Les Cahiers du GERAD ISSN: 0711–2440

A Hybrid Space-Economic Model Generating World Demo-Economic Projections

P. Hansen
C. Meyer
L.-N. Tellier

G–2011–17

March 2011

A Hybrid Space-Economic Model Generating World Demo-Economic Projections

Pierre Hansen
Christophe Meyer
GERAD and HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada, H3T 2A7
pierre.hansen@gerad.ca
christophe.meyer@gerad.ca

Luc-Normand Tellier[*]
Département d'études urbaines et touristiques
Université du Québec à Montréal
Case postale 8888, succursale Centre-ville
Montréal (Québec) Canada, H3C 3P8
tellier.luc-n@uqam.ca

March 2011

Les Cahiers du GERAD
G–2011–17

---

[*] Corresponding author

Abstract

The GRIEG model is a hybrid model of demo-economic projections that combines two approaches: the econometric approach – based on the micro-economy – of the New Economic Geography (NEG) and the stochastic approach of the topodynamic model. The data required by the NEG part of the model being only available for the United States, the hybrid version could be applied only for the United States. In contrast, the version excluding the NEG component was applied to the entire world (including USA) since the topodynamic part of the model requires data available for all countries. This paper makes a comparative analysis of the population, production and product per capita projections for the United States using the two versions of GRIEG (the version without NEG component and the hybrid version). The results of this comparison tend to confirm the validity of the purely topodynamic version where the hybrid version is not applicable, and the interest of the hybrid model where data exist.

Key Words: Topodynamic model; New Economic Geography; Fermat-Weber problem.

Résumé

Le modèle GRIEG est un modèle de projections démo-économiques hybride qui marie deux approches : l'approche économétrique, basée sur la micro-économie, de la Nouvelle Économie Géographique (NEG) et l'approche stochastique du modèle topodynamique. Les données requises par la partie NEG du modèle n'étant disponibles que pour les États-Unis, la version hybride n'a pu être appliquée qu'aux États-Unis. Par contre, la version excluant la composante NEG a été appliquée au monde entier (y compris les États-Unis) puisque la partie topodynamique du modèle requiert des données disponibles pour tous les pays ou presque. Ce texte fait une analyse comparative des projections de population, de production et de produit per capita générées pour les États-Unis à l'aide des deux versions du modèle GRIEG (la version sans composante NEG et la version hybride). Les résultats de cette comparaison tendent à confirmer la validité de la version purement topodynamique là où la version hybride n'est pas applicable, et l'intérêt de la version hybride du modèle là où les données existent.

Mots clés : Modèle topodynamique; Nouvelle Économie géographique; problème de Fermat-Weber.

Acknowledgments: The authors acknowledge the contribution of Dr. Kristian Behrens, Professor, Department of Economics, Université du Québec à Montréal, and thank him for his major assistance in the conception of the NEG component of the GRIEG model. The second author also thanks Sébastien Le Digabel for his support in the use of the NOMAD software. It is also acknowledged that Sébastien Le Digabel and Charles Audet contributed to the description of the NOMAD software, and the authors thank them for that.

1 Introduction

In order to size the challenge of producing demo-economic projections over a fifty-year period, it is useful to look backward, and wonder what demo-economic forecasts made, for instance, in 1940 did actually materialize in 1990. Who foresaw, in 1940, the post World War II Baby Boom, the radical birth rate decline of the 60s, the weakening of marriage as an institution, the incredible economic expansion of the 50s, and the stagflation of the 70s? More recently, who, in the 90s (and even in 2007), predicted the recent world economic crisis? There is, however, one thing that most people in North America were foreseeing in 1940, which materialized and did so at about the same speed as was foreseen. It is the continuation of the historical trend observed, since, at least, 1790, by which population and economic activities progressively move towards the southwest in North America.

Such long-run space-economic trends are not unusual in history. They remain maybe the most trustable basis for generating reliable projections. The GRIEG model (named after the Group of Research in International Economic Geography[1]) attempts precisely to replicate those trends to generate long-run demo-economic projections by resorting to the topodynamic model and the models of the New Economic Geography (NEG), which are macro-geography models that may be seen as stemming from the same space-economic Weberian sources through the *attraction-repulsion problem*, as Ottaviano and Thisse (2005) have pointed out. Altogether, those models aim at understanding, simulating and predicting the evolution of large spatial systems. The NEG models have contributed a lot to the micro-economic understanding of such evolutions, but they have not always been able to satisfactorily simulate and predict them. Meanwhile the topodynamic model, which can be considered as a “complex system” stochastic model, has been rather successful in simulating and predicting the evolution of large spatial systems, but its micro-economic bases are more implicit than explicit.

