aggregated Demand equilibrium and is equal to a constant multiplied by the price-demand elasticity and the relation between the variations of output and productive capacity.

During the calibration stage, every agent of the company and individual classes starts with the following coefficients θ, selected from uniform distributions:

$$\text{Select } rndm\ \mathrm{i,j,k} = \{1,2,3\}.\ \theta_\mathrm{i} = \mathrm{u}(0,1).\ \theta_\mathrm{j} = \mathrm{u}(0,1-\theta_\mathrm{i}).\ \theta_\mathrm{k} = \mathrm{u}\left(0,1-\theta_\mathrm{i}-\theta_\mathrm{j}\right)\ [6.8]$$

They compare their previous inflation expectation, the three ones obtained with $\theta_\mathrm{i} = 1$ and the one they experience in the current period. They modify the weights increasing the one(s) that proved to be more accurate by max(1- θ_i, $0.01\theta_\mathrm{i}$) and reducing the other(s) in order to maintain the sum of the three ones equal to the unit. This procedure is developed in a high number of periods until it converges to the actual aggregate outcome in order to keep enough amount of heterogeneity and assure a sufficient degree of rationality to the majority of agents. This procedure is also implemented in the agent-based model after the calibration has taken place. In figure 3 it is possible to see the initial values, those after the calibration and the ones after twenty periods in the simulation.

Figure 6.3. Calibration and simulation of weights of macroeconomic theories for inflation expectations



Source: Own elaboration

The initial values of the three coefficients are in average around to one third. Their distributions are also very similar and they have a higher amount of lower values and the median locates near one fourth. In this way a positive portion of agents starts with one of the three weights close to one and if it was the optimal macroeconomic theory to forecast inflation in the agent-based model the weight would be one unit after the calibration.

Table 6.1. Values of theta coefficients in different stages of the calibration process

Coefficient	Initial value	Value after calibration (2011Q1)	Value after simulation (2015Q4)
θ_1 - Quantitative Theory of Money	0.33	0.23	0.18
θ_2 - Phillips curve	0.33	0.2	0.16
θ_3 - Aggregated Supply and Aggregated Demand	0.33	0.67	0.76

Source: Own elaboration

The process before the simulation takes place is the initial calibration of the coefficients. Agents need an initial value for every coefficient in the equations so they can develop their decision-making processes correctly. In order to make this calibration, the first period simulated takes place a high number of times until the aggregate outcome is similar to the one of the second period. In this way, agents modify their coefficients until the simulation has a realistic performance. In this case, the first period is 2011Q1.

Once the simulation for the calibration takes place, in the initial period one third of agents assign a value equal to zero to inflation forecast based on Phillips curve, one fourth of agents assign a value below to 0.05 to inflation forecast based on Quantitative Theory of Money and only one twelfth do not rely on the Aggregated Supply and Demand equilibrium. In addition, less than 0.5% assign θ_1 or θ_2 a value higher than 0.95, while θ_3 is close to the unit for one out of six the agents. The average values are $\theta_1 = 0.23$; $\theta_2 = 0.2$; $\theta_3 = 0.67$.

Once the actual simulation from 2011Q1 to 2015Q4 takes place, the likelihood assigned to each theory is slightly modified. The percentage of agents assigning a null or small weight to the first two macroeconomic theories increases to 30% and 40% respectively. It is worth noting that the third theory decrease its mode to the range from 0.9 to 0.95. This can indicate the other two theories might have some advantages for a number of companies and individuals when they are have a low but positive weight. After the simulation the averages values are $\theta_1 = 0.18$; $\theta_2 = 0.16$; $\theta_3 = 0.76$.

Finally, inflation rate in the environment object of the database is updated as:

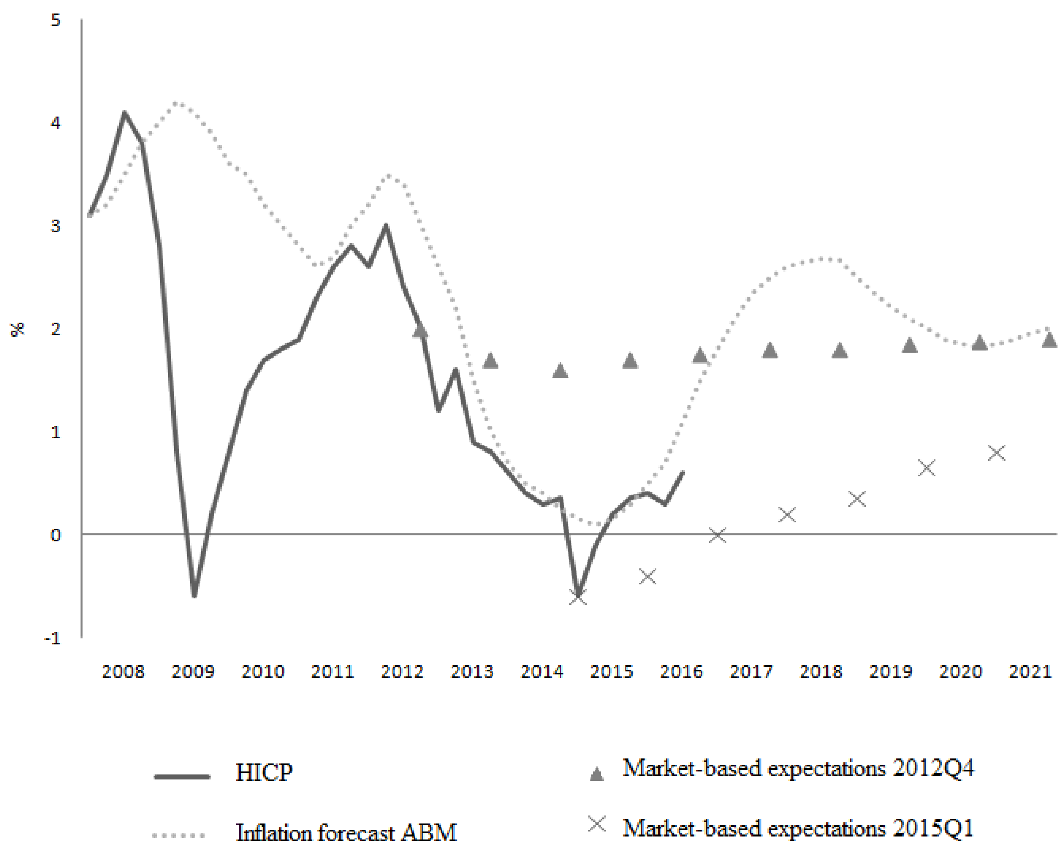
$$\pi_t = \sum \Delta price_{c,t}\left(production_{c,t} - \Delta stock_{c,t}\right) \qquad [9]$$

In this way, every company decision of updating prices is weighted for its demand in the current period, expressed as the difference between the amount produced and the change in inventory. In the case a company has no demand all the production would be stockpiled and its contribution to the inflation rate calculation of that period would be null.

