OECD Environment Working Papers No. 122

Multi-objective local
environmental simulator
(MOLES 1.0): Model
specification, algorithm
design and policy
applications

**Ioannis Tikoudis,
Walid Oueslati**

<https://dx.doi.org/10.1787/151cf08a-en>

ENVIRONMENT DIRECTORATE

MULTI-OBJECTIVE LOCAL ENVIRONMENTAL SIMULATOR (MOLES 1.0): MODEL SPECIFICATION, ALGORITHM DESIGN AND POLICY APPLICATIONS - ENVIRONMENT WORKING PAPER No. 122

By Ioannis Tikoudis and Walid Oueslati (OECD)

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Authorised for publication by Simon Upton, Director, Environment Directorate.

Keywords: spatial general equilibrium, land-use model, transport model, microsimulation, greenhouse gas emissions, air pollution.

JEL codes: C68; C60; R13; D58; D62; R00; H70; R14; R40; R52.

OECD Environment Working Papers are available at www.oecd.org/environment/workingpapers.htm

JT03413205

Complete document available on OLIS in its original format

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

OECD ENVIRONMENT WORKING PAPERS

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

This series is designed to make available to a wider readership selected studies on environmental issues prepared for use within the OECD. Authorship is usually collective, but principal author(s) are named. The papers are generally available only in their original language –English or French- with a summary in the other language.

Comments on Working Papers are welcomed, and may be sent to:

OECD Environment Directorate 2 rue André-Pascal, 75775 Paris Cedex 16, France
or by email: env.contact@oecd.org.

OECD Environment Working Papers are published on
www.oecd.org/environment/workingpapers.htm

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD (2017)

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given.

All requests for commercial use and translation rights should be submitted to rights@oecd.org.

ABSTRACT

This paper describes MOLES 1.0, an integrated land-use and transport model developed with Object-Oriented Programming principles in order to combine selected characteristics from Spatial Computable General Equilibrium and microsimulation models. MOLES 1.0 models the links between urban land use, mobility patterns, urban economic activities and their environmental impacts, in particular air pollution and emissions of greenhouse gases (GHGs). These linkages are represented in a microfounded framework in which the behavioural responses of various agents are consistent with the constraints they face and the economic stimuli they receive. The proposed model is suitable for uncovering the multiple trade-offs between environmental and economic performance as well as the synergetic effects of land-use policies on the function of the urban transport system and vice versa. The technical appendix provides the complete mathematical formulation of the model, as well as a detailed description of the employed algorithms and the accompanying programming code.

Keywords: spatial general equilibrium, land-use model, transport model, microsimulation, greenhouse gas emissions, air pollution.

JEL codes: C68; C60; R13; D58; D62; R00; H70; R14; R40; R52.

RÉSUMÉ

Ce papier décrit MOLES 1.0, un modèle intégré d'utilisation des terres et de transport élaboré avec des principes de programmation orientée objet afin de combiner des caractéristiques choisies de l'équilibre général calculable spatial et des modèles de micro-simulation. MOLES 1.0 modélise les liens entre l'utilisation des sols urbains, les modes de mobilité urbaine, les activités économiques urbaines et leurs impacts environnementaux, en particulier les pollutions atmosphériques et les émissions de gaz à effet de serre (GES). Ces liens sont représentés dans un modèle aux fondements microéconomiques dans lequel les réponses comportementales de divers agents sont cohérentes avec les contraintes auxquelles elles sont confrontées et les stimuli économiques qu'elles reçoivent. Le modèle proposé est adapté pour mettre en évidence les multiples arbitrages entre les performances environnementales et économiques ainsi que les effets synergiques des politiques d'utilisation des terres sur la fonction du système de transport urbain et vice-versa. L'annexe technique fournit la formulation mathématique complète du modèle, ainsi qu'une description détaillée des algorithmes utilisés et leurs codes.

Mots-clés : modèle d'équilibre général spatialisé, modèle d'utilisation des terres, modèle de transport, micro-simulation, émissions de gaz à effet de serre, pollution de l'air.

Classification JEL: C68; C60; R13; D58; D62; R00; H70; R14; R40; R52.

FOREWORD

This report has been authored by Ioannis Tikoudis and Walid Oueslati of the OECD Environment Directorate. The authors are grateful to delegates to the Working Party on Integrating Environmental and Economic Policies for helpful comments on earlier drafts. The paper has also benefited from useful comments by external experts Elena Irwin (Ohio State University) and JunJie Wu (Oregon State University). The authors would like to thank the OECD colleagues Ruben Bibas, Nils-Axel Braathen, Jean Chateau, Rob Dellink, Olivier Durand-Lasserve and Kurt Van Dender, for multiple insights in previous versions of the paper, and Katjusha Boffa, for her excellent editorial assistance. The authors are responsible for any remaining omissions or errors. Work on this paper was conducted under the overall responsibility of Shardul Agrawala, Head of the Environment and Economy Integration Division.

TABLE OF CONTENTS

ABSTRACT	3
RÉSUMÉ.....	3
FOREWORD.....	4
1. INTRODUCTION	7
2. THE MODEL	8
2.1. Overview	8
2.2 Spatial and network configuration	11
2.3 Vehicle and building classes	12
2.4 Individual behaviour	13
2.5 Construction sector	14
2.6 Exogenous variables	14
3. SUMMARY	14
REFERENCES	15
TECHNICAL APPENDIX.....	17
A. Building blocks	17
A.1 Spatial configuration and land-use	17
A.2 Transportation networks, routes and vehicles	18
B. Core	21
B.1 Individuals, alternatives and choice sets.....	21
B.2 Utility maximisation and choice probabilities	22
B.3 Developers	25
C. Peripheral modules.....	27
C.1 Transportation.....	27
C.2 Urban structure and environmental indicators.....	29
D. Policy instruments and exogenous variables.....	30
D.1 Fuel taxation	30
D.2 Fiscal treatment of private vehicles	31
D.3 Road pricing	31
D.4 Zoning, height restrictions, boundaries and low emission zones	32
D.5 Land and property taxation.....	32
D.6 Road capacity expansion and open space conservation	32
D.7 Exogenous variables	32
E. General, stochastic user equilibrium	33
E.1 Closing conditions	33
E.2 System of equations	33
E.3. Solution.....	37
NOTATION APPENDIX.....	44

Tables

Table 1.	Summary of policy instruments in MOLES	9
Table E.1	Equation counters in MOLES.....	34
Table E.2	Sub-system solved by Module A.....	34
Table E.3	Sub-system solved by Module B.....	35
Table E.4	Sub-system solved by the core module	35
Table E.5	Sub-system solved by the transportation module	35
Table E.6	Sub-system solved by the urban morphology & environment module	36
Table E.7	Endogenous variables block 1 (x_1)	36
Table E.8	Endogenous variables block 2 (x_2)	36
Table E.9	Endogenous variables block 3 (x_3)	37
Table E.10	Endogenous variables block 4 (x_4)	37
Table E.11	Endogenous variables block 5 (x_5)	37
Table N.1	Variables in MOLES	44
Table N.2	Indicator functions in MOLES	45

Figures

Figure 1.	Overview of MOLES.....	9
Figure 2.	Network representations in MOLES	11
Figure 3.	Representation of private vehicles in MOLES	12
Figure 4.	Hierarchical classification of residential types in MOLES	12
Figure 5.	Schematic depiction of individual choice.....	13
Figure A.1	Network representations in MOLES	18
Figure B.1	Schematic depiction of individual choice.....	22
Figure E.1	Chart flows for modules A and B.....	39
Figure E.2	Left: chart flow of the network loading. Right: chart flow of module T	40
Figure E.3	Chart flow of the traffic assignment module	41
Figure E.4	The MOLES solution algorithm	43

1. INTRODUCTION

Urbanisation has been one the main cornerstones of 20th century's vast economic expansion. For decades, uninterrupted economic growth fuelled –and got fuelled by– a constant increase in urban population and land uptake of cities. Slowly, a series of interrelated challenges that exert pressure on the ability of cities to generate prosperity have emerged. Local air pollution and greenhouse gas emissions are two of the primary environmental challenges. Rough measures (OECD, 2010) estimate that 60-80% of the global CO₂ emissions are generated in cities,¹ with a growing share of urban emissions originating from transportation and stationary sources such as residential heating and cooling. Furthermore, cities span the areas in which the largest part of local air-pollution is generated and dispersed, exposing considerable fractions of the population to it.

These environmental pressures are expected to intensify as cities expand. The current trends foreshadow the total prevail of the city. In 2050, about 70% of total population is expected to live in urban areas and the global urban land cover is projected to increase steeply.² Under the business-as-usual scenario, this expansion is highly likely to exacerbate the problem of air pollution. For instance, OECD (2016) projects that, in the absence of more stringent policies, the number of premature deaths due to outdoor air pollution will increase from approximately 3 million people annually in 2010 to 6-9 million in 2060. The associated monetised cost will increase from USD 3 trillion in 2015 to USD 18-25 trillion in 2060, with the most affected areas being those densely populated with high concentrations of PM_{2.5}.³ Moreover, there is no sign that suburbanisation is likely to slow down.⁴ The uncurbed, rapid expansion of a low-density suburban fabric into the countryside is associated with a steep increase in greenhouse gas emissions. The literature has in many cases highlighted the inverse, non-linear statistical relationship between urban density and energy consumption (Newman and Kenworthy, 1989; Mindali et al., 2004) and hypotheses on the various causal channels of it have been formulated and tested. The above hard facts upgrade the importance of local policy action, *i.e.* the implementation of policy instruments targeting land-use and transportation. Such instruments are local in the sense that their optimal values may vary in a complex way, not only across different cities but also within a given urban area. The evaluation of the environmental and economic effects of these instruments requires the development of computer-intensive tools that model explicitly the several critical interdependencies characterising the economy of an urban area.

This paper provides a full documentation of the *Multi-Objective Local Environmental Simulator* (hereafter, MOLES), a model that is unique in that it combines the internal consistency of an *urban general equilibrium model* (U-CGE) with the detail and the additional predictive power of a microsimulation

¹ The estimation of the exact fraction is a formidable task, see OECD (2015a, 2015b).

² In Angel et al. (2011) the increase is projected from 603 thousand km² in 2000 to over 3 million km² in 2050.

³ Earlier projections by the OECD (2012) estimated that the worldwide mortality due to particulate matter alone (PM_{2.5} and PM₁₀) is expected to increase from 1 million in 2000 to over 3.5 million in 2050.

⁴ In contrast, there is evidence that suburbanisation is uninterrupted even in OECD countries: 66 out of the 78 largest OECD metropolitan areas experienced a faster growth of their suburban belt than their urban core over the period 1995-2005 (OECD, 2010).

model. U-CGE is a sub-class of the CGE modelling tradition. The latter is known for its capacity to assess the economic impacts of various policy interventions. These economic impacts include the overall welfare effect of such policy-induced shocks to the average household of the city as well as its distributional impact, *i.e.* the distribution of the monetised gains or losses across the various population cohorts. In general, CGE models are particularly powerful not only in computing the above effects, but also in uncovering the various channels through which the latter are realised. That is, they disentangle the direct effects in a certain market (for instance due to the introduction of a tax) from the various secondary effects that kick in after adjustment mechanisms, following the policy shock, are in place. In the context of U-CGE models, these secondary effects capture the impact of certain land-use policies on transportation demand and, similarly, the impact of local mobility instruments on urban development. Therefore, as a U-CGE model, MOLES captures and quantifies the interactions between the various land-use and transport policies.

CGE models have been criticised –among others– for their low predictive power. Part of this weakness may be associated with their inability to model agents and tradable objects with idiosyncratic characteristics. In contrast, microsimulation frameworks model consumer choice in a detailed level. MOLES is a C# code developed completely upon Object-Oriented Programming (OOP) principles in order to handle a large number of selected elements from such a framework: heterogeneous households with different preferences, skills and wealth stocks, as well as a rich representation of the markets for housing, land and private vehicles. This facilitates forecasts over the *environmental effectiveness* of detailed policies that may be targeting the use of a narrow set of vehicle classes in a city, the use of vehicles in a narrow set of urban zones or the development of specific building types in specific areas.

Section 2 provides a non-technical introduction to MOLES, the policy instruments whose impacts it is especially designed to investigate, the way vehicles, buildings and transportation networks are represented as well as the way individual behaviour is modelled. Section 3 summarises the accompanying Technical Appendix provides the complete mathematical formulation of the model, chart flows and detailed descriptions of the solution algorithms and pseudocode disengaged from any programming language.

2. THE MODEL

2.1. Overview

MOLES is an amalgam of a U-CGE and an urban microsimulation model⁵ designed to represent the long-run evolution of urban areas. The model is tailored to evaluate the environmental and economic effects of policies targeting the land-use and urban mobility patterns. A summary of the instruments making up the policy mix is provided in Table 1. These policy effects depend on the given idiosyncratic characteristics of a city, such as its spatial lay-out, its physical morphology and the configuration of its transport networks. They are also subject to the evolution of key exogenous economic and demographic forces determined at the national and international level. The upper part of Figure 1 displays the exogenous input to the model. A summary of MOLES's policy instruments and exogenous variables is provided in Table 1.

⁵ See Anas and Kim, 1996; Anas and Xu, 1999; Anas and Liu, 2007; Tscharktschiew and Hirte, 2010; Anas and Hiramatsu, 2012, 2013; Tikoudis, 2015 for some examples of general equilibrium models at the city level. See Waddell, 1998a; 1998b; 1998c; Waddell et al., 1998; Waddell, 2002 for some representative work with urban microsimulation models.