Both NEG and topodynamic models focus on the understanding of the phenomenon of economic space polarization (why are there crowded areas, while there are empty spaces?), but they differ in many respects as illustrated in Table 1.

Table 1: Complementarity of the topodynamic and NEG models

	Topodynamic model	NEG models
Based on	Attraction-repulsion problem, and Monte Carlo approach	Econometrics, micro-economics, and general-equilibrium models
Resort to	System fitting	Regressions (straight line fitting) and simultaneous equations
Exploit mainly	Spatial logic of development	Micro-economic general-equilibrium logic of development
Stress	Macro regularities	Micro-economic logic
Distinguish forces of	Attraction and repulsion	Agglomeration and dispersion
Conceived while referring to	Infinite number of locations	Finite number of regions
Strengths	Macro and long-run coherence	Micro-economic and short-run coherence
Nature	Entropic	Deterministic
Required data	Very spatially disaggregated data for few variables	Less spatially disaggregated data for many variables

The pioneering contribution of Alfred Weber (1909) in location theory is based on the interaction between *attractive forces*. Specifically, Weber assumed that the firm aims at minimizing total transportation costs, which are defined by the sum of weighted distances to several points of reference, each weight expressing the importance of the corresponding point of reference to the firm. This amounts to assuming that a firm seeks a location that provides the best access to several markets or sources of inputs, which have different

---

[1] This research was sponsored by the International Opportunities Fund of the Social Sciences and Humanities Research Council of Canada (SSHRC). It was carried out by the researchers of GRIEG assisted by the following persons: Hakan Andic, Ph.D. student in Economics, Christophe Meyer, Ph.D. in Mathematics, Xavier Provençal, Ph.D. in Mathematics, Claude Vertefeuille, M.Sc. in Mathematics, Laurence Marien, M.Sc. student in Demography, and Nawel Saker, M.Sc. student in Geography.

sizes, relative positions and transportation rates. Tellier (1985) has extended this setting by introducing the concept of *repulsive force* and formulating the *attraction-repulsion problem*, thus making the firm's optimal location the outcome of the interplay of both attractive and repulsive forces. The attraction-repulsion problem consists in finding the optimal location with respect to reference points exerting both attractive and repulsive forces. Tellier (1985) found a trigonometric solution to the triangular attraction-repulsion problem, while Jalal and Krarup (2003) proposed a geometrical solution to the Fermat problem with arbitrary weights (N.B.: the original Fermat problem corresponds to a Weber problem with equal positive weights). The attraction-repulsion problem was further studied by Tellier and Polanski (1989). An algorithm that provides a numerical solution to the general attraction-repulsion problem has been developed by Chen et al. (1992). The topodynamic model resorts to series of interdependent attraction-repulsion problems to simulate space-economic evolutions. The main references for the topodynamic model are Tellier (1992) and Tellier (1995).

The NEG approach, whose main references are Krugman (1991), Fujita et al. (1999), Fujita and Thisse (2003), Ottaviano et al. (2002), Ottaviano and Thisse (2004), as well as Combes et al. (2006), resorts to classical micro-economic theory and econometrics to generate microeconomic general-equilibrium models based on imperfect competition, increasing returns and transportation costs. These models aim to explore the logic of the formation of economic agglomerations and economic spaces by resorting to the concepts of agglomeration (i.e. attractive) and dispersion (i.e. repulsive) forces. Their purpose is to explain why strong spatial disparities may emerge in settings that are otherwise symmetric and homogenous. The NEG models put the emphasis on the micro-economic underpinnings of the phenomenon of agglomeration and aim at dissecting the systems they attempt to replicate as well as the various elements of those systems. They stress the role of lock-in effects generated by a myriad of individual decisions.

The topodynamic model and the NEG models differ in two main ways. First, the topodynamic model is not econometric while the NEG ones are. Second, the topodynamic model is stochastic while the NEG ones are not. Third, the topodynamic model requires spatially disaggregated data about a limited number of variables, essentially population and production, while NEG models require spatially disaggregated data about numerous variables, like incomes, wages, housing costs, rents, urban costs, natural amenities, inner migrations, cost of living, and transportation costs. In fact, the data required by the NEG models are seldom available in most countries of the world, while those required by the topodynamic model are generally available throughout the world. This explains the structure of the GRIEG model, which aims at taking advantage of both approaches to generate world demo-economic projections.