6.5. Main results

The agent-based model presented in the last section allows forecasting not only inflation and the average inflation expectation for the next quarter. In the baseline scenario the actual data for all the variables and calibrated coefficients is included in the model in order to simulate the inflation rate for the period 2008Q1 to 2021Q4. Next, there are three scenarios introduced: Oil shock where the price of this commodity would increase to 150$ in the period 2016Q1 to 2021Q4. Second, inflationary tensions are incorporated in the model with an increase in public expenditure of 5% and increase in exports of 5% for the period 2016Q1 to 2021Q4. Last, deflationary tensions with an equivalent decrease in public expenditure and exports in these five years.

Figure 6.4. Inflation forecast of the agent-based model from 2008 to 2021



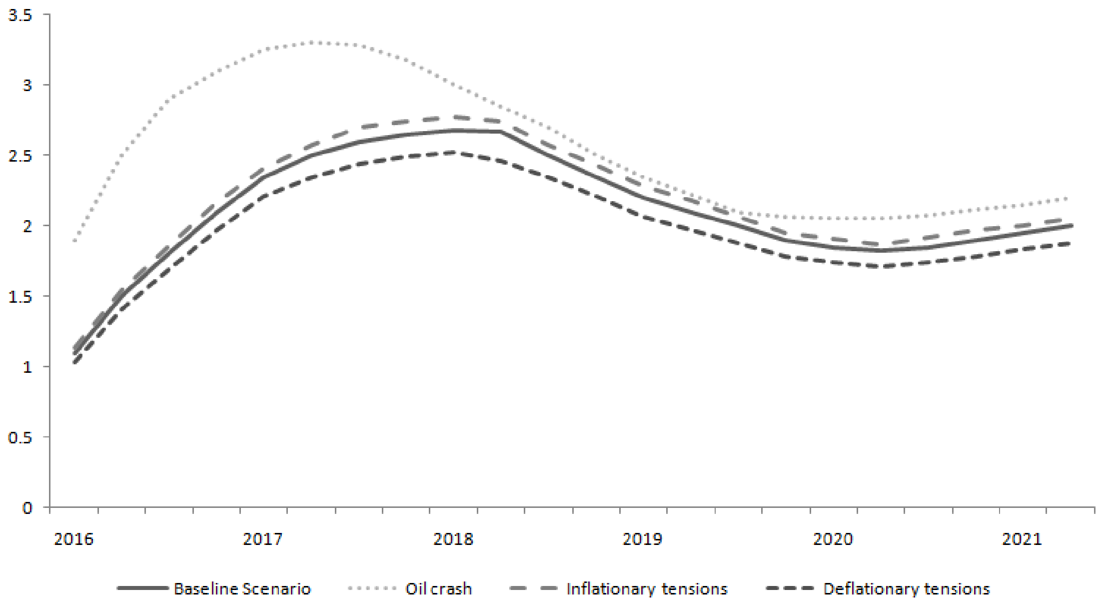
Source: European Central Bank and own elaboration

Note: Market-based inflation expectations refer to zero-coupon swaps of 1 to 10 years

The main difficulty the model with respect to previous inflation performance took place during 2009 and 2010. During this period the austerity of public administrations reduced significantly aggregated demand. In the agent-based model fiscal policy variations do not take place as they are the result of political decisions rather than the application of optimal behavioural rules of the government. After 2011 the model fits better the inflation rate of the Eurozone and it forecast a low inflation rate for the period 2016 to 2021. In the future, ceteris paribus, the model forecast an inflation increase until 2018Q1, when the inflation rate would be 2.68%. From that point it would slightly decrease and would stay near the 2% target of the ECB. As it can also be seen in figure 4, market-based expectations in 2012 estimated a constant inflation near the target, but in the beginning of 2015 the low price of oil and the uncertainty about economic growth anticipated several years of inflation rate below 1%.

An oil shock, as defined by Hamilton (2003) is the non linear relation between oil price and macroeconomic performance. When the oil price changes, suddenly GDP growth and other macroeconomic variables such as inflation and unemployment rate vary in the following periods and their changes can exceed the one in oil price that set off the dynamics. In addition, oil shocks tend to be higher when oil prices increase. In the simulated scenario the price of oil barrel increase to 150$ in 2016Q1 and continues with that price during the remaining periods of the simulation. The inflation rate increases up to 3.3% in 2017Q2, and has a slightly higher rate than the baseline scenario for the rest of the simulation. Nevertheless, the forecasted automatic adjustment in the monetary policy implemented by the ECB allows arriving to a value close to the target of 2% the inflation rate.

Figure 6.5. Inflation forecasts of the agent-based model from 2016 to 2021



Source: Own elaboration

Finally, deflationary and inflationary tensions initiated with 5% variation in public expenditure and exports have little impact on inflation in the agent-based model. On the one hand, currently companies have surplus capacity and an increase in demand will not necessarily generate a higher inflation rate. On the other hand, an additional decrease in demand would trigger additional expansionary monetary policy from the ECB and the negative effect would be partially corrected.

6.6. Conclusions and future research directions

The implemented agent-based model intends to forecast inflation rate in the Eurozone and compute market-based expectation. Although it is a promising research field, most previous attempts did not use real data. In this case the model uses actual data for individuals, companies and financial institutions of the Eurozone but still there are some assumptions that should be enhanced in order to improve the accuracy. Once this objective is achieved, the period of the simulation can be decreased to one month instead of one quarter and the model can introduce seasonality in the inflation rate.

In addition, it would be desirable if a future version of the model could generate macroeconomic variations such as oil shocks and other inflationary and deflationary tensions. In order to implement this feature it would be necessary to include regional financial markets and to model the interrelations among agents located in different countries of the Eurozone. Companies, individuals, financial institutions and the public sector in several countries face uncertainties and problems that other areas currently do not have.

The classification of agents in countries would make the behavioral rules more complex but the model would be able to forecast not only the average inflation rate in the Eurozone but the one in each country. Another direction for future research related to the previous point is the strength of the international financial aspects of the models. The forecast would benefit from the incorporation of important variables such as foreign direct investment, foreign ownership of companies, money transfers between countries and remittances. These variables might in the current version of the model are not explicitly included, and therefore their value is constant and null. The inflation forecast would change if they were explicitly incorporated and modelled, particularly at national level and presumably in a lower amount in the average of the Eurozone.