Figure 1. Overview of MOLES

Note: *Solid* arrows represent exogenous model input to the core module (spatial general equilibrium model), such as policy instruments, scenarios for the evolution of key demographic and economic variables, the city's spatial lay-out and network configuration. *Short-dashed* arrows represent the output of the core module, such as prices, population and structural densities. *Long-dashed* arrows represent the feedback sent to the core model (new speeds, travel times and pollution levels, as well as urban form indicators).

The *core* of the model (Figure 1) is a non-linear system of market clearing equations in housing and land markets solved for the corresponding equilibrium prices. The solution of this system yields the long-run development profile, population density, traffic and emission intensity at each urban zone (core's output). MOLES allows these variables to emerge endogenously as the outcome of well-defined behavioural mechanisms that characterise atomistic agents: utility maximising households that sort their location over space and a profit-maximising real-estate sector that adjusts the housing stock through construction and demolition. Each side responds to the relevant prices (*e.g.* land, housing), the aforementioned exogenous input, and a series of critical variables (feedback effects in Figure 1) that affect the willingness to pay for different housing constructs in different locations. These variables include travel times, environmental quality and a series of indicators quantifying the urban land-use patterns. The exact functions of the core module are further explained in the accompanying Technical Appendix.

Table 1. Summary of policy instruments in MOLES

Target	Type	Instrument	Section
Land-use	Tax-based	Property taxation	D.5
		Land development tax	D.5.
	Command-and-control	Building height regulations	D.4
		Urban growth boundary	D.4
		Other land supply regulations	D.4
		Other residential zoning regulations	D.4
		Industry relocation	D.4
	Investments	Open space & natural amenities	D.6

		Periurban reforestation	D.6.
Transportation	Tax-based	Cordon tolls	D.3.
		Flat kilometre tax	D.3.
		Varying kilometre tax	D.3.
		Area charging	D.3.
		Fuel taxes	D.1.
		Parking fees	D.3.
		Annual circulation taxes	D.2.
		Vehicle registration tax	
	Command-and-control	Car free areas	D.4
		Near zero emission zones	D.4
	Subsidies	Public transportation	C.1.
	Investments in infrastructure		
		New public transport networks	D.6.
		Lower-emission public transportation	Extension
		Electric vehicle charging stations	Extension
		Road capacity augmentation	D.6.
		Road network expansion	D.6.
		Support infrastructure for automated driving	Extension

Notes: *Extension* denotes that the investigation of the corresponding policy instrument requires further development or modification of the model.

The solution of the model's core equations is succeeded by an update of the aforementioned critical variables, performed by three peripheral modules displayed in the lower part of Figure 1. That is, the core's output can be passed to peripheral transport, biophysical and land-use⁶ modules to produce new feedback effects. These are then be injected back to the core of MOLES. The Technical Appendix shows that this recurrent process is a fixed point iteration that terminates when the updated feedback effects are too small to cause a considerable disturbance in the core system of equations.

⁶ The land-use module updates variables such as local visibility, access to local open public and private open space. In a future extension, the authors plan to develop further the functionality of this module in order to account for the effect of urban structure on the residential heating energy needs. The general lack of empirical work on this topic hampers the incorporation of such feedback effects in MOLES, as existing literature has concentrated mainly on the effects of urban morphology on the temperature at the exterior of buildings in urban environments (see, for instance Guo et al., 2016).

Figure 2. Network representations in MOLES

Notes: Left panel: Implicit road network in an urban area of fourteen developed zones; middle panel: highway network with nine of the developed and one undeveloped zone having direct access to it; right panel: urban and highway driving.

2.2 Spatial and network configuration

MOLES applications are built upon GIS data. A bounded region representing an urban area is partitioned into separate undeveloped, residential and employment zones where economic activity takes place. The partition is flexible and can be adapted to different types of urban morphology, geospatial information, administrative delimitations, economic data or geometrical criteria. Zone delimitations are fixed, *i.e.* the partitioned zones do not expand or contract. Thus, the total land surface of the urban area under consideration remains unchanged throughout the time horizon of the simulations. As a result of this, city growth is represented in two ways: (i) densification of existing residential zones and (ii) expansion of urban fabric to previously undeveloped zones. *Densification* occurs by raising a residential zone's average floor-to-area ratio via erection of taller buildings or via the development of vacant land plots.⁷ *Expansion* occurs via conversion of an undeveloped zone to residential or employment. Thus, MOLES models urban development explicitly by incorporating undeveloped periurban areas that could, in the horizon of a simulation, be granted for development.

Road transportation takes place upon two different networks: (i) a *highway network* that consists of larger capacity links, where vehicles can attain a high speed in the absence of traffic and (ii) an *urban road network* where, due to a series of hampering factors such as traffic lights and intersections, vehicles attain only a moderate speed even under free-flow conditions. For reasons related to computational tractability, the environmental externalities from vehicle use in the *urban road network* are captured without modelling the actual route choices that take place inside it, *i.e.* without an explicit representation of it. Instead, the aggregate traffic and the average road capacity determine the vehicle speeds and emissions within each cell of a high resolution grid.⁸ The aggregate traffic in each cell is approximated under the assumption that urban routes are straight lines connecting the origin and destination of the trips.

⁷ That is the ratio of housing floor space to the built area.

⁸ This choice introduces the level of abstraction that is desirable in this modelling approach: it keeps the model focused on the environmental aspects of total vehicle use in small to medium capacity road segments without shifting the interest to local traffic details, which (despite being the focus of traffic engineering models) bear only secondary environmental importance.

Figure 3. Representation of private vehicles in MOLES

Figure 4. Hierarchical classification of residential types in MOLES

Notes: Level 1: single versus multi-family housing, level 2: degree of detachment in single family housing, level 3: coverage coefficient, level 4: structural density. Coverage coefficient: the fraction of the land plot occupied by the footprint of the building. Structural density: the ratio between the total floor space of a building and the surface of its footprint.

2.3 Vehicle and building classes

MOLES considers various classes of private vehicles, derived from combinations of three differentiators: engine type, fuel type and size. Figure 3 displays an indicative hierarchical classification with two internal combustion engine classes (*i.e.* diesel, unleaded), one electric engine class and four size classes. The reader is referred to Section C.1 of the Technical Appendix for the representation of public transportation in MOLES. A similar classification of residential types is displayed in Figure 4. The critical differentiators include the presence of other residences in the host building (single versus multi-family), the degree of detachment from other single family residences, the fraction of the land plot occupied by the

building footprint (coverage coefficient), the number of floors (structural density) and the energy efficiency of the building. For expositional convenience, Figure 4 abstracts from defining the exact number of classes derived from the last differentiator. The land-use patterns in employment zones and the emissions from buildings used by the industrial, commercial and service sectors are treated as exogenous, for reasons explained in Section 2.6. However, the impact of these emissions can be altered by industry relocation and other policies giving rise to near-zero emission zones (see Table 1 and Section D.4 in the Technical Appendix).

Figure 5. Schematic depiction of individual choice

Notes: The choice of a feasible alternative (a1) implies a series of discrete choices (sample alternative). Combined with the individual characteristics, these choices yield the input for the conditional (to a1) household optimisation problem. Solving that problem yields the optimal consumption and residential size.

2.4 Individual behaviour

Each individual type considered in MOLES represents a population cohort with similar observable characteristics (such as skills) and preferences (taste variation). These characteristics are exogenous in the model, *i.e.* individuals cannot alter the characteristics that identify the group to which they belong. Each individual maximises utility by engaging in a series of discrete and continuous choices. *Discrete* choices include residential location, residential type, job location and sector, vehicle ownership, vehicle type, commuting mode and route choice. Any such choice combination is referred to as an *alternative*. *Continuous* choices include consumption, residential size, leisure time and the kilometres travelled for shopping or leisure purposes. Continuous choices must respect the constraints placed by the disposable income and the time endowment of a household. MOLES constructs these constraints and solves the household optimisation problem for each alternative, as depicted schematically in Figure 5.⁹ Upon solution, MOLES assigns a choice probability to each specific alternative using a statistical model for the

⁹ This may be a particularly burdensome computational task, given the large number of alternatives arising even from specifications of medium resolution. For instance an application with 50 housing types, 12 vehicle types, 150 residential zones, 5 employment zones, 3 sectors and an average of 10 competing routes for each origin destination pair may yield 13500000 alternatives.

distribution of the unobserved factors that influence individual choice. The reader is referred to the Technical Appendix for the exact formulation and solution of the household optimisation problem.

2.5 Construction sector

On the supply side, profit-maximising developers convert vacant land into different types of real estate developments or adjust the floor-to-area ratio by demolishing existing buildings to raise higher ones in developed land parcels. This choice is probabilistic in order to allow unobserved factors to affect expected profits from development or re-adjustment. Modelling the real estate sector this way allows for postponed development: a fraction of the vacant land in each zone remains undeveloped at each time period, reflecting a low current profitability. Free entry in the construction sector and competition for land between developers will raise land prices to a level in which profits from development are zero.

2.6 Exogenous variables

MOLES does not model explicitly the price setting mechanisms in the markets of vehicles and fuel, as these are more relevant for analyses at the national or international level. Therefore, pre-tax vehicle and fuel prices are introduced directly from scenarios and are treated as exogenous variables in the model. Real wages are also treated as exogenous, as modelling explicit wage setting mechanisms would require a complete model of firm behaviour. This could be a formidable task, as firm output levels (and therefore employment and wages) usually depend on demand and regulations determined at the national and, very often, international level. Therefore, the real wage by skill and location evolves exogenously. Other exogenous input to the model includes variables such as city population growth and population composition, *i.e.* the percentage of each individual type in the city population.

3. SUMMARY

This paper summarised MOLES, an environmental economic model for the long-run evolution of urban areas, focusing on the quantification of costs and benefits arising from various urban policies targeting land-use and urban mobility. The model helps policy makers distinguish possible best practices, *i.e.* progressive policy interventions that increase economic efficiency and improve environmental quality, worst practices, *i.e.* regressive interventions that harm both the economy and the environment, and the contexts in which these practices may arise. It also sheds light into the trade-offs between environmental effectiveness and economic efficiency that characterise second-best practices when a first-best intervention is absent. Finally, the model is designed to systematically explore the environmental and economic consequences of a fragmented governance structure, in which different authorities with non-compatible objectives control different policy instruments. Together with the accompanying Technical Appendix, where the reader will find the complete documentation of the model available, this paper highlights the novelties of MOLES, as well as commonalities with existing frameworks.

REFERENCES

- Anas, A., Hiramatsu, T. (2012), The effect of the price of gasoline on the urban economy: from route choice to general equilibrium. *Transportation Research Part A*, 46, 855-873.
- Anas, A., Hiramatsu, T. (2013), The economics of cordon tolling: general equilibrium and welfare analysis. *Economics of Transportation*, 2 (1), 18-37.
- Anas, A., Kim, I. (1996), General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. *Journal of Urban Economics*, 40 (2), 232-256.
- Anas, A., Liu, Y. (2007), A regional economy, land use, and transportation model (RELU-TRAN): formulation, algorithm design, and testing. *Journal of Regional Science*, 47 (3), 415-455.
- Anas, A., Xu, R. (1999), Congestion, land use, and job dispersion: a general equilibrium model. *Journal of Urban Economics*, 45 (3), 451-473.
- Angel et al. (2011), Making room for a planet of cities. Report, Lincoln Institute of Land Policy, Cambridge, MA, USA.
- Bureau of Public Roads (1964), Traffic Assignment Manual, U.S. Department of Commerce, Urban Planning Division, Washington D.C.
- Guo et al. (2016), Characterizing the impact of urban morphology heterogeneity on land surface temperature in Guangzhou, China. *Environmental Modelling and Software*, 84, 427-439.
- Mindali et al., (2004), Urban density and energy consumption: a new look at old statistics. *Transportation Research Part A*, 38 (2), 143-162.
- Newman, P., Kenworthy, J. (1989), Cities and Automobile Dependence: An International Sourcebook. Avebury Technical, Great Britain.
- OECD (2016), The economic consequences of outdoor air pollution, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264257474-en>.
- OECD (2015a), Climate Change Mitigation: Policies and Progress, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264238787-en>.
- OECD (2015b), The Economic Consequences of Climate Change, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264235410-en>.
- OECD (2012), OECD Environmental Outlook to 2050, OECD Publishing, Paris,
<http://dx.doi.org/10.1787/9789264122246-en>.
- OECD (2010), Cities and Climate Change, OECD Publishing, Paris,
<http://dx.doi.org/10.1787/9789264091375-en>.

- Tscharaktschiew, S., Hirte, G. (2010), The drawbacks and opportunities of carbon charges in metropolitan areas. A spatial general equilibrium approach. *Ecological Economics*, 70, 339-357.
- Waddell, P. (1998a), *The Oregon Prototype Metropolitan Land Use Model*, in: Proceedings of the ASCE Conference on Transportation, Land Use and Air Quality: Making the Connection, Portland, OR, USA, May, 549–558.
- Waddell, P. (1998b), *Final Report on the Oregon Prototype Metropolitan Land Use Model, Phase II*. Report (Salem, OR: Oregon Department of Transportation).
- Waddell, P. (1998c), An urban simulation model for integrated policy analysis and planning: residential location and housing market components of UrbanSim. Paper presented at the 8th World Conference on Transport Research, Antwerp, Belgium, July 1998.
- Waddell, P. (2002), UrbanSim: modelling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association*, 68(3), 297–314.
- Waddell, P., Moore, T., Edwards, S. (1998), *Exploiting parcel-level GIS for land use modelling*, in: Proceedings of the ASCE Conference on Transportation, Land Use and Air Quality: Making the Connection, Portland, OR, USA, May, pp. 632–641.