The GRIEG model has been applied to the United States with and without its NEG component. This allowed comparing the projections stemming from the 'with and without NEG component' versions of the model, and to look for an answer to the interesting following question: does the introduction of the NEG component with the hard-to-find data it requires really improve the projections of the GRIEG model ?

2 The basic structure of the GRIEG model

The GRIEG model aims to preserve the long-run and macro reliability of the topodynamic approach, and the micro-economic coherence of the NEG one. This is done thanks to the clear distinction that exists in the topodynamic model between the determination of the optimal values of the five characteristic parameters of the model, and the estimation of the micro "corrections" that are introduced into the model for each location of the considered space once the optimal values of the parameters have been selected. The GRIEG model is made of three components:

1. a topodynamic component TP focused on the population spatial evolution;
2. a topodynamic component TY focused on the production spatial evolution;
3. a NEG component linked to the TP sub-model; the NEG component modifies the micro "corrections" of the traditional topodynamic model by taking into account the impact of various micro-economic variables on the population spatial evolution, and, on the basis of the "so-corrected" population projections, it generates its own production projections, which differ from the production projections stemming from the TY component.

The NEG component is estimated by traditional econometric methods whereas the selection of the optimal values of the two topodynamic components is made by means of NOMAD (Abramson et al.; Le Digabel, 2011), a software that implements the Mesh Adaptive Direct Search algorithm (MADS) (Audet and Dennis, 2006). MADS targets black-box optimization problems. A black-box is typically a computer simulator that takes a point as an input and that returns the values of the objective and of the constraints as outputs. It is called a black-box because its internal properties, such as derivatives, are not available. It may be also noisy, nonsmooth, discontinuous, have several local optima, and even fail to compute at a priori feasible points. In addition, the black-box may be costly to compute, and for this reason, an optimizer such as MADS aims at finding the best solution as possible with a limited budget of evaluations. Depending on the degree of smoothness of the functions, this solution is guaranteed to meet some local optimality conditions, based on the Clarke calculus (Clarke, 1983). In the GRIEG application, the ‘multistart’ version of NOMAD was used, that is to say that several executions of NOMAD were performed from different starting points, and the default parameter values from Abramson et al. (2009) were used. The objective function is the global conformity index I described in the next section; an evaluation of this function requires to perform first a projection phase (also described in the next section) for a given set of “grami” parameters, an operation that can be very time consuming.

The topodynamic sub-models require very spatially disaggregated data about few variables (the more spatially disaggregated, the better), whereas the NEG sub-model uses data, which are much less spatially disaggregated, about much more variables (it must be stressed that, in the NEG sub-model, the level of breaking-up cannot exceed the level of the less disaggregated variable). So the GRIEG model resorts to compromises in order to benefit from both approaches.

3 The topodynamic TP and TY components

The topodynamic sub-models TP and TY are stochastic. They look at the spatial evolution of populations and productions as the product of a complex system of interdependent location decisions relating to both consumption and production activities. It is considered that such a complex system results from the interaction of innumerable attractive and repulsive forces, which generate “system effects” and “spatial trends” marked by some form of “topodynamic inertia”. The TP and TY components attempt to “model” these system effects and spatial trends by means of a large number of randomly selected interdependent attraction-repulsion problems.

Let us describe the TP component based on population while keeping in mind that the TY component is entirely similar from a mathematical point of view. The TP-submodel starts with two initial distributions of population P_i and P_f, at respective times $t_i < t_f$. Let P_i^{tot} and P_f^{tot} the total population at these two times. Let $\Delta P = P_f^{tot} - P_i^{tot}$. Each projection phase of the TP-submodel will correspond to an increase by ΔP of the total population. The model stops when a specified total population target is reached. In our study, this target will correspond to the predicted total population for the years 2030 and 2060.

At the end of a projection phase of the TP-submodel, corrections are applied to the generated distribution of population. Essentially, these (topodyn) corrections are such that the generated distribution of population coincides with the observed distribution of population P_f, at the end of the first projection phase (more details follow).

A projection phase involves a certain number of iterations. At each iteration a certain number of persons is added or subtracted to some location. As indicated before, the projection phase ends when the net total number of persons added equals ΔP. For subtracting a given number of persons, the location of the subtracted persons is randomly selected amongst the existing populated locations. By contrast, for adding persons, an attraction-repulsion problem is generated by randomly selecting three reference points (attraction or repulsion points) and their corresponding location forces, and the resulting attraction-repulsion problem is solved on the set of ’populated’ locations. It must be noted that some reference points are chosen only amongst the existing populated points, while the others are selected amongst all the possible locations (populated or unpopulated) according to a given parameter i. The persons are added to the location that solves this attraction-repulsion problem.