The results of the model would improve if interbank markets were modelled in order to allow systemic risk. This would increase the difficulty of ECB supervision but the effects on inflation in the short and long run could be significant. In addition, this improvement would benefit from the modelling of real estate market in order to link construction activities with credit market of households in a realistic way. However, institutional agents always react in an appropriate manner with anti-cyclical policies. This might not always be the case but it would be difficult to incorporate to the model the correct ideological bias in a democratic union of countries such as the Eurozone, currently formed by 19 countries.

Finally, this agent-based model shows it is possible and convenient to include standard macroeconomic models in the behavioural rules of agents when their use is extended and the only alternatives would be the estimation with time series or the assumption of the same value of the variable as in the previous period.

Previous attempts to forecast inflation with agent-based models are recent and they did not include macroeconomic theories with microfoundations in agents' behavioural rules. This can improve predictions and make models more realistic in terms of decision-making processes and agents' learning.

A meaningful agent-based model to study inflation has to model individuals, companies, financial institutions, public sector and the central bank in order to simulate both the real and financial economy and their interrelations. The central bank has to implement the monetary policy adjusting its action to changes in inflation but also taking into account economic growth and potential problems in financial institutions.

Public administrations have a key and sometimes neglected impact in the real economy, purchasing goods and services, collecting taxes and providing subsidies. In addition they fund public deficit issuing government bonds, which are one importance part of financial transactions. Finally, the incorporation of the rest of the world in the model is highly advisable, not only for the importance of imports and exports of services and manufactures goods, but especially because importation of commodities in advanced economies has a great impact on the macroeconomic performance.

The developed model allows agents to have into account the forecasts with the Quantitative Theory of Money, the Phillips curve and the Aggregated Supply and

Aggregated Demand equilibrium. During the period 2008 to 2015 the latest one is the most useful for the majority of agents but the three are incorporated to the forecast of inflation in a significant portion of agents.

Currently, the Eurozone has low inflation and in the simulations the inflation rate will remain low, increasing until 2018 and exceeding the target of 2%. The simulation is performed with no unforeseen change in issues such as public expenditure, financial stability, foreign trade and price of commodities. The model was able to replicate the inflation rate in the Eurozone for the period 2011 to 2015. However, contractive fiscal policies in the previous years were not taken into account and the inflation rate forecasted was higher than the actual one. Nevertheless, this is not a drawback of the model and can be seen as a counterfactual analysis studying the case of the macroeconomic performance without discretionary actions of European governments during the period immediately after the recent financial crisis.

CHAPTER 7

Final remarks

This Ph.D. Thesis tries to give solutions for two current problems in Economics that are threats as well as opportunities. On the one hand, we must keep in mind the utter failure of traditional models to predict the crisis started in 2008. On the other hand, the challenges that economic modelling face with the increasing availability of big data sources should be also taken into account. Economists still have not resolved the old controversy between the social scientists who hold that the methods of natural sciences should incorporate to this social science and those that think that the nature and the society are different and therefore the research tools must be diametrically different.

This work clearly defends the necessity of using multiagent systems as natural sciences do with the help of the last developments in systems engineering that make it possible to cope with large amounts of data in complex models.

Economics has increasingly used sophisticated mathematical tools with different results but still refuses the usage of other techniques such agent-based modelling. Macroeconomic equilibrium models are aimed to forecast but in most of the cases they only produce a single and incorrect result. The time to incorporate tools from Physics and Computer Engineering has arrived and economists have to adapt or perish. With the tools presented in this Ph.D. Thesis it is possible to rebuild the work of macroeconomists including evolutionary and institutional points of view in order to achieve more useful results to policymakers, businesspeople, academics and society.

The economic theory has been developing different tools to assess and simulate the impact of policies according to the dominant theories in every moment. Currently, many central banks and governments use Computable General Equilibrium models to forecast the effect of their decisions and choose among alternatives. The unpredictability of the crisis by these models and its narrow contribution to the spatial impact of policies or the income distribution, make this kind of model a lacklustre tool to evaluate and simulate policies. Spatial econometric models appear to be an alternative to computable general equilibriums but they have other drawbacks: they do not resist the Lucas critique and their specification is too rigid to incorporate the real performance of economic activity. Industrial Organization and Business Theory research about companies but they put the focus on game theory, again a tool that aims for equilibrium, and neglect the dynamics of the markets they study. In the case of Labour Economics workers and firms can be modelled and overcome the partial focus of researches in this field, usually focused on employees or employers. It is possible to find examples of how multiagent systems can be incorporated in all areas of Economics and Business Administration in order to solve unanswered questions. In fact, an increasing number of researchers are in the fields of Evolutionary Economics, Artificial Economics, Socioeconomic Engineering, Complex Social Sciences and other fields in which multiagent systems are considered one of the main approaches in the toolbox to represent the economic reality.

In conclusion, the main ideas of the five chapters presenting applications can be summarized as follow:

i. Multiagent systems are capable to model accurately population dynamics correcting some of the problems of alternative estimation methods. (Ch. 2)
ii. It is possible to combine real agents such as individuals with artificial functions that represent the interactions with other agents (Ch. 2)
iii. MAS can be compartmentalized in a set of modules in order to make a flexible and adaptable tool. (Ch. 2; Ch. 3)
iv. Multiagent systems can forecast aggregate variables such as market concentration or inflation using a bottom-up approach. (Ch. 3; Ch. 6)
v. Market power in traditional manufactures in Spain decreases during the crises and increases or remain equal during periods of economic growth. (Ch. 3)
vi. MAS are a better tool than game theory because they not only focus on the potential equilibrium but also on the dynamics of the decision process and interaction between agents. (Ch. 4)
vii. MAS can incorporate contagion, free-riding and coordination mechanisms without explicitly modelling them. (Ch. 4)
viii. Multiagent systems can predict the effects of economic policies with a high level of sector and spatial disaggregation (Ch. 5)
ix. Spatio-temporal databases are among the most promising tools to model the real economy and achieve results in time periods less than 24 hours. (Ch. 5)
x. Trading hours deregulation in Madrid contributed to increasing the average

number of employees in shops. Therefore it negatively affected the smallest establishments (Ch. 5)

xi. Multiagent systems can incorporate different macroeconomic theories and test their effectiveness in realistic scenarios. (Ch. 6)

xii. The Phillips curve and the Quantitative Theory of Money do not show results as good as the ones of the aggregate supply and demand equilibrium in the Eurozone during the period 2008-2016 (Ch. 6)

As stated in the starting quote by the philosopher of science Karl Popper, whenever a theory appears as the only possible one it is a sign of the lack of understanding of the theory and the problem intended to be solved. This is what happened to some macroeconomists with dynamic stochastic general equilibrium models. However, most economists using the agent-based computational economics are willing to improve their models including all the possible techniques (Dilaver, Jump & Levine, 2016), including DSGE modelling.