TECHNICAL APPENDIX

A. Building blocks

A.1 Spatial configuration and land-use

A.1.1 Zones, nodes and surfaces

All activities take place in an ordered set of *employment zones*, indexed by $j = \{1, 2, \dots, J\}$ and an ordered set of *residential zones*, indexed by $i = \{J + 1, J + 2, \dots, J + I\}$. The concatenation of the above sets yields the ordered set of zones, indexed by $z = \{1, 2, \dots, J, J + 1, J + 2, \dots, J + I\}$. Zone delimitations are fixed, *i.e.* zones do not expand or contract; thus, the total land surface covered by zone z , denoted by X_z , is fixed. Each zone is represented by its centroid, whose coordinates are given by the vector $\mathbf{cn}_z = (cn_z^x, cn_z^y)$. Thus the Euclidean distance between zones z and z' is: $d_{zz'} = d_{z'z} = \|\mathbf{cn}_z - \mathbf{cn}_{z'}\|$. Some of the zones serve as entrances or exits to the highway network (see below). The indicator $I(z, \mathcal{H})$ obtains the value one if the centroid of zone z functions as a *highway node* and zero otherwise. The set that contains the indexes of the highway nodes is denoted by $\mathcal{H}$.

A.1.2 Land use and residential development

Every unit of land surface in an arbitrary zone i corresponds to a unique type of land-use. The ordered set of residential development types (see below) is indexed by $r = \{1, 2, \dots, R\}$ with $r = 0$ denoting the absence of any development and the indicator function $I_p(i, r)$ returning the value one if the residential type r is allowed to be developed in zone i (zero otherwise). Let $0 \leq \beta(i, r) \leq 1$ denote the portion of surface X_i occupied by residential development r and $0 \leq \beta(i, 0) \leq 1$ the undeveloped portion of surface X_i , *i.e.* vacant land. Similarly, the portions of X_i that represent exogenously determined open space (A), road and other types of transport infrastructure (I) are denoted by $\beta(i, A)$ and $\beta(i, I)$ respectively. Finally, $\beta(i, E)$ denotes the fraction of X_i occupied by any other land-use.

The above enumeration is exhaustive, therefore in each zone:

$$\beta(i, A) + \beta(i, E) + \beta(i, I) + \beta(i, 0) + \sum_r \beta(i, r) = 1. \quad [\text{T.1}]$$

As shown in Figure 4, residential development types r are differentiated between single-family and multi-family constructs, as well as with respect to the *coverage coefficient*, μ_r , and *structural density*, ξ_r . The coverage coefficient is the fraction of the land plot surface covered by the footprint of development type r ; ¹⁰ the structural density is the number of floors characterising any building of type r . The domains of μ_r and ξ_r are finite sets denoted by $\mathcal{D}_\mu$, $\mathcal{D}_\xi$ respectively. A residential building type r is *feasible* in zone i if its (μ, ξ) combination complies with the regulations imposed in that zone.

¹⁰ For instance, $\mu_r = 0.8$ implies that for each m^2 of land occupied by development type r , there must be $\frac{1-0.8}{0.8} = 0.25$ m^2 used as backyard open space.

A.2 Transportation networks, routes and vehicles

To move from the centroid of any zone z to the centroid of any zone z' , private vehicles may use two different road networks: (i) a *highway network* that consists of larger capacity links, where vehicles can attain a high speed under absence of traffic (referred to as the free-flow speed) and (ii) an *urban road network* where, due to a series of hampering factors such as traffic lights and intersections, vehicles attain only a moderate speed even under free-flow conditions. For reasons related to computational tractability, the environmental externalities from vehicle use inside the *urban road network* are captured without modelling the actual route choices that take place inside it, *i.e.* without an explicit network representation. As it will be shown later, the traffic conditions and emissions are determined in an aggregate fashion inside each zone z and are, therefore, uniform across all virtual road links within it. This choice introduces the level of abstraction that is desirable for an urban environmental model: it keeps the model focused on the environmental aspects of total vehicle use in small to medium capacity road segments without shifting the interest to local traffic details, which (despite being the focus of traffic engineering models) bear only secondary environmental importance.

Figure A.1 Network representations in MOLES

Notes: Left panel: Implicit road network in an urban area of fourteen zones; middle panel: highway network with nine of the zones serving as highway nodes; right panel: comparing an urban route from zone 1 to zone 5 (straight line) to a mixed route consisting of (a) the urban route $1 \rightarrow 10$, (b) the highway route $10 \rightarrow 7 \rightarrow 6$ and (c) the urban route $6 \rightarrow 5$.

A.2.1 Highway network representation and highway routes

Two arbitrary nodes of the highway network, $z' \in \mathcal{H}$ and $z \in \mathcal{H}$, are *neighbouring* if there is at least one *highway link* $l_{z' \rightarrow z}^H$ starting at (zone) node z' and ending at (zone) node z . Links are directed, thus $l_{z' \rightarrow z}^H \neq l_{z \rightarrow z'}^H$. However, for the objectives of this paper it is assumed that the network is symmetric, *i.e.* every $l_{z' \rightarrow z}^H$ is mirrored by an identical link $l_{z \rightarrow z'}^H$. A *highway route* (q^H) is defined as a *sequence* (ordered list) of links such that, for each pair of consecutive links in the sequence, $l_{z' \rightarrow z}^H$ and $l_{z \rightarrow z'}^H$, it holds that $z = \hat{z}$. The length of an arbitrary link is denoted by $d_{z' \rightarrow z}^H$. Therefore, the length of highway route q^H is:

$$d(q^H) = \sum_{l_{z' \rightarrow z}^H \in q^H} d_{z' \rightarrow z}^H. \quad [\text{T.2}]$$

No route q^H can reach the same node twice, *i.e.* *cyclical* routes that contain any pair of links $l_{z \rightarrow z'}^H$ and $l_{z' \rightarrow z}^H$, for which it holds that $z = \hat{z}'$, are excluded. By assumption, the highway network is used exclusively by private vehicles. An example of a highway network is displayed on the middle panel of Figure A.1.

A.2.2 Implicit urban road network representation

The rest of the transportation infrastructure is accessible by both public transportation and private vehicles and is composed out of low or medium capacity road segments. This is represented by an *implicit urban network* that lacks concrete road links. Therefore, the model abstracts from urban route choice, assuming instead a unique virtual route that connects each node z' to each node z with a straight line. Thus, each urban route, denoted by q^U , is a straight line of length equal to $d_{zz'} = \|\mathbf{cn}_z - \mathbf{cn}_{z'}\|$. The m -th element of the *urban route assignment vector* $\mathbf{b}_{q^U} = (b_{q_1^U}, \dots, b_{q_J^U}, b_{q_{J+1}^U}, \dots, b_{q_{J+I}^U})$, *i.e.* $0 \leq b_{q_m^U} < 1$, represents the fraction of $d_{zz'}$ that passes through the territory of the arbitrary zone m . Therefore, it holds that $\sum_{m=1}^{J+I} (b_{q_m^U}) = 1$. Knowing vector $\mathbf{b}_{q^U}$ allows the decomposition of the length of urban route q^U into the separate distances travelled on the territory of each zone. That is:

$$d_{zz'} = d(q^U) = \underbrace{d_{zz'} b_{q_1^U}}_{\text{distance into zone 1}} + \dots + \underbrace{d_{zz'} b_{q_{J+I}^U}}_{\text{distance into zone } J+I}. \quad [\text{T.3}]$$

It is straightforward that the distance travelled into zones for which $b_{q_m^U} = 0$ will be zero. The decomposition in [T.3] is useful to calculate the travel time of the urban route as a sum of the travel times in each zone through which route q^U passes, *i.e.* each zone m for which $b_{q_m^U} > 0$. This exercise is postponed until Section A.2.4. An example of an implicit road network is displayed on the left panel of Figure A1.

A.2.3 Mixed routes

For private vehicles, it is possible to reach a zone z' from some other zone z using both networks, provided that at least one of the zones z' and z is not serving as a network zone. That is, a *mixed route* from z' to z is considered in the model if and only if: $I(z', \mathcal{H}) \cdot I(z, \mathcal{H}) = 0$.

Such a mixed route is denoted by $q^M = (q_{\text{in}}^U, q^H, q_{\text{out}}^U)$ where: (i) q_{in}^U is an urban route from the origin node z' to the closest highway node $\hat{z}$, (ii) q_{out}^U is an urban route to the destination node z from the closest highway node $\hat{z}$ and (iii) q^H is a highway route from highway node $\hat{z}$ to highway node $\check{z}$. Parts (i) and (ii) are necessary only if $z' \neq \hat{z}$ and $\hat{z} \neq z$ respectively. That is, if $z' = \hat{z}$ then $q_{\text{in}}^U = \emptyset$ and if $\hat{z} = z$ then $q_{\text{out}}^U = \emptyset$. The right panel of Figure A1 illustrates an example of a mixed route from zone 1 to zone 5 in which $z' \neq \hat{z}$ and $\hat{z} \neq z$. In that example, q_{in}^U is represented by the urban route from zone 1 to zone 10, q^H by the sequence of highway links $l_{10 \rightarrow 7}^H$ and $l_{7 \rightarrow 6}^H$ and q_{out}^U by the urban route from zone 6 to zone 5.

The total distance covered in the implicit network by a mixed route $q^M = (q_{\text{in}}^U, q^H, q_{\text{out}}^U)$ is:

$$d(q_{\text{in}}^U) + d(q_{\text{out}}^U) = (1 - I(z', \mathcal{H}))d_{z'\hat{z}} + (1 - I(z, \mathcal{H}))d_{\hat{z}z}, \quad [\text{T.4}]$$

and can be distributed to the various zones using the decomposition equation in [T.3]. Similarly, $d(q^H)$ is computed from [T.2].

A.2.4 Road technology

The *resulting* speed (expressed in km/h) in an arbitrary highway link $l_{z' \rightarrow z}^H$ is given by the following BPR (Bureau of Public Roads, 1964) volume-delay function:

$$SP(l_{z' \rightarrow z}^H) = \frac{1}{\xi_0^H(l_{z' \rightarrow z}^H) \left(1 + \xi_1^H \left(\frac{TL(l_{z' \rightarrow z}^H)}{RK(l_{z' \rightarrow z}^H)} \right)^{\xi_2^H} \right)}, \quad [\text{T.5}]$$

where $TL(l_{z' \rightarrow z}^H)$ is the total vehicle load using the link, that is going to be computed in equation [T.44]; $RK(l_{z' \rightarrow z}^H)$ is the serving capacity of the link; $\xi_0^H(l_{z' \rightarrow z}^H)$ is the inverse of the link's free flow speed, *i.e.* the speed that prevails in total absence of traffic, when setting $TL(l_{z' \rightarrow z}^H) = 0$; ξ_1^H and ξ_2^H are shape parameters that govern the sensitivity of travel speed to changes in the load-to-capacity ratio, $TL(l_{z' \rightarrow z}^H)/RK(l_{z' \rightarrow z}^H)$.

Multiplying the inverse of the resulting speed (expressed in h/km) with the length of the link and summing across a route's links yields the travel time required to drive through q^H . This is:

$$t_c(q^H) = \sum_{l_{z' \rightarrow z}^H \in q^H} \left\{ \left(\frac{d_{z' \rightarrow z}^H}{SP(l_{z' \rightarrow z}^H)} \right) \right\}. \quad [\text{T.6}]$$

A similar road technology is assumed for the urban routes. As seen already, these routes are not composed by concrete road links, but can be decomposed into trip segments, each one passing through a different zone m with an aggregate level of traffic denoted by TL_m and a road capacity $RK_m = \beta(m, I)X_m$. The car speed in zone m is given by:

$$SP_m^{U,C} = \frac{1}{\xi_{0m}^{U,C} \left(1 + \xi_1^{U,C} \left(\frac{TL_m}{\beta(m, I)X_m} \right)^{\xi_2^{U,C}} \right)}, \quad [\text{T.7}]$$

where the urban free-low parameters $\xi_{0m}^{U,C} > \xi_0^H$ to reflect the critical factors that differentiate urban driving from highway driving. Such speed-reducing factors include traffic lights, road intersections and interactions with pedestrians. The corresponding speed of public transport modes in zone m is given by:

$$SP_m^{U,PT} = \frac{1}{\xi_{0m}^{U,PT} \left(1 + \xi_1^{U,PT} \left(\frac{TL_m}{\beta(m, I)X_m} \right)^{\xi_2^{U,PT}} \right)}, \quad [\text{T.8}]$$

where the free-flow parameter $\xi_{0m}^{U,PT} > \xi_{0m}^{U,C}$ reflects the fact that public transportation requires frequent stops, transit time and that it might be prone to disruptions, factors that increase the expected free-flow travel time.

The corresponding travel times of urban route q^U for car and public transport are:

$$t_c(q^U) = \sum_{m=1}^{J+I} \left\{ \left(\frac{d(q^U) b_{qm}^U}{SP_m^{U,C}} \right) \right\}, \quad [\text{T.9}]$$

and

$$t_{PT}(q^U) = \sum_{m=1}^{J+I} \left\{ \left(\frac{d(q^U) b_{qm}^U}{SP_m^{U,PT}} \right) \right\}, \quad [\text{T.10}]$$

respectively. Knowing the travel times for all highway and urban routes is sufficient to compute the travel times for mixed routes. Recalling that a mixed route $q^M = (q_{in}^U, q^H, q_{out}^U)$ can only be performed by private vehicles, its corresponding travel time is:

$$t_c(q^M) = \begin{cases} t_c(q_{in}^U) + t(q^H) + t_c(q_{out}^U) & \text{if } q_{in}^U \neq \emptyset \text{ and } q_{out}^U \neq \emptyset \\ t(q^H) + t_c(q_{out}^U) & \text{if } q_{in}^U = \emptyset \text{ and } q_{out}^U \neq \emptyset \\ t_c(q_{in}^U) + t(q^H) & \text{if } q_{in}^U \neq \emptyset \text{ and } q_{out}^U = \emptyset \end{cases} \quad [T.11]$$

A.2.5 Vehicles and fuel types

Let $v = \{1, 2, \dots, V\}$ index an ordered set of vehicles and introduce the indicator function $I_V(v, e)$, which takes the value one if the vehicle v is electric and zero if it has an internal combustion engine. Vehicles are differentiated according to a set of attributes: fixed costs (covering registration, license and annual fees) denoted by c_v^F ; depreciation costs, denoted by c_v^D ; size, denoted by s_v . Finally, each internal combustion engine vehicle uses a unique fuel type from the ordered set of fuels indexed by $g = \{1, 2, \dots, G\}$. Let the subscript $g(v)$ denote fuel type g compatible to vehicle type v .