It has been observed (see, e.g., Tellier, 1992):

1. that the more interdependent attraction-repulsion problems are, the higher the level of polarization;
2. similarly, the more important the repulsive forces are with respect to the attractive forces, the more the polarizing process benefits to the periphery of the considered space;
3. finally, the more the located activities die, the more radical the changes in the location pattern are.

On the basis of such observations, the topodynamic sub-models TP and TY aim to define a theoretical location system involving a large number of interdependent attraction-repulsion problems[2] that simulate as perfectly as possible an observed evolution in order to generate projections. In the topodynamic sub-models, interdependent attraction-repulsion problems are randomly selected according to various critical parameters. Once the parameter values that best replicate on a computer the observed evolution of population or production spatial distributions have been found, projecting the future evolution of the location system is possible if it is assumed that those optimal parameter values are stable through time.

Each of the topodynamic sub-models involves five major steps:

1. the mathematical characterization of the observed spatial evolution of population (in the case of TP) or production (in the case of TY) by means of mathematical indices;
2. the selection of the number of persons or the production value that are added (or subtracted) at each iteration; that parameter as well as parameter m (see below) determine the net number of iterations between time t_i and time t_f; it is selected by the user; the smaller the value of the parameter, the more accurate the results;
3. the testing of numerous "scenarios" corresponding to different values of five basic "*grami*" parameters, that is:
   - parameter g: the neutralization of "space friction" expressed by a limit imposed on the maximum total distance between the reference points involved by the location problems; this parameter, indirectly, takes into account the deterrence effect of distance usually expressed by "gravity" models;
   - parameter r: relative range of the repulsive forces with respect to the attractive forces in the system;
   - parameter a: proportion of attractive forces;
   - parameter m: mortality of activities;
   - parameter i: interdependence expressed as the probability for a new attraction-repulsion problem to involve optimal locations of previous attraction-repulsion problems;
4. the estimation of the optimal values of the five parameters by means of a synthetic conformity index measuring the conformity of each scenario with the observed evolution;
5. the estimation of the location-by-location 'topodyn' corrections that will be integrated into the model based on the optimal scenario;
6. the production of projections.

In order to replicate on computer the observed evolution of the spatial population distribution, an optimization criteria must be defined for selecting the optimal values of the five "*grami*" parameters (N.B.: a set of five values of the "*grami*" parameters constitutes a "scenario"). That criterion takes the form of an index of global conformity I, which corresponds to the mean of the three following partial conformity indices:

---

[2] The topodynamic components TP and TY of the GRIEG model use only three-reference-points attraction-repulsion problems. It is considered that including more complex attraction-repulsion problems would increase the mathematical complexity to no avail. As for one or two reference points attraction-repulsion problems, they are trivial and do not allow for simulating complex spatial evolutions.

1. the conformity index *RMS** defined with respect to the root mean squares, this index being equal to *(RMSmax – RMSobtained) / RMSmax*, where:

   RMSmax = the maximum possible sum of the root mean squares in the analyzed case;

   RMSobtained = the sum of the root mean squares generated by a given scenario;

2. the conformity index *C** defined with respect to the concentration index C (explained below), this conformity index being equal to *(∆Cmax – ∆Cobtained) / ∆Cmax*, where:

   ∆Cmax = the maximum possible difference between the observed and simulated values of the concentration index C;

   ∆Cobtained = the difference between the observed and simulated values of the concentration index C that has been obtained with the given scenario;

3. the conformity index *H** defined with respect to the movement of the center of gravity, this index being equal to *(∆Hmax – ∆Hobtained) / ∆Hmax*, where:

   ∆Hmax = the maximum possible distance between the observed and simulated centers of gravity;

   ∆Hobtained = the distance between the observed and simulated centers of gravity that has been obtained with the given scenario.

The value of the three indices of partial conformity and of the index of global conformity varies between 0 and 1, the value 1 corresponding to a total conformity of the simulated evolution with the observed one.