In this way, recalling the famous quote by G.E.P. Box, “all models are wrong, but some are useful” and as I try to highlight, multiagent systems and agent-based models have the potential to be extraordinarily useful in Economics and Business Administration.

One of the main problems of agent-based models is the criticism due to lack of understanding and academic territorialism. “Your model is too simple”, “Your model is too complex” and “your model is a black box” are three of the main responses scientists in social simulation reported receiving in a recent survey (Waldherr &. Wijermans, 2013). Therefore, the quote by Paul Valéry is especially important to multiagent systems because “everything simple is false. Everything which is complex is unusable”. In the elaboration of the five applications included in this Thesis I always kept in mind the concept KISS (Keep It Sophisticatedly Simple) and I tried to build models and systems usable, adaptable and simple enough. However, human behaviour most of the times proves to be illogical and we will never arrive to a model capable of perfectly reproducing it.

References

Aiginger, K., & Finsinger, J. (Eds.). (2013). *Applied industrial organization: Towards a theory-based empirical industrial organization*. Springer Science & Business Media.

Akerlof, G.A. and J.L. Yellen. (1985). A Near-Rational Model of the Business Cycle with Wage and Price Inertia. *Quarterly Journal of Economics*, 100, 823-838.

Anderson, R. (2001). Why information security is hard-an economic perspective. In *Computer Security Applications Conference. ACSAC 17th Proceedings* (pp. 358-365). IEEE.

Anderson, R., & Moore, T. (2006). The Economics of Information Security, *Science*, 314(5799), 610-613.

Anderson, R., & Moore, T. (2009). Information security: where computer science, economics and psychology meet. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1898), 2717-2727.

Aranda García, E. & Santos del Cerro, J. (2012). Aproximación analítica y empírica a los horarios comerciales. *Información Comercial Española* 868, 31-56.

Archibugi, D., Filippetti, A., & Frenz, M. (2013). Economic crisis and innovation: Is destruction prevailing over accumulation?. *Research Policy*, 42(2), 303-314.

Arentze, T. A., & Timmermans, H. J. P. (2001). Deriving performance indicators from models of multipurpose shopping behavior. *Journal of Retailing and Consumer Services*, 8(6), 325-334.

Arrow, K.J. and G. Debreu. (1954). The existence of an equilibrium for a competitive economy. *Econometrica*, 22(3), 265-290.

Asensio, J. (2013). Regulación comercial regional y entrada de supermercados en España. In Merino-Castelló, A. (Coord.) *Regulación y competencia: en busca de efectos* *no deseados* (pp. 115-126). Autoridad Catalana de la Competencia, Generalitat de Catalunya.

Ashraf, Q., Gershman, B., & Howitt, P. (2013). How Inflation Affects Macroeconomic Performance: An Agent-Based Computational Investigation. *Macroeconomic Dynamics*, 20(2), 558-581.

August, T., & Tunca, T. I. (2006). Network software security and user incentives. *Management Science*, 52(11), 1703-1720.

Audretsch, D. B., & Mahmood, T. (1995). New firm survival: new results using a hazard function. *The Review of Economics and Statistics*, 77(1), 97-103.

Axelrod, R. (2007). Simulation in Social Sciences. In J. Rennard (Ed.), *Handbook of Research on Nature-Inspired Computing for Economics and Management* (pp. 90-100). IGI Global.

Ballot, G., Mandel, A., & Vignes, A. (2014). Agent-based modeling and economic theory: where do we stand?. *Journal of Economic Interaction and Coordination*, 10(2), 199-220.

Baptista, M. L., Martinho, C. R., Lima, F., Santos, P. A., & Prendinger, H. (2014). Applying Agent-Based Modeling to Business Simulations. *Developments in Business Simulation and Experiential Learning*, 41, 179-183

Barro, R.J. and H. Grossman. (1971). A General Disequilibrium Model of Income and Employment. *American Economic Review*, 61(1), 82-93.

Beas, D. (2012). *Obama 2012: de los medios sociales a la estrategia de datos*. Memorando nº 182/2012, Fundación Alternativas.

Belleflamme, P., & Peitz, M. (2015*). Industrial organization: markets and strategies*. Cambridge University Press.

Box, G.E.P., and N. R. Draper. (1987). *Empirical model-building and response surfaces*. John Wiley & Sons.

Bowles, S., & Gintis, H. (1993). The revenge of homo economicus: contested exchange and the revival of political economy. *The Journal of Economic Perspectives,* 7(1), 83-102.4

Browning, M., & Crossley, T. F. (2001). The life-cycle model of consumption and saving. The Journal of Economic Perspectives, 15(3), 3-22.

Bröcker, J. (1998) Operational spatial computable general equilibrium modelling. *Annals of Regional Science*, 32(3), 367-387.

Buliung, R. N. (2001). Spatiotemporal patterns of employment and non-work activities in Portland, Oregon. In *Proceedings of the 2001 ESRI International User Conference*.

Buliung, R. N., & Kanaroglou, P. S. (2004). On design and implementation of an objectrelational spatial database for activity/travel behaviour research. *Journal of Geographical Systems*, 6(3), 237-262.

Caballero, R.J. (2010) Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome. *Journal of Economic Perspectives*, 24(4), 85-102.

Caggese, A., & Cuñat, V. (2013). Financing constraints, firm dynamics, export decisions, and aggregate productivity. *Review of Economic Dynamics*, 16(1), 177-193.

Canova, F. (2011) *Methods for applied macroeconomic research*. Vol. 13. Princeton University Press.

Castelfranchi, C. (2014). Making Visible "The Invisible Hand": The Mission of Social Simulation. In D. Adamatti, G. Dimuro, & H. Coelho (Eds.) *Interdisciplinary Applications of Agent-Based Social Simulation and Modeling* (pp. 1-19). IGI Global.

Cavusoglu, H., Mishra, B., & Raghunathan, S. (2004). A model for evaluating IT security investments. *Communications of the ACM*, 47(7), 87-92.

Chang, M. H. (2011). Agent-based Modeling and Computational Experiments in Industrial Organization: Growing Firms and Industries in silico. *Eastern Economic Journal*, 37(1), 28.

Chie, B. T., & Chen, S. H. (2014). Competition in a New Industrial Economy: Toward an Agent-Based Economic Model of Modularity. *Administrative Sciences*, 4(3), 192-218.