The pecuniary cost of vehicle use per kilometre is given by:

$$c_{v,km}^x = \begin{cases} l_v^x \cdot p_{g(v)} & \text{if } I_V(v, e) = 0 \\ e_v^x \cdot p_e & \text{if } I_V(v, e) = 1 \end{cases} \quad [T.12]$$

where p_e is the per unit price of electricity; p_g is the per litre final price of fuel type g compatible to vehicle type v ; l_v^x is the fuel consumption per kilometre travelled by the internal combustion engine vehicle v ; e_v^x is the electricity consumption per kilometer travelled by the electric vehicle v . These consumption coefficients are independent of traffic conditions, thus consumption per kilometre is not affected by the level of congestion. However, they are differentiated between urban and highway driving by the superscript x , with $x = U$ denoting the former case and $x = H$ the latter.

B. Core

B.1 Individuals, alternatives and choice sets

The model assumes an exogenous ordered set of N individual types, with an arbitrary individual type being indexed by n . Individual types are differentiated by exogenous characteristics (such as skill category and age cohort) and preferences. Each individual chooses an alternative:

$$\mathbf{a} = \{i, r, j, v, q\}, \quad [T.13]$$

which is a unique combination of residential zone i , residential type r , employment zone j , vehicle type v and route q . The choice set formed by these alternatives is individual-specific for two reasons. The first reason is *ex-ante compatibility*. That is, some individual types may be incompatible to certain alternatives. For instance, employment zone j may host production sectors that do not employ the skill category s to which individual n belongs to. In that case any alternative $\mathbf{a}$ that contains employment zone j is *ex-ante* incompatible. The alternatives $\mathbf{a}$ that are *ex-ante* compatible form the temporary choice set of individual n , denoted by $\mathcal{C}_n^T$.

Furthermore, an *ex-ante* compatible alternative may not be *economically affordable*, i.e. one or more of the continuous variables (optimal consumption, residential size and leisure time) associated with the discrete choice $\mathbf{a} = \{i, r, j, v, q\}$ may not be strictly positive. The economic affordability condition will become apparent to the reader after reading Section B.2, where the optimal levels of the continuous variables for any choice of alternative $\mathbf{a}$ are derived analytically.

Figure B.1 Schematic depiction of individual choice

B.2 Utility maximisation and choice probabilities

Suppose that individual n chooses an *ex-ante* compatible alternative $\mathbf{a} = \{i, r, j, v, q\}$ from the temporary choice set $\mathcal{C}_n^T$ (see *sample alternative* in Figure B1). Conditional on that choice, the individual chooses the residential size (denoted by $H_{\mathbf{a}}$) and consumption (denoted by $C_{\mathbf{a}}$) that maximize the utility function:

$$U_{\mathbf{a}} = \Omega_{\mathbf{a}} + u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n) + u_B(C_{\mathbf{a}}, H_{\mathbf{a}}, \ell_{\mathbf{a}}) + \varepsilon_{\mathbf{a}n}, \quad [\text{T.14}]$$

where $\Omega_{\mathbf{a}}$ is an alternative-specific constant that represents the average utility of all unobserved factors that vary across alternatives $\mathbf{a} = \{i, r, j, v, q\}$ in $\mathcal{C}_n^T$; u_A is a subutility function (see below and in Section 3.5); u_B is the remaining part of the systematic utility that is strictly increasing and concave in consumption, C , residential size, H , and leisure time, ℓ ; $\varepsilon_{\mathbf{a}n}$ is the random part of the utility that is drawn from the i.i.d. type I extreme value distribution with joint cumulative density function:

$$F_{\varepsilon}(\varepsilon_{\mathbf{a}}^0, \dots, \varepsilon_{\mathbf{a}}^{\mathcal{N}(\mathcal{C}_n^T)}) = \exp(-e^{-\varepsilon_{\mathbf{a}}}), \quad [\text{T.15}]$$

where $\mathcal{N}(\mathcal{C}_n^T)$ is the size of the temporary choice set, *i.e.* the dimensions of the F_{ε} domain. This distribution gives rise to the logit choice probabilities, provided in equation [T.34].

The alternative specific constant can be decomposed as follows:

$$\Omega_{\mathbf{a}} = \Omega_i + \Omega_r + \Omega_j + \Omega_v + \Omega_q = 0. \quad [\text{T.16}]$$

Therefore, it is assumed to be additive in the choice elements that make up the composite alternative $\mathbf{a}$.

The subutility $u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n)$ is derived from observed characteristics varying across residential zones and residential types ($\mathbf{x}_{ir}$). These characteristics include a series of non-market environmental attributes such as: open space in the nearest proximity, visibility and exposure to sunlight and local population density. Since the level of these attributes is determined endogenously in the equilibrium, their discussion is postponed until Section C.2.

The non-linear part of the systematic utility is given by the Constant Elasticity of Substitution (CES) function:

$$u_B = (\alpha_C C_{an}^\rho + \alpha_H H_{an}^\rho + \alpha_F \ell_{an}^\rho)^{1/\rho}, \quad [\text{T.17}]$$

where the parameter ρ determines the elasticity of substitution (σ) between composite consumption, residential size and leisure via the relationship $\sigma = 1/(1 - \rho)$ and $\alpha = (\alpha_C, \alpha_H, \alpha_F)$ is the corresponding vector of CES shift parameters.

Equation [T.17] is maximized subject to the budget constraint:

$$C_{an} + p_{ir}^H H_{an} = e_n + w_{js_n} \bar{N}_L \bar{L} - c_a^{PK} - C_{vq}^T, \quad [\text{T.18}]$$

where the price of consumption has been normalized to one; e_n denotes the non-labor income of individual type n ; w_{js_n} is the per hour labor remuneration for the skill type s of individual n in employment zone j ; $\bar{N}_L$ is an exogenous annual labor supply in days; $\bar{L}$ is the exogenous duration of the working day (in hours); p_{ir}^H is the per m² annual rental rate of residential type r in zone i ; c_a^{PK} are the annual parking fees which are determined by residential location i , employment location j , vehicle type v and housing type r ; and C_{vq}^T are the annual transport costs of using route q and vehicle type v , where $v = 0$ denotes the choice of public transport. These costs are:

$$C_{vq}^T = \begin{cases} c_v^F + c_v^D + 2\bar{N}_L c_{vq}^V & \text{if } v \neq 0 \\ c_q^{PT} & \text{if } v = 0 \end{cases}, \quad [\text{T.19}]$$

where c_v^F and c_v^D are the annual fixed costs (registration-licensing and depreciation respectively) associated with the ownership of vehicle v ; c_{vq}^V is the pecuniary cost of one trip in route q with vehicle v ; c_q^{PT} is the annual cost of using route q with public transport. The per trip cost with vehicle v can be decomposed as:

$$c_{vq}^V = \tau_{qv}^{RP} + (c_{v,km}^U d_q^U + c_{v,km}^H d_q^H), \quad [\text{T.20}]$$

where τ_{qv}^{RP} are the per trip road-use fees, discussed in detail in Section D.3; $c_{v,km}^U$ and $c_{v,km}^H$ are the kilometer costs of urban and highway driving provided in [T.12]; d_q^U is the total distance covered in the implicit network, given by:

$$d_q^U = \begin{cases} 0 & \text{if } q \text{ is a highway route} \\ d(q) & \text{if } q \text{ is an urban route} \\ d(q_{in}^U) + d(q_{out}^U) & \text{if } q \text{ is a mixed route} \end{cases}, \quad [\text{T.21}]$$

where $d(q_{in}^U)$ and $d(q_{out}^U)$ for mixed routes can be computed from [T.4]. The distance covered in the highway network, d_q^H , can then be computed using:

$$d_q^H = d(q) - d_q^U, \quad [\text{T.22}]$$

The annual private cost of public transport is assumed to be a linear function of the commuting distance, *i.e.*:

$$c_q^{PT} = c_{km}^{PT} d(q), \quad [\text{T.23}]$$

where c_{km}^{PT} is the annual kilometre cost. The time constraint is:

$$\ell_{an} = \bar{T} - \bar{N}_L(\bar{L} + t_{rvq}), \quad [\text{T.24}]$$

where $\bar{T}$ is the total time endowment and t_{rvq} is the daily travel time given by:

$$t_{rvq} = \begin{cases} 2t_{vq} & \text{if } v \neq 0 \\ 2t_{ij}^{PT} & \text{if } v = 0 \end{cases}. \quad [\text{T.25}]$$

In equation [T.25], t_{vq} is the travel time of route q using vehicle v and t_{ij}^{PT} is the travel time required to move from residential zone i to employment zone j using public transport.

From the preference relation assumed in [T.17] and the constraints in [T.18] and [T.24] it follows that the alternative under examination, $\mathbf{a} \in \mathcal{C}_n^T$, is affordable if and only if the total income net of transportation costs and the resulting leisure time are both strictly positive. Letting the indicator $I(\mathbf{a}, n)$ take the value one if both conditions are satisfied, affordability is expressed by:

$$I(\mathbf{a}, n) = \begin{cases} 1 & \text{if } \underbrace{e_n + w_{js_n} \bar{N}_L \bar{L} - c_{\mathbf{a}}^{PK} - C_{vq}^T}_{\text{income net of transport}} > 0 \text{ and } \underbrace{\bar{T} - \bar{N}_L(\bar{L} + t_{rvq})}_{\text{resulting leisure time}} > 0 \\ 0 & \text{otherwise} \end{cases}. \quad [\text{T.26}]$$

This holds because consumption, residential space and leisure are all strictly positive (as required) only when both conditions in [T.26] are satisfied with strict inequality. That is, inserting [T.18] and [T.24] into [T.17], differentiating with respect to H and solving the first order condition for H yields the optimal residential size for the individual type n that chose alternative $\mathbf{a} = \{i, r, j, v, q\}$. This is:

$$H_{an}^* = \left(\frac{e_n + w_{js_n} \bar{N}_L \bar{L} - c_{\mathbf{a}}^{PK} - C_{vq}^T}{p_{ir}^H} \right) \underbrace{\left(\frac{1}{1 + (p_{ir}^H)^{\rho/(1-\rho)} \left(\frac{\alpha_C}{\alpha_H} \right)^{1/(1-\rho)}} \right)}_{\text{housing share of net expenditure}}, \quad [\text{T.27}]$$

where, unlike Cobb-Douglas preferences, the expenditure share of housing varies with housing costs, p_{ir}^H . Equation [T.27] is strictly positive only if $e_n + w_{js_n} \bar{N}_L \bar{L} - c_{\mathbf{a}}^{PK} - C_{vq}^T > 0$. The corresponding optimal consumption can subsequently be computed by inserting [T.27] into [T.18]. This is:

$$C_{an}^* = (e_n + w_{js_n} \bar{N}_L \bar{L} - c_{\mathbf{a}}^{PK} - C_{vq}^T) \underbrace{\left(\frac{(p_{ir}^H)^{\rho/(1-\rho)} \left(\frac{\alpha_C}{\alpha_H} \right)^{1/(1-\rho)}}{1 + (p_{ir}^H)^{\rho/(1-\rho)} \left(\frac{\alpha_C}{\alpha_H} \right)^{1/(1-\rho)}} \right)}_{\text{consumption share of net expenditure}}, \quad [\text{T.28}]$$

which again is strictly positive only if $e_n + w_{js_n} \bar{N}_L \bar{L} - c_a^{PK} - C_{vq}^T > 0$. The indirect utility corresponding to u_{Bn} can now be computed by inserting [T.27] and [T.28] into [T.17]. This yields:

$$V_{Ban}^* = \begin{cases} V_{Ban} & \text{if } I(\mathbf{a}, n) = 1 \\ -\infty & \text{if } I(\mathbf{a}, n) = 0 \end{cases} \quad [\text{T.29}]$$

where:

$$V_{Ban} = \left[g \left(\frac{e_n + w_{js_n} \bar{N}_L \bar{L} - c_a^{PK} - C_{vq}^T}{1 + (p_{ir}^H)^{\rho/(1-\rho)} \left(\frac{\alpha_C}{\alpha_H} \right)^{1/(1-\rho)}} \right)^\rho + \alpha_F \left(\bar{T} - \bar{N}_L (\bar{L} + t_{T_{vq}}) \right)^\rho \right]^{1/\rho}. \quad [\text{T.30}]$$

and g is an auxiliary function of parameters and prices, such that:

$$g = \left(\alpha_H (p_{ir}^H)^{-\rho} + \alpha_C \left((p_{ir}^H)^{\rho/(1-\rho)} \left(\frac{\alpha_C}{\alpha_H} \right)^{1/(1-\rho)} \right)^\rho \right). \quad [\text{T.31}]$$

Finally, the systematic utility corresponding to the maximization problem in [T.14] can be written as:

$$V_{an}^* = \Omega_a + u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n) + V_{Ban}^*, \quad [\text{T.32}]$$

re-expressing this way [T.14] as:

$$U_{an} = V_{an}^* + \varepsilon_{an}, \quad [\text{T.33}]$$

and the choice probability for the alternative $\mathbf{a}$ in $\mathcal{C}_n^T$, given the assumption over the distribution of the error terms in [T.15], can be expressed as:

$$P_{an} = \frac{e^{(V_{an}^*/\lambda)}}{\sum_{\mathbf{b} \in \mathcal{C}_n^T} (e^{(V_{bn}^*/\lambda)})}. \quad [\text{T.34}]$$

The reader should note that, as long as the utility function in [T.32] is assumed to be additive in the systematic and random part, changing the assumption over the distribution of the error terms in [T.15] changes the choice probability formula in [T.34] but leaves the deterministic part of the utility, derived in [T.16]-[T.32], intact.