Concentration index C measures the extent to which the population (or production) is concentrated in space. It varies between 0 and 1, the value 1 corresponding to the concentration of the total population (or production) at a single point. Index C is obtained by dividing a considered space successively according to six different grids whose lines must not coincide. The first grid is 2 x 2; the second one, 3 x 3; the third one, 5 x 5; the fourth one, 11 x 11, and the fifth one, 23 x 23. As for the sixth grid, it refers to the smallest division of space: here, it corresponds to a certain number of urban regions resulting from the division of the world space in 2,397 urban regions. The mathematical expression of the concentration index C is the following:

$$C = \left(\frac{1}{2MR}\right) \sum_{r=1}^{R} \left(\frac{J_r}{TJ_r - S}\right) \sum_{j=1}^{J_r} \left|\sigma_{r,j}^{M} T - M \sigma_{r,j}^{T}\right|$$

where:

M: total mass (of population or production);
R: number of grids;
J_r: number of cells in the r-th grid;
T: total inhabitable area;
S: total area of the considered region;
$\sigma_{r,j}^{M}$: total mass of the population (or production) located in the j-th cell of the r-th grid;
$\sigma_{r,j}^{T}$: inhabitable area located in the j-th cell of the r-th grid.

The optimization of the topodynamic components TP and TY is done independently from that of the NEG component of the model.

Once the optimal values of the five "*grami*" parameters have been determined both for the TP and TY components, an "average" scenario is computed in order to minimize the biases stemming from the stochastic nature of the model. The resulting spatial distribution is called the "average" scenario. Comparing the projected values of the "average" scenario with the observed values at each location yields the 'topodyn' corrections, which are then integrated into the model.

For instance, if the simulated population of location i obtained for the most recent year of observation in the "average" scenario corresponding to the optimal values of the parameters is 1,498,653 inhabitants while the observed population of that location is 1,567,342 inhabitants, a topodynamic "correction" of + 68,689 is computed. Positive corrections and negative corrections are treated differently: basically, positive corrections are based on absolute values, but negative corrections are computed in relative terms. For instance, if the simulated population of location i obtained in the "average" scenario corresponding to the optimal values of the parameters is 1,678,450 inhabitants while the observed 2007 population of that location is just 1,437,652 inhabitants, the topodynamic "correction" will be -14.35% (instead of $-240,798$).

The reason why positive corrections are expressed in absolute terms whereas negative corrections are expressed in relative terms is that the reverse has radical consequences. Positive corrections expressed in relative terms make the fastest growing regions literally "explode", whereas negative corrections expressed in absolute terms make the declining regions literally disappear.

In the traditional topodynamic models, those corrections were assumed to remain the same throughout the period of projection, which was not always obvious since it assumed, for instance, that the "underperformance" of location i was to go on. Moreover, the population and production corrections were assumed to be independent from one another. This is where the NEG component intervenes in the GRIEG model: it is used to compute NEG corrections that link the evolution of population to economic behavior according to the micro-economic logic developed in the context of the New Economic Geography. It must be stressed that, while the rationale of the NEG corrections is micro-economic, that of the topodynamic corrections is basically spatial. For instance, topodynamic corrections reflect the fact that a given location corresponds to a port or that other location, to a major crossroads, while NEG corrections take into account that, if strong regional imbalances develop, people are likely to migrate from a less to a more developed region. Both topodynamic and NEG corrections are worth being taken into account.

In the GRIEG model as in the traditional topodynamic model, corrections are applied at the end of each "phase of projection", each "phase" adding the same additional population or production as those registered during the period of observation. First, the population projections obtained thanks to the TP component are corrected by applying the 'topodyn' TP corrections. Second, the topodynamic corrected population projections are introduced as an input into the NEG component, which performs two tasks:

1. by means of the estimated per capita income and the median urban costs elasticities with respect to population size, it computes the NEG population corrections and corrects the topodynamic corrected population projections;
2. by means of the same elasticities, it generates 'topoNEG' production projections, which differ from the TY production projections generated by the TY component.

Of course, if the data required by the NEG component are not available, the GRIEG model generates just TP and TY projections.