Chari, V. V. (2010). Statement for the committee on Science and Technology, Subcommittee on Investigations and Oversight, US House of Representatives, *Building a Science of Economics for the Real World.* (pp. 32-37) 111th Congress, Government Printing Office.

Chaturvedi, A., Mehta, S., Dolk, D., & Ayer, R. (2005). Agent-based simulation for computational experimentation: Developing an artificial labor market. *European Journal of Operational Research*, 166(3), 694-716.

Cincontti, S., R. M. Raberto and A. Teglio. (2010). Credit money and macroeconomic instability in the agent-based model and simulator Eurace. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 1-32.

Cincotti, S., Raberto, M. & Teglio, A. (2012). The Eurace macroeconomic model and simulator. In Agent-based Dynamics, Norms, and Corporate Governance. *Proceedings of the 16th World Congress of the International Economic Association*, Palgrave.

Clarida, R., J. Gali and M. Gertler. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*. 37(4), 1661-1707.

Clementi, G. L., & Palazzo, B. (2013). *Entry, exit, firm dynamics, and aggregate fluctuations* (Working Paper 19217). National Bureau of Economic Research.

Colander, D., Howitt, P., Kirman, A., Leijonhufvud, A., & Mehrling, P. (2008). Beyond DSGE models: toward an empirically based macroeconomics. *The American Economic Review*, 98(2), 236-240.

Conlisk, J. (1988). Optimization cost. *Journal of Economic Behavior and Organization*, 9, 213–228.

Davis, J P., Eisenhardt, K.M. & Bingham, C.B. (2007). Developing theory through simulation methods. *Academy of Management Review*, 32(2), 480-499.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2009). Optimal structure, market dynamism, and the strategy of simple rules. *Administrative Science Quarterly*, 54(3), 413-452.

Dawid, H, S. Gemkow, S. Van Der Hoog and M. Neugart. (2011). The Eurace@ Unibi Model: An Agent-Based Macroeconomic Model for Economic Policy Analysis. *Working paper*. Universität Bielefeld.

De los Llanos, M., & Mora-Sanguinetti, J. S. (2012). Comercio minorista y regulación autonómica: efectos en la densidad comercial, el empleo y la inflación. *Revista de Economía Aplicada*, 20(59), 5-54.

De Marchi, V. (2012). Environmental innovation and R&D cooperation: Empirical evidence from Spanish manufacturing firms. *Research Policy*, 41(3), 614-623.

Deissenberg, C., Van Der Hoog, S. And H. Dawid. (2008) EURACE: A massively parallel agent-based model of the European economy. *Applied Mathematics and Computation* 204(2), 541-552.

Dejonj, D.N. and C. Dave. (2011) *Structural macroeconometrics*. Princeton University Press.

Díaz de Rada, V. (2012). ¿Deben regularse los horarios comerciales? Delimitación de los perfiles más y menos favorables?. *Distribución y consumo*, 22(123), 67-75.

Diebold, F.X. (1998). The Past, Present, and Future of Macroeconomic Forecasting. *Journal of Economic Perspectives*, 12, 175-192.

Dilaver, O., Jump, R., & Levine, P. (2016). Agent-based Macroeconomics and Dynamic Stochastic General Equilibrium Models: Where Do We Go from Here?. *Discussion Papers in Economics*, 01/16. School of Economics, University of Surrey.

D'Orazio, M., M. Di Zio and M. Scanu. (2006) *Statistical matching: Theory and practice*. Wiley.

Dosi, G., G.Fagiolo and A. Roventini. (2010). Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles. *Journal of Economic Dynamics and Control*, 34(9), 1748-1767.

Dunne, T., Roberts, M.J. & Samuelson, L. (1989). Patterns of firm entry and exit in U.S. manufacturing industries, *Rand Journal of Economics*, 19, 495-515.

Ebenhoh, E. & Pahl-Wostl, C. (2007). Agent-Based Modeling with Boundedly Rational Agents. In J. Rennard (Ed.), *Handbook of Research on Nature-Inspired Computing for Economics and Management* (pp. 225-245). IGI Global.

Elgin, C., & Tumen, S. (2012). Can sustained economic growth and declining population coexist?. *Economic Modelling*, 29(5), 1899-1908.

Epstein, J.M. and R.Y. Axtell. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press.

Evans, D.S. (1987). The relationship between firm growth, size and age: estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 35(4) 567-581

Fernández, E. and Rubiera, F. (2013) *Defining the Spatial Scale in Modern Regional Analysis: New Challenges from Data at Local Level*. Springer, Advances in Spatial Science.

Fernández Menéndez de Andés, F. (2012). Liberalización del comercio y de los horarios comerciales: falacias habituales y evidencia empírica. *Información Comercial Española, ICE: Revista de economía*, (868), 57-74.

Fingleton, B. (2000). Spatial econometrics, economic geography, dynamics and equilibrium: a ‘third way’? *Environment and Planning A*, 32(8), 1481-1498.

Fisher, F.M. (1983) *Disequilibrium Foundations of Equilibrium Economics*, Cambridge: Cambridge University Press.

Friedman, M. (1957) *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Friedman, M. and A.J. Schwartz. (1963) *A Monetary History of the United States, 1867-1960*. Princeton: Princeton University Press.

Frihida, A., Marceau, D. J., & Theriault, M. (2002). Spatio-temporal object-oriented data model for disaggregate travel behavior. *Transactions in GIS*, 6(3), 277-294.

Galán, J.M., L.R: Izquierdo, S.S. Izquierdo, J.I. Santos, R. Del Olmo, A. López and B. Edmonds. (2009). Errors and artefacts in agent-based modelling. *Journal of Artificial Societies and Social Simulation*, 12(1), 1.

Gallego, F.J. (2010). A population density grid of the European Union. *Population and Environment*, 31, 460–473.

Gal-Or, E., & Ghose, A. (2005). The economic incentives for sharing security information. *Information Systems Research*, 16(2), 186-208.

García-Milá, T. and Montalvo, J.G. (2007). The impact of new highways on business location: New evidence from Spain. *Working Paper* Centre de Recerca en Economia Internacional, Universitat Pompeu Fabra,

García-Tabuenca, A., Pablo-Martí, F. & Crecente-Romero, F. (2015). Women entrepreneurship: Changes in access to credit and business results (2003-2013). *Proceedings of the 55th Conference of the European Regional Science Association*.

Gatti, D. D., Gaffeo, E., & Gallegati, M. (2010). Complex agent-based macroeconomics: a manifesto for a new paradigm. *Journal of Economic Interaction and Coordination*, 5(2), 111-135.