B.3 Developers

The model assumes a representative developer in each residential zone i . The developer acquires each unit of land (m^2) at price p_{Xi} and faces a construction cost per m^2 (denoted by k^D) that is constant across residential development types r and the amount of floor space being developed. The net profit from choosing to develop one m^2 of land into a residential type r is:

$$\pi_{ir}^D = \begin{cases} \mu_r \xi_r (p_{ir}^H - k^D) - p_{Xi} & \text{if } I_P(i, r) = 1 \\ -\infty & \text{if } I_P(i, r) = 0 \end{cases}, \quad [\text{T.35}]$$

where $I_P(i, r)$ is the policy indicator that equals one if residential type r is allowed in zone i (and zero otherwise) and the product of the coverage (μ_r) and height (ξ_r) coefficients represents the amount of floor space that results per unit of land converted to development type r . This product is multiplied with $(p_{ir}^H - k^D)$ to give the gross profit per m² of land being developed. Subtracting the land remuneration from the gross profit yields the deterministic net profit in [T.35]. All unobserved factors that determine the profit but are not modelled explicitly are encapsulated into an error term ε_{ir}^D . Adding this error to the deterministic profit in [T.35] yields the conditional (on r) stochastic profit:

$$\hat{\pi}_{ir}^D = \pi_{ir}^D + \varepsilon_{ir}^D. \quad [\text{T.36}]$$

If the developer decides to keep the land undeveloped the deterministic profit will be negative and equal to the land costs:

$$\pi_{i0}^D = -p_{Xi}, \quad [\text{T.37}]$$

with the stochastic counterpart being equal to:

$$\hat{\pi}_{i0}^D = -p_{Xi} + \varepsilon_{i0}^D. \quad [\text{T.38}]$$

Assuming that ε_{ir}^D is i.i.d. type I extreme value distributed yields the expected maximum net profit in residential zone i . This is:

$$\tilde{\pi}_i^D = \lambda^D \left[\mathcal{E} + \log \left(e^{(\pi_{i0}^D/\lambda^D)} + \sum_r e^{(\pi_{ir}^D/\lambda^D)} \right) \right] = 0, \quad [\text{T.39}]$$

where $\mathcal{E}$ is the Euler's constant and λ^D is the scale parameter of the error term distribution. The zero profit condition in [T.39] requires that the equilibrium land price, p_{Xi} , should be such that all potential profits from the real estate sector will capitalize into it. Therefore, in an equilibrium land prices will reflect not only the characteristics that drive demand for housing in each zone i (e.g. accessibility and environmental quality), but also the active zoning policies that constrain the supply of floor space in each zone i .

The probability of converting one m² of land into housing type r in zone i is:

$$P_{i,r}^D = \frac{e^{(\pi_{ir}^D/\lambda^D)}}{e^{(\pi_{i0}^D/\lambda^D)} + \sum_r e^{(\pi_{ir}^D/\lambda^D)}}, \quad [\text{T.40}]$$

and the expected aggregate supply of type r floor space in zone i is:

$$AS_{ir}^H = (\mu_r \xi_r) P_{i,r}^D \underbrace{X_i (1 - \beta(i, A) - \beta(i, E) - \beta(i, I))}_{\text{developable land in zone } i}, \quad [\text{T.41}]$$

where the total amount of developable land in zone i is computed by solving equation [T.1] for the fraction of land that can be developed and then multiplying this fraction with the zone's total land surface, X_i . Subsequently, multiplying the developable surface with $P_{i,r}^D$ yields the expected footprint of type r residences in zone i . Multiplying this footprint with the coverage coefficient (μ_r) and the number of floors (ξ_r) yields the expected aggregate supply of type r floor space in zone i .

Apart from the surface $X_i(\beta(i, A) + \beta(i, E) + \beta(i, I))$ that is exogenously reserved for open space, infrastructure and other land uses, some part of the developable area $X_i(1 - \beta(i, A) - \beta(i, E) - \beta(i, I))$ will stay undeveloped. The developer's probability of keeping a unit of land vacant is:

$$P_{i,0}^D = \frac{e^{(\pi_{i0}^D/\lambda^D)}}{e^{(\pi_{i0}^D/\lambda^D)} + \sum_r e^{(\pi_{ir}^D/\lambda^D)}} = \beta(i, 0), \quad [\text{T.42}]$$

Therefore, the model allows for both exogenously and endogenously determined undeveloped urban fabric. For the arbitrary zone i , the latter is:

$$VS_i = P_{i,0}^D \underbrace{X_i(1 - u(i, A) - u(i, E) - u(i, I))}_{\text{developable land in zone } i}. \quad [\text{T.43}]$$

C. Peripheral modules

C.1 Transportation

The choice probabilities derived in Section B.2 can now be used to compute the aggregate level of traffic in the highway links and in the roads of the implicit urban network. The total vehicle load in the highway link $l_{z' \rightarrow z}^H$ is given by the traffic assignment equation:

$$TL(l_{z' \rightarrow z}^H) = \sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(q, l_{z' \rightarrow z}^H) \varphi_v P_{\mathbf{a}n} \right), \quad [\text{T.44}]$$

where the indicator $I(q, l_{z' \rightarrow z}^H)$ equals one if the chosen route q in alternative $\mathbf{a}$ includes the highway link $l_{z' \rightarrow z}^H$ (zero otherwise) and φ_v is the size of the private vehicle. The load in [T.44] consists of private vehicles only, since public transport is assumed to operate only in the implicit road network. The corresponding vehicle load in urban roads of the arbitrary zone z is given by the traffic assignment equation:

$$TL_z = \sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} \left(I(v) I(q, z) \varphi_v P_{\mathbf{a}n} + (1 - I(v)) I(q, z) \frac{\varphi_{PT}}{K_{PT}} P_{\mathbf{a}n} \right) \right), \quad [\text{T.45}]$$

where the ownership indicator $I(v)$ equals one if $v \neq 0$ (and zero if $v = 0$); the indicator $I(q, z)$ equals one if the chosen route q is either urban or mixed and passes through zone z (zero otherwise); φ_{PT} is the average size of public transport vehicles; K_{PT} is the average passenger capacity of public transport vehicles.

The total demand for type v vehicles is given by:

$$AD_v = \sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(\mathbf{a}, v) P_{\mathbf{a}n} \right), \quad [\text{T.46}]$$

where $I(\mathbf{a}, v)$ equals one if alternative $\mathbf{a}$ implies the choice of vehicle type v (zero otherwise). The aggregate demand for fuel type g from private vehicles is:

$$AD_g = \sum_n N_n \left(\sum_v (1 - I_v(v, e)) \left(\sum_{a \in \mathcal{C}_n^T} I(a, v) I(g, v) (d_q^U l_v^U + d_q^H l_v^H) P_{an} \right) \right), \quad [T.47]$$

where $I(a, v)$ equals one if the vehicle v under consideration is being chosen through the choice of alternative a (zero otherwise); the indicator function $I_v(v, e)$ equals one if vehicle v is electric and zero if it has an internal combustion engine; the indicator $I(g, v)$ equals one if vehicle v uses fuel type g (zero otherwise); l_v^U and l_v^H are the liters of gasoline consumed per kilometer driven by vehicle v in the implicit urban network and highways, respectively; d_q^U and d_q^H are the kilometres driven in the implicit urban network and highways, respectively.

The aggregate passenger kilometres performed in zone z by public transportation are:

$$PKM_z = \sum_n N_n \left(\sum_{a \in \mathcal{C}_n^T} ((1 - I(v)) I(q, z) d(q^U) b_{qz}^U P_{an}) \right), \quad [T.48]$$

where the product $d(q^U) b_{qz}^U$ represents the distance covered inside the territory of zone z under the choice of urban route q . This route is by definition urban (q^U) whenever $I(v) = 0$. Summing up across zones yields the aggregate passenger kilometres performed by public transportation. That is:

$$PKM = \sum_n N_n \left(\sum_{a \in \mathcal{C}_n^T} ((1 - I(v)) d(q^U) P_{an}) \right), \quad [T.49]$$

To illustrate the full capacity of MOLES for environmental and economic analysis, a crude representation of public transportation is provided in this section. To some extent, the following assumptions may be perceived by the reader as *ad-hoc*. The expansion of the present setting in order to account explicitly for a public transport operator and its behaviour is left as a future exercise.

The assumptions that follow allow the mapping of the aggregate passenger kilometres produced in the territory of each zone, PKM_z , into resulting costs without modelling explicitly the behaviour of the public transport operator. These costs are assumed to be concave and increasing in the aggregate passenger kilometres. The quadratic cost function:

$$C_{PTz} = \rho_0 + \rho_{1z} PKM_z + \rho_{2z} (PKM_z)^2, \quad [T.50]$$

yields the average zone-specific cost per passenger kilometer:

$$AC_{PTz} = \frac{\rho_0}{PKM_z} + \rho_{1z} + \rho_{2z} PKM_z, \quad [T.51]$$

and the marginal zone-specific cost per passenger kilometer:

$$MC_{PTz} = \rho_{1z} + 2\rho_{2z} PKM_z. \quad [T.52]$$

Therefore an endogenous degree of scale economies emerges; this is equal to AC_{PTz}/MC_{PTz} . Subtracting the kilometer price c_{km}^{PT} from [T.51], and multiplying with PKM_z yields the public transport deficit of zone z . The total public transport deficit, therefore, is:

$$D_{PT} = \sum_z (\rho_0 + \rho_{1z} PKM_z + \rho_{2z} PKM_z^2 - c_{km}^{PT} PKM_z), \quad [T.53]$$

C.2 Urban structure and environmental indicators

The discussion of the components making up the subutility function $u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n)$ in [T.14] was postponed because some of the variables composing $\mathbf{x}_{ir}$ are in turn functions of the choices made in Section B.2. The local *footprint density* (FD) indicator for residential zone i , FD_i , is the fraction of the zone's footprint, X_i , that remains undeveloped exogenously, through the exogenous policy variable $\beta(i, A)$, or endogenously, through decisions of the construction sector.

$$FD_i = \beta(i, A) + \underbrace{\beta(i, 0)}_{\text{vacant land}} + \underbrace{\sum_r \beta(i, r) \mu_r}_{\text{backyard spaces}}. \quad [T.54]$$

The above indicator may be valued positively (negatively) by individuals who are attracted (repelled) by low footprint density and open spaces in their proximity, as well as by individuals who appreciate (disparage) a relatively fragmented development dominated by vacant open spaces and large private backyards.

The *population density* (PD) of residential zone i is given by:

$$PD_i = \left(\underbrace{\sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(\mathbf{a}, i) P_{an} \right)}_{\text{population in zone } i} \right) / X_i, \quad [T.55]$$

where $I(\mathbf{a}, i)$ is the indicator function that equals one if the choice of alternative $\mathbf{a}$ implies the choice of residential zone i (zero otherwise). This indicator captures the preference towards higher or lower population density.

The *accessibility to natural amenities* of the entire urban area from zone i is given by:

$$NA_i = \sum_z \beta(z, A) X_z d_{zi}, \quad [T.56]$$

where d_{zi} is the Euclidean distance between zone i and an arbitrary zone z . Therefore, increasing the fraction of land reserved for natural amenities or open space in any region z will have a positive effect to the utility of those residing in zone i , but the effect decays linearly with distance to z .

Finally, *local visibility* and *exposure to sunlight* are captured by the indicator:

$$EV_{ir} = \xi_r - \left(\underbrace{\sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(\mathbf{a}, i, r) P_{an} H_{an}^* \right)}_{\text{floor-to-area ratio in } i} \right) / X_i, \quad [T.57]$$

which expresses the difference between the height of the chosen residential type r and the average building height in zone i . Therefore, increasing the latter imposes a negative externality to those valuing positively the above indicator.

The functional form of $u_A(\mathbf{x}_{ir}; \boldsymbol{\theta})$ is assumed to be linear in the preference parameters:

$$u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n) = \theta_n^{FD} FD_i + \theta_n^{PD} PD_i + \theta_n^{NA} NA_i + \theta_n^{EV} EV_{ir}, \quad [\text{T.58}]$$

and allows the various indicators to be valued differently by different individuals. The reader will note that $u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n)$ abstracts from local air quality indicators. In the current paper this abstraction is useful since it allows a relatively simpler documentation of the model's functions without directing the reader to the more complex air pollution propagation mechanisms.¹¹

In the context of this work, greenhouse gases (GHG) are not assumed to be a source of externality. However, the equilibrium amount of an arbitrary greenhouse gas can be approximated by multiplying the amount of each fossil fuel demanded in the equilibrium, provided in [T.47], with the amount of the greenhouse gas produced by each combusted liter of that fuel and then by summing over fuels. For instance, the total CO₂ will be:

$$\text{CO}_2 = \sum_g (AD_g E_{\text{CO}_2}^g). \quad [\text{T.59}]$$

Subsequently, the negative value of GHG emissions can be computed for different costs per unit of CO₂ assumed.