Gaudet, G., & Salant, S. W. (2015). The Hotelling model with multiple demands. *Handbook on the Economics of Natural Resources*, 23-40.

Geard, N., McCaw, J. M., Dorin, A., Korb, K. B., & McVernon, J. (2013). Synthetic Population Dynamics: A Model of Household Demography. *Journal of Artificial Societies and Social Simulation*, 16(1), 8.

Geroski, P. A., Mata, J., & Portugal, P. (2010). Founding conditions and the survival of new firms. *Strategic Management Journal*, 31(5), 510-529.

Gibrat, R. (1931). *Les Inegalites Economiques*. Paris: Sirey.

Gintis, H. (2006). The emergence of a price system from decentralized bilateral exchange. *BE Journal of Theoretical Economics*, 6(1), 1302-1322.

Giri, F., Riccetti, L., Russo, A., & Gallegati, M. (2016). *Monetary Policy and Large Crises in a Financial Accelerator Agent-Based Model*. MPRA Paper No. 70371

Gordon, R.J. (1990). U.S. Inflation, Labor's Share, and the Natural Rate of Unemployment. In *Economics of Wage Determination*, Springer-Verlag.

Gordon, L. A., & Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC),*5(4), 438-457.

Grazzini, J., & Richiardi, M. (2015). Estimation of ergodic agent-based models by simulated minimum distance. *Journal of Economic Dynamics and Control*, *51*, 148-165.

Grossklags, J., Christin, N., & Chuang, J. (2008 a). Secure or insure?: a game-theoretic analysis of information security games. In *Proceedings of the 17th international conference on World Wide Web* (pp. 209-218). ACM.

Grossklags, J., Christin, N., & Chuang, J. (2008b). Security Investment (Failures) in Five Economic Environments: A Comparison of Homogeneous and Heterogeneous User Agents. In *Proceedings of the 7th Workshop of Economics of Information Security.*

Grossklags, J., Christin, N., & Chuang, J. (2008c). Predicted and Observed User Behavior in the Weakest-link Security Game. In *Usability, Psychology and Security (UPSEC 2008) online proceedings.*

Grossklags, J., Johnson, B., & Christin, N. (2010). The price of uncertainty in security games. In *Economics of Information Security and Privacy* (pp. 9-32). Springer US.

Gualdi, S., Tarzia, M., Zamponi, F., & Bouchaud, J. P. (2015). Monetary policy and dark corners in a stylized agent-based model. *Journal of Economic Interaction and Coordination*, in press.

Hamilton, J. D. (2003). What is an oil shock?. *Journal of econometrics*, 113(2), 363-398.

Hare, M., & Deadman, P. (2004). Further towards a taxonomy of agent-based simulation models in environmental management. *Mathematics and computers in simulation*, *64*(1), 25-40.

Harting, P. (2015). Stabilization policies and long term growth: Policy implications from an agent-based macroeconomic model. Technical Report 6, *Bielefeld Working Papers in Economics and Management*.

Henson, R., & Garfield, J. (2015). What Business Environment Changes Are Needed to Cause SME's to Take a Strategic Approach to Information Security?. In: *12th Annual International Conference on SME's, Entrepreneurship and Innovation*: Management – Marketing – Economic – Social Aspects, Athens, Greece.

Herfindahl, O.C. (1950). *Concentration in the steel industry*, PhD dissertation, Columbia University.

Heriques de Gusmão, A. P., Camara e Silva, L., Silva, M. M., Poleto, T., & Cabral Seixas Costa, A. P. (2016). Information security risk analysis model using fuzzy decision theory. *International Journal of Information Management*, 36(1), 25-34.

Hicks, J.R. (1937) Mr. Keynes and the "Classics" - A Suggested Interpretation. *Econometrica*, 5(2) 147-159.

Hicks, J. R. (1980) IS-LM: An Explanation, *Journal of Post Keynesian Economics*, 3, 139-155.

Hirschman, A. O. (1945). *National power and the structure of foreign trade* (Vol. 17). Berkeley: University of California Press.

Hotelling, H. (1931). The economics of exhaustible resources. *The journal of political economy*, 39(2), 137-175

Howitt, P. (2006). Coordination Issues in Long-Run Growth, Pp. 1605-1624 in *Agent-Based Computational Economics. Agent-Based Computational Economics. Vol 2*, coord. By L. Tesfatsion y K. Judd, Amsterdam: North Holland.

Isard, W. (1960). *Methods of Regional Analysis*, Technology press of MIT and Wiley, New York.

Janssen, M.A. (2005). *Agent-based modelling*. Entry prepared for the Internet Encyclopaedia of Ecological Economics.

Jiang, L., Anantharam, V., & Walrand, J. (2008). Efficiency of selfish investments in network security. In *Proceedings of the 3rd international workshop on Economics of networked systems* (pp. 31-36). ACM.

Jiang, L. (2009). *Optimization and Incentives in Communication Networks*. Doctoral dissertation, University of California, Berkeley).

Jiang, L., Anantharam, V., & Walrand, J. (2011). How bad are selfish investments in network security? *IEEE/ACM Transactions on Networking*, 19(2), 549-560.

Joyce, M., Miles, D. Scott, A. & Vayanos, D. (2012). Quantitative Easing and unconvencional monetary policy – an introduction. *The Economic Journal*, 122(564), 271-288.

Kaszowska, J., & Santos, J. L.(2014). The Role of Risk Perception in the Systemic Risk Generation and Amplification: Agent-Based Approach. *ACRN Journal of Finance and Risk Perspectives*, 3(4), 146-170.

Kavouras, M. (2003). Understanding and modelling spatial change. Life and Motion of Socio-Economic Units. GISDATA Series, 8.

Kim, B. C., & Shapiro, J. K. (2005). On the efficacy of detecting and punishing selfish peers. In *Internet and Network Economics* (pp. 836-847). Springer Berlin Heidelberg.

Kirman, A. (2010) The Economic Crisis is a Crisis for Economic Theory. *CESifo Economic Studies*, 56(4), 498-535.

Klügl, F., & Bazzan, A. L. (2012). Agent-based modeling and simulation. *AI Magazine*, *33*(3), 29-40.

Klügl, F. (2008). A validation methodology for agent-based simulations. In *Proceedings of the 2008 ACM symposium on Applied computing* (pp. 39-43). ACM.

Kunreuther, H., & Heal, G. (2003). Interdependent security. Journal of risk and uncertainty, 26(2-3), 231-249.

Kupiszewska, D., & Kupiszewski, M. (2005). A revision of the traditional multiregional model to better capture international migration: The MULTIPOLES model and its application. *Central European Forum for Migration Research*.

Kydland, F., and E.C. Prescott. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50(6), 1345-1370.

Lanne, M., Luoto, J., & Nyberg, H. (2014). *Is the Quantity Theory of Money Useful in Forecasting US Inflation?* Working Paper No. 2014-26. School of Economics and Management, University of Aarhus.

Lazonick, W. (2005). The innovative firm. *The Oxford book of Innovation*, pp. 29-55.

Lesage, J.P. (2008). An introduction to Spatial Econometrics. *Revue d'économie industrielle*. 123, 18-44.

Little, I.M.D. (1982). *Economic development: theory, policy, and international relations.* New York: Basic Books.

Long, J.B. y C. Plosser. (1983). Real business cycles. *Journal of Political Economy*, 91(1), 39-69.

Lucas, R (1976) Econometric Policy Evaluation: A Critique, pp. 19–46 in *The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy, Vol. 1*, coord. By Brunner, K y A. Meltzer, New York: American Elsevier.

Maddala, G. (2001) *Econometrics,* McGraw-Hill, New York.

Mancha, T. and J. E. Villena. (1993) El estado actual de la macroeconomía: implicaciones básicas desde la perspectiva de la política económica. *Información Comercial Española*, 718, 161-178.

Mancha, T,, Pablo, F., Santos, J.L., Garcia Tabuenca, A., del Val M.T., Gallo, M.T., Buendia, R., Atallah, N. & D´Arminio, P. (2012). Evaluación y simulación del impacto de políticas públicas sobre las PYMES a través del uso de modelos basados en agentes. *Documento de Trabajo* nº 4/2012 del Instituto Universitario de Análisis Económico y Social de la Universidad de Alcalá.

Mandel, A., Jaeger, C., Fuerst, S., Lass, W., Lincke, D., Meissner, F., Pablo-Martí, F & Wolf, S. (2010). Agent-based dynamics in disaggregated growth models. *Documents de Travail* du Centre d'Economie de la Sorbonne, 2010.77

Mankiw, N.G. (1985) Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly. *Quarterly Journal of Economics,* 100(2), 529-537.

Marimon, R. and A. Scott, eds. (1999) *Computational methods for the study of dynamic economies*. Oxford University Press.

McCallum, B.T. and E. Nelson. (1999). An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis. *Journal of Money, Credit and Banking*, 31(3), 296-316.

McCune, J. T., Beatty, R. W., & Montagno, R. V. (1988). Downsizing: Practices in manufacturing firms. *Human Resource Management*, 27(2), 145-161.

Nguyen, C.T. (2008) *Developing a conceptual architecture for a generalized agent-based modeling environment.* Master's Thesis. Naval Postgraduate School. Monterey, California.

Okun, A. M. (1962). Potential GNP: its measurement and significance. In *Proceedings of the business and economic statistics section of the American Statistical* Association Vol. 7, pp. 89-104.

Pablo-Martí, F. (1999) *La movilidad empresarial en la industria española*. Ph. D. Dissertation. University of Alcala.

Pablo-Martí, F. & Arauzo-Carod, J. M. (2012). Concentration Analysis Using Microgreographic Data. In *Defining the Spatial Scale in Modern Regional Analysis* (pp. 311-326). Springer Berlin Heidelberg.

Pablo-Martí, F & Muñoz-Yebra, C. (2009). Localización empresarial y economías de aglomeración: el debate en torno a la agregación espacial. *Investigaciones regionales*, (15), 139-166.

Pablo-Martí, F., García-Tabuenca, A., Gallo, M. T., Santos, J. L., del Val, M. T., & Mancha, T. (2013). MOSIPS Agent-based model for predicting and simulating the impact of public policies on SMEs. In *Proceedings of the European Conference on Complex Systems 2012* (pp. 399-413). Springer International Publishing.

Pablo-Marti, F., Garcia-Tabuenca, A., Santos, J. L., Gallo, M. T., del Val, M. T., & Mancha, T. (2014). Macroeconomic Forecasting with Agent-Based Models: Prediction and Simulation of the Impact of Public Policies on SMEs. In *Advances in Computational Social Science* (pp. 139-155). Springer Japan.

Pablo-Martí, F., Santos, J. L., & Kaszowska, J. (2015). An agent-based model of population dynamics for the European regions. *Emergence: Complexity and Organization*, 17(2), C1.

Paelink, J.H.P. and L.H. Klaassen. (1979) *Spatial Econometrics*, Gower, Westmead, Farnborough.

Phelps, E.S. (1968) Money-wage dynamics and labor-market equilibrium, *The Journal of Political Economy*, 76(4). 678-711.

Phillips, A. W. (1958). The Relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *economica*, 25(100), 283-299.

Popoyan, L., Napoletano, M., & Roventini, A. (2016). Taming Macroeconomic Instability: Monetary and Macro Prudential Policy Interactions in an Agent-Based Model. Working Paper 6/2015. *ISI Growth.*

Rahman, A., A. Harding, R. Tanton and S. Liu. (2013). Simulating the characteristics of populations at the small are level: New validation techniques for a spatial micro simulation model in Australia. *Computational Statistics & Data Analysis*, 57(1), 149-165.

Romer, P. (2016). *The trouble with Macroeconomics*. Delivered January 5, 2016 as the Commons Memorial Lecture of the Omicron Delta Epsilon Society. Forthcoming in The American Economist.

Saari, D.G. (1985) Iterative price mechanisms, *Econometrica*, 53(5), 1117–1131.

Salle, I., Senegas, M. A., & Yildizoglu, M. (2013). How transparent about its inflation target should a Central Bank be? *An Agent-Based Model Assessment*. Working paper No. 2013-24. Groupe de Recherche en Economie Théorique et Appliquée.

Samuelson, P. A., & Solow, R. M. (1960). Analytical aspects of anti-inflation policy. *The American Economic Review*, 50(2), 177-194.

Sandulli, F. D., Fernandez-Menendez, J., Rodriguez-Duarte, A., & Lopez-Sanchez, J. I. (2012). Testing the Schumpeterian hypotheses on an open innovation framework. *Management Decision*, 50(7), 1222-1232.

Schechter, S. E., & Smith, M. D. (2003). How much security is enough to stop a thief? In *Financial Cryptography* (pp. 122-137). Springer Berlin Heidelberg.

Schechter, S. E. (2004). *Computer security strength & risk: A quantitative approach* (Doctoral dissertation, Harvard University Cambridge, Massachusetts).

Schenk, T. A., Löffler, G., & Rauh, J. (2007). Agent-based simulation of consumer behavior in grocery shopping on a regional level. *Journal of Business Research*, 60(8), 894-903.