D. Policy instruments and exogenous variables

The model assumes a government that controls a series of tax-based policy instruments and regulatory mechanisms that are summarised in this section. The policy interventions are grouped into several categories: fuel taxes, fiscal treatment of private vehicles, road pricing, land-use regulations, land and property taxes and investments in road capacity and open space conservation.

D.1 Fuel taxation

The final price of fuel type g that is used to compute the kilometre cost of private vehicle use in equation [T.12], can be decomposed as:

$$p_g = p_g^{IM} + \tau_g, \quad [\text{T.60}]$$

where p_g^{IM} and τ_g denote the import price and the tax imposed on the consumption of fuel type g . The government revenue from fuel taxation is:

¹¹ Such mechanisms, which allow the mapping of the city's traffic profile (*i.e.* the aggregate traffic level in each zone z and each highway link) into the concentration of each pollutant in each location of the city, will be embodied in a future version of the model, designed specifically for the study of policies targeting air pollution.

$$TR_F = \sum_g \tau_g AD_g, \quad [\text{T.61}]$$

where AD_g is the aggregate demand of fuel type g given in [T.47] and $\tau_g AD_g$ is the revenue from fuel type g .

D.2 Fiscal treatment of private vehicles

Similarly, the government controls the annual vehicle registration and licensing charges, denoted by c_v^F . The total revenue from these charges is:

$$TR_V = \underbrace{\sum_v I_v(v, e) c_v^F AD_v}_{\text{aggregate tax or subsidy to electric vehicles}} + \underbrace{\sum_v (1 - I_v(v, e)) c_v^F AD_v}_{\text{aggregate tax or subsidy to internal combustion engine vehicles}}. \quad [\text{T.62}]$$

the indicator function $I_v(v, e)$ equals one if vehicle v is electric and zero if it has an internal combustion engine.

D.3 Road pricing

The road-use fee per commuting trip mentioned in **Section B.2**, τ_{qv}^{RP} , can take various forms according to the pricing scheme applied. Under a vehicle type-specific flat kilometre tax, $\bar{\tau}_v^{FK}$, it holds that:

$$\tau_{qv}^{RP} = \bar{\tau}_v^{FK} d(q) = \bar{\tau}_v^{FK} (d_q^H + d_q^U). \quad [\text{T.63}]$$

A technologically more demanding scheme, the varying kilometre charge, could spatially differentiate the burden. For a mixed route $q^M = (q_{in}^U, q^H, q_{out}^U)$ the charge could be:

$$\tau_{qv}^{RP} = d(q_{in}^U) \sum_z (\tau_{zv} b_{q_{in,z}^U}) + d(q_{out}^U) \sum_z (\tau_{zv} b_{q_{out,z}^U}) + \sum_{l_{z' \rightarrow z}^H \in q^H} \{(d_{z' \rightarrow z}^H) \tau_{lv}\}, \quad [\text{T.64}]$$

where the first sum is the total charge for the distance covered from the trip origin to the highway entrance and the second sum is the total charge for the distance covered from the highway exit to the trip destination, both determined with a zone- and vehicle type-specific kilometre charge, τ_{zv} ; finally, the third term represents the total charge for the distance covered in the highway, determined with a link- and vehicle type-specific kilometre charge, τ_{lv} (see **Section A.2.2** for the interpretation of terms $b_{q_z^U}$). For a purely urban route, the above charge becomes:

$$\tau_{qv}^{RP} = d(q^U) \sum_z (\tau_{zv} b_{q_z^U}). \quad [\text{T.65}]$$

A cordon toll can be expressed as a special case of [T.64] and [T.65], by choosing to charge only a number of selected zones and highway links that form a bounded area, to which access from outside is possible only with a predetermined fee.

D.4 Zoning, height restrictions, boundaries and low emission zones

MOLES is designed to evaluate the impact of key urban planning instruments. Zoning is performed by setting values (zero or one) to the various housing type zoning indicators $I_p(i, r)$. This allows for a direct control of the minimum and maximum building height and the maximum floor space that can be produced in a residential zone i . It also allows for partial control of the average footprint coverage and building height by altering the choice sets of the developers. Complete zoning of a residential area i is possible by setting each $I_p(i, r)$ equal to zero. The urban growth boundary is determined implicitly by restricting any type of development in periurban undeveloped areas.

The land-use patterns in employment zones and the emissions from buildings used by the industrial, commercial and service sectors are treated as exogenous, for reasons explained in **Section 2.6**. However, the impact of these emissions can be altered by industry relocation and other policies giving rise to near-zero emission zones. The conversion of an employment zone to residential zone can be by moving an arbitrary employment zone j from the set to employment zones to the set of residential zones and by adding a new zone j' (of footprint equal to that of area j) outside the city boundaries.

Low emission areas and car-free zones may be considered by replacing all the urban paths by detour paths that circumvent a restricted area. A *detour path* to circumvent zone z can be defined as a sequence of urban routes $(q^{U1}, q^{U2}, \dots, q^{UD})$ such that the origin of q^{Ud} is the destination of q^{Ud-1} and $b_{q^{Ud}} = 0$ for all d .

D.5 Land and property taxation

The land and housing prices mentioned in Section B.2 and B.3 may incorporate a tax component. For instance, a differentiated by property type and zone *ad-valorem* tax, τ_{ir}^H , incurred by the individuals translates into:

$$p_{ir}^H = (1 + \tau_{ir}^H) \hat{p}_{ir}^H. \quad [\text{T.66}]$$

A similar modification would be relevant for an *ad-valorem* tax on developed land.

D.6 Road capacity expansion and open space conservation

Investments in road capacity can be captured by adjusting capacity $RK(l_{z'z}^H)$ in the highway segment $l_{z' \rightarrow z}^H$ and adjustment of the fraction of land occupied by urban road infrastructure in zone m , $\beta(m, I)$. Similarly, the total exogenous open space is adjusted through $\beta(i, A)$. Highway and rail networks can be expanded by adding new links in the existing ones.

D.7 Exogenous variables

MOLES does not model explicitly the price setting mechanisms in the markets of vehicles and fuel, as these are more relevant for analyses at the national or international level. Therefore vehicle and fuel prices are introduced directly from scenarios and serve as exogenous variables in the model. A similar argument is valid for the labor market, as modeling explicit wage setting mechanisms would require a complete model of firm behaviour. This could be a formidable task, as firm output levels (and therefore employment and wages) usually depend on demand and regulations determined at the national and, very often, international level. Therefore, the real wage by skill and location evolves exogenously. Other exogenous input to the model includes variables such as city population growth and the population composition, *i.e.* the percentage of each individual type in the city population.

E. General, stochastic user equilibrium

E.1 Closing conditions

The choice probabilities derived in **Section B.2** can also be used together with the conditional demands to compute the aggregate demands for the various housing types in residential zones. The aggregate demand for floor space of residential type r in zone i is:

$$AD_{ir}^H = \sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(\mathbf{a}, i, r) P_{\mathbf{a}n} H_{\mathbf{a}n}^* \right), \quad [\text{T.67}]$$

where N_n is the population of type n individuals and the indicator function $I(\mathbf{a}, i, r)$ equals one if alternative $\mathbf{a}$ implies the choice of residential zone i and housing type r (zero otherwise).

Using [T.67] and [T.41], the excess demand function for housing type r in residential zone i can be expressed as:

$$\sum_n N_n \left(\sum_{\mathbf{a} \in \mathcal{C}_n^T} I(\mathbf{a}, i, r) P_{\mathbf{a}n} H_{\mathbf{a}n}^* \right) - (\mu_r \xi_r) P_{i,r}^D X_i (1 - \beta(i, A) - \beta(i, E) - \beta(i, I)) = 0. \quad [\text{T.68}]$$

Furthermore, in equilibrium the profits of the construction sector should be zero, reflecting free entry and a profit absorbing price setting mechanism in land markets. The total land income generated in the equilibrium is:

$$\Theta = \sum_i p_{Xi} X_i (1 - \beta(i, A) - \beta(i, E) - \beta(i, I)). \quad [\text{T.69}]$$

The equilibrium non-labor income of individual type n that appears in equations [T.18] and [T.26]-[T.29] will be:

$$e_n = d_{Ln} \Theta + \bar{M}_n. \quad [\text{T.70}]$$

where d_{Ln} is the land dividend of individual type n (with $\sum_n d_{Ln} \leq 1$) and $\bar{M}_n$ is the exogenous income stream from other sources (*e.g.* capital income).

E.2 System of equations

This section describes the system of equations and unknowns that characterise the *general, stochastic user equilibrium* without policy intervention. To facilitate exposition, Table E.1 elaborates on the various equation counters used throughout the text. Solving the highly non-linear system described below for the general equilibrium requires very good initial values of all endogenous variables, but even when these are available convergence will not be guaranteed. Furthermore, the values of the endogenous variables may differ substantially between equilibria supported by different policies rendering convergence from one equilibrium to the other an uncertain process. To mitigate the computational turbulence that can occur in these cases, the paper proposes a high-modularity algorithm which is developed upon object-oriented programming principles, but should be also possible to implement with functional programming techniques. The proposed algorithm is based on a partition of the square system (28 blocks of equations in

28 vectors of unknowns) in five subsystems. That partition is by no means arbitrary. As shown later, it allows for a smooth and fast convergence to the general, stochastic user equilibrium.

Table E.1 Equation counters in MOLES

<i>Counter</i>	<i>Description</i>
I	The total number of land markets and residential zones.
$I + J$	The total number of zones in the model.
$\sum_i I_P(i, r)$	The total number of floor space markets: one for each type r allowed in each residential zone i .
N	The number of individual types
$\mathcal{N}(\mathcal{C}_n^T)$	Number of <i>ex-ante</i> compatible (but not necessarily economically affordable) alternatives forming the <i>temporary</i> choice set of individual n .
$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$	Total number of <i>ex-ante</i> compatible (but not necessarily economically affordable) alternatives in the model.
$\mathcal{N}(l^H)$	Total number of links in the highway network.
$\mathcal{N}(q^H)$	Total number of highway routes in the model.
$(I + J)^2$	Total number of urban routes in the model.
$\mathcal{N}(q^M)$	Total number of mixed routes in the model.

E.2.1 Equations

The *household subsystem* (denoted by $\mathbf{f}_A$) consists of: N non-labor income equations as the one in [T.70]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ conditional housing demands equations as in [T.27]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ conditional consumption levels equations as in [T.28]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ conditional leisure time equations as in [T.24]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ indirect utility levels as in [T.29]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ systematic utilities as in [T.32]; $\sum_n \mathcal{N}(\mathcal{C}_n^T)$ choice probability equations as in [T.34]. The subsystem is summarised on Table E.2.

Table E.2 Sub-system solved by Module A

<i>Equation group</i>	<i>Number of equations</i>	<i>Equation group</i>	<i>Number of equations</i>
[T.27]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$	[T.29]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$
[T.28]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$	[T.32]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$
[T.24]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$	[T.70]	N
[T.34]	$\sum_n (\mathcal{N}(\mathcal{C}_n^T))$		

The *real estate subsystem* (denoted by $\mathbf{f}_B$) consists of: $\sum_i I_P(i, r)$ conditional deterministic profit equations as in [T.35]; I conditional deterministic loss equations as in [T.37]; $\sum_i I_P(i, r)$ development probability equations as in [T.40]; I “postpone-development” probabilities as in [T.42]. The subsystem is summarised on Table E.3.

Table E.3 Sub-system solved by Module B

<i>Equation group</i>	<i>Number of equations</i>	<i>Equation group</i>	<i>Number of equations</i>
[T.35]	$\sum_i I_p(i, r)$	[T.40]	$\sum_i I_p(i, r)$
[T.37]	I	[T.42]	I

The subsystem representing the *economic and land-use equilibrium* (denoted by $\mathbf{f}_C$) consists of: $\sum_i I_p(i, r)$ excess demand equations as in [T.68], I zero profit conditions as in [T.39] and the total land income equation in [T.69]. The core system is summarised on Table E.4.

Table E.4 Sub-system solved by the core module

<i>Equation group</i>	<i>Number of equations</i>
[T.68]	$\sum_i I_p(i, r)$
[T.39]	I
[T.69]	1

The subsystem representing the *stochastic user equilibrium* in transportation (denoted by $\mathbf{f}_D$) consists of: $\mathcal{N}(l^H)$ highway traffic equations as in [T.44]; $(I + J)$ implicit network traffic equations as in [T.45]; $\mathcal{N}(l^H)$ highway speed equations as in [T.5]; $(I + J)$ zonal speed equations for car as in [T.7]; $(I + J)$ zonal speed equations for public transport as in [T.8]; $\mathcal{N}(q^H)$ highway route travel time equations as in [T.6]; $(I + J)^2$ urban route travel time equations for private vehicles, as in [T.9]; $(I + J)^2$ urban route travel time equations for public transport, as in [T.10]; $\mathcal{N}(q^M)$ mixed route travel time equations as in [T.11]. Subsystem D is summarised on Table E.5.

Table E.5 Sub-system solved by the transportation module

<i>Equation group</i>	<i>Number of equations</i>	<i>Equation group</i>	<i>Number of equations</i>
[T.44]	$\mathcal{N}(l^H)$	[T.6]	$\mathcal{N}(q^H)$
[T.45]	$I + J$	[T.9]	$(I + J)^2$
[T.5]	$\mathcal{N}(l^H)$	[T.10]	$(I + J)^2$
[T.7]	$I + J$	[T.11]	$\mathcal{N}(q^M)$
[T.8]	$I + J$		

The subsystem representing the *environmental equilibrium* (denoted by $\mathbf{f}_E$) consists of: I footprint density equations as in [T.54]; I population density equations as in [T.55]; I natural amenity accessibility equations as in [T.56]; $\sum_i I_p(i, r)$ visibility and sunlight exposure equations as in [T.57]; $\sum_n \mathcal{N}(C_n^T)$ subutility equations as in [T.58]. Subsystem E is summarised on Table E.6.

Table E.6 Sub-system solved by the urban morphology & environment module

Equation group	Number of equations	Equation group	Number of equations
[T.54]	I	[T.57]	$\sum_i I_P(i, r)$
[T.55]	I	[T.58]	$N \sum_i I_P(i, r)$
[T.56]	I		

E.2.2. Endogenous variables

The corresponding 28 blocks of endogenous variables, are grouped into five blocks. Group $\mathbf{x}_1$ (see Table E.7) consists of: non-labor incomes e_n ; the conditional residential sizes H_{an}^* ; the conditional consumption levels C_{an}^* ; the conditional leisure times ℓ_{an} ; the choice probabilities P_{an} ; the indirect utilities V_{Ban}^* and the systematic utilities V_{an}^* . Group $\mathbf{x}_2$ (see Table E.8) consists of: the conditional deterministic profits π_{ir}^D ; the conditional deterministic losses π_{i0}^D ; the development probabilities $P_{i,r}^D$ and the postponed development probabilities $P_{i,0}^D$. Group $\mathbf{x}_3$ (see Table E.9) consists of: the housing prices p_{ir}^H , the land prices p_{xi} and the total land income Θ . Group $\mathbf{x}_4$ (see Table E.10) consists of: highway network link loads $TL(l_{z,z'}^H)$; implicit network zonal loads TL_z ; highway link speeds $SP(l_{z' \rightarrow z}^H)$; zonal car speeds $SP_m^{U,C}$; zonal public transport speeds $SP_m^{U,PT}$; highway route travel times $t(q^H)$; urban route travel times for private vehicles $t_c(q^U)$; urban route travel times for public transportation $t_{PT}(q^U)$; mixed route travel times for private vehicles $t_c(q^M)$. Finally, group $\mathbf{x}_5$ (see Table E.11) consists: of the footprint density indicators FD_i ; the population density indicators PD_i ; the natural amenity accessibility indicators NA_i ; the visibility and sunlight exposure indicators EV_{ir} .

Table E.7 Endogenous variables block 1 ($\mathbf{x}_1$)

Endogenous variable	Number of variables	Endogenous variable	Number of variables
H_{an}^*	$\sum_n (\mathcal{N}(C_n^T))$	e_n	N
C_{an}^*	$\sum_n (\mathcal{N}(C_n^T))$	P_{an}	$\sum_n (\mathcal{N}(C_n^T))$
ℓ_{an}	$\sum_n (\mathcal{N}(C_n^T))$	V_{Ban}^*	$\sum_n (\mathcal{N}(C_n^T))$
V_{an}^*	$\sum_n (\mathcal{N}(C_n^T))$		

Table E.8 Endogenous variables block 2 ($\mathbf{x}_2$)

Endogenous variable	Number of variables	Endogenous variable	Number of variables
π_{ir}^D	$\sum_i I_P(i, r)$	π_{i0}^D	I
$P_{i,r}^D$	$\sum_i I_P(i, r)$	$P_{i,0}^D$	I

Table E.9 Endogenous variables block 3 ($\mathbf{x}_3$)

Endogenous variable	Number of variables
p_{ir}^H	$\sum_i I_P(i, r)$
p_{Xi}	I
Θ	1

Table E.10 Endogenous variables block 4 ($\mathbf{x}_4$)

Endogenous variable	Number of variables	Endogenous variable	Number of variables
$TL(l_{z' \rightarrow z}^H)$	$\mathcal{N}(l^H)$	$t_C(q^H)$	$\mathcal{N}(q^H)$
TL_z	$I + J$	$t_C(q^U)$	$(I + J)^2$
$SP(l_{z' \rightarrow z}^H)$	$\mathcal{N}(l^H)$	$t_{PT}(q^U)$	$(I + J)^2$
$SP_m^{U,C}$	$I + J$	$t_C(q^M)$	$\mathcal{N}(q^M)$
$SP_m^{U,PT}$	$I + J$		

Table E.11 Endogenous variables block 5 ($\mathbf{x}_5$)

Endogenous variable	Number of variables	Endogenous variable	Number of variables
FD_i	I	EV_{ir}	$\sum_i I_P(i, r)$
PD_i	I	$u_A(\mathbf{x}_{ir}; \boldsymbol{\theta}_n)$	$N \sum_i I_P(i, r)$
NA_i	I		

E.2.3 System representation

The reader can now verify that the system described above can be written in compact form as:

$$\begin{aligned}
 \mathbf{e}_A &= \mathbf{f}_A(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5) \\
 \mathbf{e}_B &= \mathbf{f}_B(\mathbf{x}_2, \mathbf{x}_3) \\
 \mathbf{e}_C &= \mathbf{f}_C(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
 \mathbf{e}_D &= \mathbf{f}_D(\mathbf{x}_1, \mathbf{x}_3) \\
 \mathbf{e}_E &= \mathbf{f}_E(\mathbf{x}_1, \mathbf{x}_4)
 \end{aligned}
 \tag{T.71}$$

E.3. Solution

E.3.1 Auxiliary modules A and B

Together with the exogenous variables and parameters, vectors $\mathbf{x}_3$, $\mathbf{x}_4$ and $\mathbf{x}_5$ contain sufficient information to solve subsystem A. This is, there is a mapping $\mathbf{g}_A$ such that $\mathbf{x}_1 = \mathbf{g}_A(\mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5)$ and $\mathbf{e}_A = \mathbf{f}_A(\mathbf{g}_A(\mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5), \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5) = \mathbf{0}$. This mapping from $\mathbb{R}_+^{(\omega_3 + \omega_4 + \omega_5)} \rightarrow \mathbb{R}_+^{(\omega_1)}$, (where ω_i denotes the number of elements in vector $\mathbf{x}_i$) is performed by module A (see left panel of Figure E.1), which solves the household optimisation problem for each alternative in the temporary choice set of each individual,

returning choice probabilities and the rest of endogenous variables in $\mathbf{x}_1$. Similarly, vector $\mathbf{x}_3$ contains sufficient information to solve subsystem B. Again, there is a mapping $\mathbf{g}_B$ such that $\mathbf{x}_2 = \mathbf{g}_B(\mathbf{x}_3)$ for which $\mathbf{e}_B = \mathbf{f}_B(\mathbf{g}_B(\mathbf{x}_3), \mathbf{x}_3) = \mathbf{0}$. That mapping from $\mathbb{R}_+^{(\omega_3)} \rightarrow \mathbb{R}_+^{(\omega_2)}$ is performed by module B (see right panel of Figure E.1), which solves the developer's optimisation problem, using the prices embodied in $\mathbf{x}_3$ to return the various choice probabilities of development in $\mathbf{x}_2$.

That is, for every alternative in the temporary choice set of an individual type, Module A uses the travel time, non-labor income guesses, the exogenous wages and travel costs to compute equations [T.19]-[T.25]. It then evaluates the inclusion condition in [T.26]. If the temporary alternative is economically affordable for individual n , the associated optimal residential size and consumption are computed from [T.27] and [T.28]. Subsequently, indirect utility is computed using sequentially [T.31], [T.30] and [T.29]. If on the other hand the temporary alternative is not affordable the above computations are skipped and indirect utility is set to $-\infty$ using directly [T.29]. Then, the rest of the relevant information is retrieved in order to compute [T.32]. Finally, the local variable x is incremented by $e^{(V_{an}^*/\lambda)}$. When the inner loop is completed, that auxiliary variable obtains a value equal to this of the denominator in [T.34]. The numerator has also been computed for each alternative using [T.32], so going through the temporary choice set once more using [T.34] yields the corresponding choice probabilities. The entire process is repeated for each individual type.

For every residential type allowed to be developed in a given zone, Module B uses the land and housing prices in vector $\mathbf{x}_3$ to compute the deterministic profit in [T.35]. Then, the local variable x is incremented by $e^{(\pi_{ir}^D/\lambda^D)}$. If on the other hand the residential type is forbidden then profit is set to $-\infty$. When the loop is completed, Module B computes the profit from postponing development given in [T.37] and increments the local variable x by $e^{(\pi_{io}^D/\lambda^D)}$. After that, the value of x is equal to that of the denominator in [T.40] and [T.42]. Also, this denominator is embodied into the expected maximum profit formula, so the latter can then be computed using [T.39]. Going through each development type again, [T.40] can yield the probability of each residential type to be developed in the zone. Using [T.41], Module B computes the aggregate supply of each type. When the loop is completed, the probability of postponing development is computed using [T.42] and the aggregate amount of land staying undeveloped is computed from [T.43]. The entire process is repeated for each residential zone.

Figure E.1 Chart flows for modules A and B

Figure E.2 Left: chart flow of the network loading. Right: chart flow of module T

E.3.2 Solving the Core model: economic and land use equilibrium

Whenever module A and module B are performed, the system reduces to:

$$\begin{aligned} \mathbf{e}_C &= \mathbf{f}_C(\mathbf{g}_A(\mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5), \mathbf{g}_B(\mathbf{x}_3), \mathbf{x}_3) \\ \mathbf{e}_D &= \mathbf{f}_D(\mathbf{g}_A(\mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5), \mathbf{x}_4) \\ \mathbf{e}_E &= \mathbf{f}_E(\mathbf{g}_A(\mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5), \mathbf{x}_5) \end{aligned} \quad [\text{T.72}]$$

It is now possible to solve subsystem $\mathbf{f}_C$ iteratively for $\mathbf{x}_3$ using any fixed set of travel times and other transport-related variables $\bar{\mathbf{x}}_4$, as well as environmental and land-use pattern variables $\bar{\mathbf{x}}_5$. This solution is referred to as *economic and land-use equilibrium* and can be performed, for instance, using the Newton method. At the k -th iteration, the adjustment will be:

$$\mathbf{x}_{3(k+1)} - \mathbf{x}_{3(k)} = -s_k \mathbf{J}_C(\mathbf{x}_{3(k)}, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)^{-1} \mathbf{e}_C, \quad [\text{T.73}]$$

where s_k is the optimal step computed from any separate line search method and $\mathbf{J}_C(\mathbf{x}_{3(k)}, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$ is the Jacobian matrix of subsystem $\mathbf{f}_C$. The i -th column of the Jacobian matrix, i.e. the vector $\mathbf{J}_{Ci}$, can be approximated using the finite difference:

$$\mathbf{J}_{Ci} \approx \frac{\mathbf{e}'_{Ci} - \mathbf{e}_C}{\varepsilon} = \frac{\mathbf{f}_C(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}_3 + \Delta \mathbf{x}_{3i}) - \mathbf{f}_C(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)}{\varepsilon}, \quad [\text{T.74}]$$

where $\mathbf{x}'_1 = \mathbf{g}_A(\mathbf{x}_3 + \Delta \mathbf{x}_{3i}, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$, $\mathbf{x}'_2 = \mathbf{g}_B(\mathbf{x}_3 + \Delta \mathbf{x}_{3i})$, $\mathbf{x}_1 = \mathbf{g}_A(\mathbf{x}_3, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$, $\mathbf{x}_2 = \mathbf{g}_B(\mathbf{x}_3)$ and $\Delta \mathbf{x}_{3i}$ is the vector whose i -th element equals the perturbation ε (the rest are zero). Thus, to evaluate the difference in the numerator, modules A and B are called to re-map through $\mathbf{g}_A$ and $\mathbf{g}_B$ the new point $(\mathbf{x}_3 + \Delta \mathbf{x}_{3i}, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$, i.e. to compute the new choice probabilities and demands for households and new development probabilities for the real estate sector. Then, these choice probabilities and quantities are inserted in $\mathbf{f}_C$ yielding $\mathbf{e}'_{Ci}$: new excess demands, profits and a new value for the excess land income

equation in [T.69]. Repeating for each column yields $\mathbf{J}_C(\mathbf{x}_{3(k)}, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$, whose inverse can be used in [T.73] in order to iterate forward. Iterations continue until a stopping criterion is satisfied.