Schivardi, F., & Viviano, E. (2011). Entry barriers in retail trade. *The Economic Journal*, 121(551), 145-170.

Segarra, A., & Teruel, M. (2014). High-growth firms and innovation: an empirical analysis for Spanish firms. *Small Business Economics*, 43(4), 805-821.

Silverman, E., J. Bijak, J. Noble, V. Cao, and J. Hilton. (2014). Semi-Artificial Models of Populations: Connecting Demography with Agent-Based-Modelling . In *Advances in Computational Social Science* (pp. 177-189). Springer Japan.

Solow, R. (1971). Some implications of alternative criteria for the firm. In Marris, R. y Wood, A. (eds.) *The Corporate Economy*, London, Macmillan.

Solow, R. (2006) Reflections on the Survey. In D. Colander (coord.). *The Making of an Economist Redux*, (pp. 234-238) Princeton University Press.

Solow, R. (2010) Statement for the committee on Science and Technology, Subcommittee on Investigations and Oversight, US House of Representatives, *Building a Science of Economics for the Real World (pp. 12-15).* 111th Congress, Government Printing Office.

Stillwell, J., & Clarke, M. (Eds.) (2011). *Population dynamics and projection methods* (Vol. 4). Springer.

Sutherland, H., R. Taylor and J. Gomulka. (2002) Combining household income and expenditure data in policy simulations, The Review of income and wealth, 48(4), 517-536.

Takadama, K., Kawai, T. and Koyama, Y. (2008) Micro- and Macro-Level Validation in Agent-Based Simulation: Reproduction of Human-Like Behaviors and Thinking in a Sequential Bargaining Game, *Journal of Artificial Societies and Social Simulation*, 11(2), 9.

Teece, D. J.(2007) Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic management journal*, 28(13), 1319-1350.

Teller, C. & Reutterer, T. (2008). The Evolving concept of Retail Attractiveness: what makes retail agglomerations attractive when customers shop at them? *Journal of Retailing and Consumer Services*, 15(3), 127-143.

Tesfatsion, L. (2007). Agents Come to Bits: Towards a Constructive Comprehensive Taxonomy of Economic Entities, *Journal of Economic Behavior and Organization*, 63 (2), 333-346.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

United Nations, Department of Economic and Social Affairs, Population Division. (2011). *World Population Prospects*: The 2012 Revision

van der Weijde, A. H., Verhoef, E. T., & van den Berg, V. A. (2014). A Hotelling Model with Price-sensitive Demand and Asymmetric Distance Costs. The Case of Strategic Transport Scheduling. *Journal of Transport Economics and Policy (JTEP)*, 48(2), 261-277.

van Eeten, M. J., & Bauer, J. M. (2008). The economics of malware, in: OECD, *Computer viruses and other malicious software*, Paris: Organisation for Economic Co-operation and Development.

Velarde, J. (2012). Problemática actual de la distribución comercial española (un estado de la cuestión). *Revista Icade*. Publicación de las Facultades de Derecho y Ciencias Económicas y Empresariales, 83-84, 413-416.

Velupillai, K.V. (2005). The unreasonable ineffectiveness of mathematics in economics *Cambridge Journal of Economics*, 29(6), 849-872.

Velupillai, K.V. (2012a). Taming the incomputable, reconstructing the nonconstructive and deciding the undecidable in mathematical economics, *New Mathematics and Natural Computation*, 8(1), 5-51.

Velupillai, K.V. (2012b). *The Epistemology of Simulation, Computation and Dynamics in Economics*, ASSRU-Algorithmic Social Science Research Unit, Discussion Papers 121, Trento.

Velupillai, K. V., & Zambelli, S. (2015). Simulation, computation and dynamics in economics. *Journal of Economic Methodology*, 22(1), 1-27.

Vermeir, A., & Bersini, H. (2015). Best Practices in Programming Agent-Based Models in Economics and Finance. In *Advances in Artificial Economics* (pp. 57-68). Springer International Publishing.

Viegas, E., Cockburn, S. P., Jensen, H. J., & West, G. B. (2014). Why free markets die: An evolutionary perspective. *arXiv preprint* arXiv:1401.5314.

Waldherr, A. and N. Wijermans. (2013). Communicating Social Simulation Models to Skeptical Minds, *Journal of Artificial Societies and Social Simulation.* 16(4), 13

Walras, L. [1877] (1954) *Elements of Pure Economics or the Theory of Social Wealth*. Boston: Harvard University Press.

West, G. R. (1995) Comparison of input–output, input–output+ econometric and computable general equilibrium impact models at the regional level. *Economic Systems Research*, 7(2), 209-227.

Wilson, T. (2011). *A Review of Sub-Regional Population Projection Methods*. The University of Queensland Working Paper.

Winter, S. G. (1971). *Economic" natural selection" and the theory of the firm*. Yale economic Essays, 4, pp. 224-272.

Wolf, S., Bouchaud, J.P., Cecconi, F., Cincotti, S., Dawid, D., Gintis, H., Hoog, S., Jaeger, C.C., Kovalevsky, D.V., Mandel, A. And Paroussos, L. (2011) Describing economic agent-based models, Dahlem ABM documentation guidelines. In: *Proceedings of the 100th Dahlem conference on new approaches in economics after the financial crisis*.

Wolf, S., S. Fürst, A. Mandel, W. Lass, D. Lincke, F. Pablo-Marti and Jaeger, C. (2013) A multi-agent model of several economic regions, *Environmental Modelling and Software,* 44, 25-43

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. The knowledge engineering review, 10(2), 115-152.

Xavier, S. R., Kelly, D., Kew, J., Herrington, M., & Vorderwülbecke, A. (2012). *Global entrepreneurship monitor 2012.* Global Report.

Závacká, J. (2015) Monetary policy during the business cycle, in *Proc. 11th International Conference on Strategic Management and its Support by Information Systems 2015*, pp. 329–338.

Zhang Q, Jager W (2011). Agent-based modeling of population dynamics in municipalities: Migration in the Derbyshire & Nottinghamshire cases in the UK. Report for the EU Prima project, WP3. University of Groningen, Center for Social Complexity Studies.

Zhang, M., Zheng, Z., & Shroff, N. B. (2015). A Game Theoretic Model for Defending Against Stealthy Attacks with Limited Resources. In *Decision and Game Theory for Security* (pp. 93-112). Springer International Publishing.

Zott, C. (2003). Dynamic capabilities and the emergence of intraindustry differential firm performance: insights from a simulation study. *Strategic management journal*, 24(2), 97-125.