E.3.3 Stochastic-user and environmental equilibria

Denoting the solution vector of the economic and land-use subsystem by $\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$ the system becomes:

$$\begin{aligned} \mathbf{e}_C &= \mathbf{f}_C(\mathbf{g}_A(\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \mathbf{g}_B(\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)), \mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)) = \mathbf{0} \\ \mathbf{e}_D &= \mathbf{f}_D(\mathbf{g}_A(\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \bar{\mathbf{x}}_4) \\ \mathbf{e}_E &= \mathbf{f}_E(\mathbf{g}_A(\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5), \bar{\mathbf{x}}_5) \end{aligned} \quad [\text{T.75}]$$

The subsystems D and E, can be solved simultaneously for a fixed value of $\mathbf{x}_3$ if there are mappings $\mathbf{g}_T$ (from $\mathbb{R}_+^{(\omega_4)} \rightarrow \mathbb{R}_+^{(\omega_4)}$) and $\mathbf{g}_E$ (from $\mathbb{R}_+^{(\omega_5)} \rightarrow \mathbb{R}_+^{(\omega_5)}$) such that:

$$\begin{aligned} \mathbf{e}_D &= \mathbf{x}_4 - \mathbf{g}_T(\mathbf{g}_A(\bar{\mathbf{x}}_3, \mathbf{x}_4, \bar{\mathbf{x}}_5)) \\ \mathbf{e}_E &= \mathbf{x}_5 - \mathbf{g}_E(\mathbf{g}_A(\bar{\mathbf{x}}_3, \mathbf{x}_4, \bar{\mathbf{x}}_5)) \end{aligned} \quad [\text{T.76}]$$

Figure E.3 Chart flow of the traffic assignment module

These mappings are performed through modules T (which is displayed in the right panel of Figure E.2) and module E. Each fixed point iteration for the stochastic user equilibrium in transport starts with a set of assumed link (zonal) flows, speeds and route travel times. The right panel of Figure E.2 refers to these input values as $\mathbf{x}_4^{IN}$. Then, module A is called to solve the household optimisation problem that

returns a new vector $\mathbf{x}_1 = \mathbf{g}_A(\bar{\mathbf{x}}_3, \mathbf{x}_4^{IN}, \bar{\mathbf{x}}_5)$, that contains the updated route and mode choice probabilities. Then, module T calls a traffic assignment module which allocates drivers in the network, computes the demand for each highway link and for the implicit road network of each urban zone. The flowchart of the traffic assignment module is provided in Figure E.3. When that process is complete, Module T calls a network loading module which uses the volume delay functions to update link and zone speeds (left panel of Figure E.2). It then aggregates these speeds to produce new route travel times. The new flows, speeds and travel times are denoted by $\mathbf{x}_4^{OUT} = \mathbf{g}_T(\mathbf{g}_A(\bar{\mathbf{x}}_3, \mathbf{x}_4^{IN}, \bar{\mathbf{x}}_5))$. Finally, module T evaluates whether $\mathbf{x}_4$ has converged to a fixed point, i.e. whether $\mathbf{x}_4^{OUT}$ is sufficiently close to $\mathbf{x}_4^{IN}$. If that is case, the fixed point iteration is terminated. If not, the resulting speeds, flows and travel times ($\mathbf{x}_4^{OUT}$) become the input of the next iteration and Module T reiterates.

Module E (which has a similar structure to that of Module T) performs the respective fixed point iteration for the environmental and land-use part of the model. Starting with some initial value of the environmental and land-use indicator values, $\mathbf{x}_5^{IN}$, Module E calls Module A, which returns $\mathbf{x}_1 = \mathbf{g}_A(\bar{\mathbf{x}}_3, \bar{\mathbf{x}}_4, \mathbf{x}_5^{IN})$. This vector contains (among others) the new choice probabilities for residential zones and types, as well as the corresponding conditional demands for floor space. This is all the information needed to update the indicators in [T.54]-[T.57], and then use the new values to compute the utility each individual type derives from these indicators, using equation [T.58]. This process yields $\mathbf{x}_5^{OUT} = \mathbf{g}_E(\mathbf{g}_A(\bar{\mathbf{x}}_3, \bar{\mathbf{x}}_4, \mathbf{x}_5^{IN}))$. The fixed point iteration is terminated if the update is sufficiently small. While the loop is executed, the resulting values of the indicators and the utilities derived ($\mathbf{x}_5^{OUT}$) become the input of the next iteration and Module E reiterates.

E.3.4 MOLES equilibrium

The MOLES equilibrium is reached with a set of vectors $(\mathbf{x}_3^*, \mathbf{x}_4^*, \mathbf{x}_5^*)$ such that:

$$\begin{aligned} \mathbf{e}_C &= \mathbf{f}_C(\mathbf{g}_A(\mathbf{x}_3^*(\mathbf{x}_4^*, \mathbf{x}_5^*), \mathbf{x}_4^*, \mathbf{x}_5^*), \mathbf{g}_B(\mathbf{x}_3^*(\mathbf{x}_4^*, \mathbf{x}_5^*)), \mathbf{x}_3^*(\mathbf{x}_4^*, \mathbf{x}_5^*)) = \mathbf{0} \\ \mathbf{e}_D &= \mathbf{x}_4^* - \mathbf{g}_T(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}_4^*, \mathbf{x}_5^*)) = \mathbf{0} \\ \mathbf{e}_E &= \mathbf{x}_5^* - \mathbf{g}_E(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}_4^*, \mathbf{x}_5^*)) = \mathbf{0} \end{aligned} \quad [\text{T.77}]$$

is found. Figure E.4 summarises the solution algorithm. That is, MOLES begins with initial guesses of $\mathbf{x}_3, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5$ and solves the economic and land-use equilibrium for $\mathbf{x}_3^*(\bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)$. After that, it evaluates whether $\bar{\mathbf{x}}_4$ is such that $\mathbf{e}_D = \bar{\mathbf{x}}_4 - \mathbf{g}_T(\mathbf{g}_A(\mathbf{x}_3^*, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)) = \mathbf{0}$. That is, MOLES evaluates whether a stochastic user equilibrium in transport has been reached. If that is true, MOLES evaluates whether $\bar{\mathbf{x}}_5$ solves the system $\mathbf{e}_E = \bar{\mathbf{x}}_5 - \mathbf{g}_E(\mathbf{g}_A(\mathbf{x}_3^*, \bar{\mathbf{x}}_4, \bar{\mathbf{x}}_5)) = \mathbf{0}$, i.e. whether an environmental equilibrium has been reached. If this is the case as well, MOLES terminates. In any other case MOLES initiates a fixed point iteration, calling initially Module T, that upon convergence updates the value of $\mathbf{x}_4$ from $\bar{\mathbf{x}}_4$ to $\mathbf{x}_4'$. Having ensured that $\mathbf{x}_4'$ is such that $\mathbf{e}_D = \mathbf{x}_4' - \mathbf{g}_T(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}_4', \bar{\mathbf{x}}_5)) = \mathbf{0}$, the next step is to evaluate whether $\mathbf{e}_E = \bar{\mathbf{x}}_5 - \mathbf{g}_E(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}_4', \bar{\mathbf{x}}_5)) = \mathbf{0}$. If that is the case, the fixed point iteration is terminated, returning the updated values $(\mathbf{x}_4', \bar{\mathbf{x}}_5)$. If not, MOLES calls Module E to find a new fixed point by solving $\mathbf{e}_D = \mathbf{x}_5' - \mathbf{g}_E(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}_4', \mathbf{x}_5')) = \mathbf{0}$.

Figure E.4 The MOLES solution algorithm

Then, it checks whether $\mathbf{e}_D = \mathbf{x}'_4 - \mathbf{g}_T(\mathbf{g}_A(\mathbf{x}_3^*, \mathbf{x}'_4, \mathbf{x}'_5)) = \mathbf{0}$. The program continues iterating between Module T and Module E until one of the updates on $\mathbf{x}_4$ or $\mathbf{x}_5$ is small enough to push $\mathbf{e}_E$ or $\mathbf{e}_D$ to a considerable distance away from zero. At that point MOLES terminates the fixed point iteration and uses the outcome $(\bar{\mathbf{x}}'_4, \bar{\mathbf{x}}'_5)$ to evaluate $\mathbf{e}_C = \mathbf{f}_C(\mathbf{g}_A(\mathbf{x}_3^*, \bar{\mathbf{x}}'_4, \bar{\mathbf{x}}'_5), \mathbf{g}_B(\mathbf{x}_3^*), \mathbf{x}_3^*)$. If $\mathbf{e}_C$ is still sufficiently close to $\mathbf{0}$, the updates of vectors $\mathbf{x}_4$ and $\mathbf{x}_5$ during the fixed point iteration were too small to induce changes in the excess demand functions, zero profit conditions and the excess land income, i.e. in the core subsystem $\mathbf{f}_C$. Then, the triple equilibrium has been achieved and MOLES terminates. But if the updated travel time, environmental and land-use indicators (*i.e.* the new values $\bar{\mathbf{x}}'_4, \bar{\mathbf{x}}'_5$) induce a significant shock on subsystem $\mathbf{f}_C$, the old endogenous vector $\mathbf{x}_3^*$ (which encapsulates prices and the land income) is no longer a solution to $\mathbf{f}_C$ and MOLES returns to its starting point. Then, a new cycle begins by solving the economic and land-use model for $\mathbf{x}_3^{**}(\bar{\mathbf{x}}'_4, \bar{\mathbf{x}}'_5)$.

NOTATION APPENDIX

Table N.1 Variables in MOLES

Variable	Description
$d(q)$	The total distance (km) of an arbitrary route q (q is urban q^U , highway q^H or mixed q^M)
d_q^U	The total distance (km) of an arbitrary route q in the urban road network
d_q^H	The total distance (km) of an arbitrary route q in the highway road network
b_{qm}^U	The fraction of the length $d(q^U)$ falling into the territory of zone m
$\xi_0^H(l_{z' \rightarrow z}^H)$	Reciprocal of the free flow speed in arbitrary highway link $l_{z' \rightarrow z}^H$
$\xi_{0m}^{U,C}$	Reciprocal of the car free flow speed in arbitrary zone m
$\xi_{0m}^{U,PT}$	Reciprocal of the public transport mode free flow speed in arbitrary zone m
$t_c(q)$	The total travel time of an arbitrary route q (q can be urban q^U , highway q^H or mixed q^M)
$t_{PT}(q^U)$	The total travel time of an arbitrary urban route
$TL(l_{z' \rightarrow z}^H)$	Total vehicle load in the link $l_{z' \rightarrow z}^H$
TL_m	Total vehicle load (aggregate level of traffic) in zone m
l_v^U	Fuel consumption per urban km travelled by the internal combustion engine vehicle v
l_v^H	Fuel consumption per highway km travelled by the internal combustion engine vehicle v
e_v^U	Electricity consumption (litres) per urban km travelled by the electric vehicle v
e_v^H	Electricity consumption (kW) per highway km travelled by the electric vehicle v
$p_g(v)$	Per litre price of fuel type g compatible with the internal combustion engine of vehicle v
p_e	Per kW price of electricity
$\bar{N}_L$	Annual labor supply in days
$\bar{L}$	Duration of the working day
w_{js_n}	Per hour labor remuneration for the skill type s of individual n in employment zone j
e_n	Non-labor income of individual type n
p_{ir}^H	Per m ² annual rental rate of residential type r in zone i
C_{vq}^T	Annual variable transportation costs of route q with vehicle v ($v = 0$ for public transport)
c_{vq}^V	Cost of one trip in route q with vehicle v ($v \neq 0$)
$c_{v,km}^U$	Kilometre cost for urban driving with vehicle v
$c_{v,km}^H$	Kilometre cost for highway driving with vehicle v
c_q^{PT}	Cost of one trip in route q with public transport
t_{Tvq}	Daily travel time of route q with vehicle v
t_{vq}	Travel time of route q using vehicle v
t_{ij}^{PT}	Travel time required to move from residential zone i to employment zone j using public
C_{an}^*	Optimal consumption of individual type n conditional on the choice of alternative $\mathbf{a}$
H_{an}^*	Optimal residential size of individual type n conditional on the choice of alternative $\mathbf{a}$

V_{Ban}^*	Indirect utility of individual type n conditional on the choice of alternative $\mathbf{a}$
V_{an}^*	Systematic utility of individual type n conditional on the choice of alternative $\mathbf{a}$
P_{an}	Choice probability of alternative $\mathbf{a}$ by individual n
π_{ir}^D	Deterministic profit from developing one m^2 of land into residential type r in zone i
$\tilde{\pi}_i^D$	Expected maximum profit from developing one m^2 of land in zone i
π_{i0}^D	Deterministic profit from leaving one m^2 of land undeveloped in zone i
$P_{i,r}^D$	Probability of developing residential type r in zone i
$P_{i,0}^D$	Probability of leaving zone i undeveloped
p_{ir}^H	Annual rental rate (per m^2) of residential type r in zone i
k^D	Per period depreciation costs (per m^2)
p_{xi}	Annual rental rate of land (per m^2) in zone i
AS_{ir}^H	Aggregate supply of residential type r in zone i (in m^2)
VS_i	Undeveloped land in in zone i (in m^2)
PKM_m	Aggregate passenger km in arbitrary zone m
C_{PTm}	Public transport total costs in arbitrary zone m
AC_{PTm}	Average cost of a passenger km in arbitrary zone m
MC_{PTm}	Marginal cost of a passenger km in arbitrary zone m
AD_v	Aggregate demand for type v vehicles
AD_{ir}^H	Aggregate demand for floor space of residential type r in zone i (in m^2)
AD_g	Aggregate demand for fuel type g
AD_v	Aggregate demand for vehicle type v
Θ	Total land income

Table N.2 Indicator functions in MOLES

Indicator	Description	equals 1 if:
$I(v)$	Vehicle ownership	$v \neq 0$
$I_v(v, e)$	Electric vehicle	$v \neq 0$ and v is electric
$I(z, \mathcal{H})$	Zonal highway membership	Zone (node) z belongs to highway nodes $\mathcal{H}$
$I(\mathbf{a}, i)$	Zone identifier	$\mathbf{a} = \{\bar{i}, r, j, v, q\}$ is such that $\bar{i} = i$
$I(\mathbf{a}, i, r)$	Zone & housing type identifier	$\mathbf{a} = \{\bar{i}, \bar{r}, j, v, q\}$ is such that $\bar{i} = i$ and $\bar{r} = r$
$I(\mathbf{a}, v)$	Vehicle type identifier	$\mathbf{a} = \{i, r, j, \bar{v}, q\}$ is such that $\bar{v} = v$
$I(g, v)$	Vehicle fuel compatibility	Vehicle v uses fuel type g
$I(q, l_{z' \rightarrow z}^H)$	Link-route membership	Link $l_{z' \rightarrow z}^H$ belongs to highway route q
$I(q, z)$	Route-zone membership	Urban route q passes through zone z
$I_p(i, r)$	Housing type zoning	Housing type r is allowed in residential zone i
$I_H(q)$	Highway route identifier	Route q is highway
$I_U(q)$	Urban route identifier	Route q is urban