4 The NEG component

The TP projections do not explicitly take into account the incomes and urban costs associated with each location, when, in fact, these variables, which are very influenced by the population size, guide the migratory flows. The NEG component measures the effect of the population redistribution projected by the TP component on urban costs and productivity at each location. In order to do so, the first step consists in estimating the per capita income and the median urban costs elasticities with respect to population size. With regard to housing, two estimations are made: one for the rents, and the other for the dwelling values. A simple way to estimate those elasticities consists in taking advantage of the panel structure of the data for calculating the following log-log regressions:

$$\begin{aligned}
\ln(w_i) &= \beta_{0w} + \beta_{1w}\ln(L_i) + \text{controls} + \varepsilon_i \\
\ln(r_i) &= \beta_{0r} + \beta_{1r}\ln(L_i) + \text{controls} + \varepsilon_i \\
\ln(h_i) &= \beta_{0h} + \beta_{1h}\ln(L_i) + \text{controls} + \varepsilon_i
\end{aligned}$$

where:

$L_i =$ population at location i
$w_i =$ per capita GDP (income) at location i
$r_i =$ median rent at location i
$h_i =$ median dwelling value at location i

This yields the $\widehat{\beta}_{1w}, \widehat{\beta}_{1r}$ and $\widehat{\beta}_{1h}$ elasticities. The controls used are the classic ones in the literature: the skilled-share, the year fixed effects and the MSA fixed effects. The year fixed effects allow to rid (as much as possible) the macroeconomic variations, while the MSA fix effects control for a part of the unobserved heterogeneity (the identification of the elasticities then comes from within-MSA variations; alternatively, we could use a difference-in-difference estimation). In the regressions r_i and h_i, we use the median income as control. Table 2 presents a summary of our results for the United States.

Table 2: The estimated elasticities for the United States

	Number of observations	Elasticities for population L_i	Education	Per capita income	Fixed effects included for MSAs	Fixed effects included for years	R^2
Income	506	0.0359 (0.038)	0.4330 (0.000)		Yes	Yes	0.9867
Rent	506	0.1050 (0.000)		.7810 (0.000)	Yes	Yes	0.9703
Housing values	506	0.1255 (0.000)		1.8820 (0.000)	Yes	Yes	0.9441

Notes: 253 MSAs in 1990 and 2000; standard errors p-values in parentheses; to stay close to the data used in the 'topodyn' model and to have a larger sample of MSAs, we use 1990 and 2000 Census data.

Table 2 coefficients are interpreted in the following way: on the average, doubling the size of an MSA population increases its income by $(2^{\widehat{\beta}_{1w}} - 1) \times 100 = 2.52\%$, its rents by $(2^{\widehat{\beta}_{1r}} - 1) \times 100 = 7.55\%$ and its dwelling values by $(2^{\widehat{\beta}_{1h}} - 1) \times 100 = 9.09\%$ (the two last percentages assuming a constant income). It must be noted that these coefficients are all precisely estimated and their values match with those generally found in the literature. Rosenthal and Strange (2004), as well as Melo et al. (2009) estimate that the income elasticities generally vary between 3% to 8%, and our estimate is closer to 3% since we control, at the MSA level, for education and unobserved heterogeneity (see, for example, Combes et al., 2008 and Melo et al., 2009). Moreover, since housing represents about one third of the budget of American households and about two thirds of Americans own their dwelling, doubling the population slightly reduces the real wages. A back-of-the-envelope calculation suggests that an increase of 2.25% of the nominal wages leads to an increase of 0.33 x (7.55% × 0.66 + 9.09) = 2.83% of the urban costs, which results in a slight decrease of the real wages. Consequently, our estimates suggest that the real wages decrease by about 0.3% with respect to the population size, which reflects the presence of amenities our approach did not measure (Glaeser et al., 2001).[3] Finally, the NEG component uses the estimates based on the real estate values as a measure of urban costs. Since, in the United States, tenants are not representative of the general population, the results little vary whether rents are taken into account or not for the sake of such estimates.

Component TP provides a projected population distribution $\widetilde{L} = \{\widetilde{L}_1, \widetilde{L}_2, \ldots, \widetilde{L}_i, \ldots\}$ for year 2030, for example. Given the $\widehat{\beta}_{1w}, \widehat{\beta}_{1r}$ and $\widehat{\beta}_{1h}$ elasticities, it is possible to build the counterfactual variables $\widetilde{w}_i, \widetilde{r}_i$ and $\widetilde{h}_i$, which take into account the urban costs and the per capita GDP observed in 1990. These variables allow introducing into the GRIEG model considerations that are not explicit in the TP and TY components.

In order to specify the preferences, let us suppose that economic agents consume dwelling (price h_i), non-traded goods (price n_i) and traded goods (price t_i). Assuming simple Cobb-Douglas preferences, the indirect utility of a MSA i agent earning a salary w_i is given by:

$$V_i \equiv \frac{w_i}{h_i^{\alpha} n_i^{\beta} t_i^{1-\alpha-\beta}} \mathrm{e}^{\kappa A_i} L_i^{\theta},$$

where $0 < \alpha < 1$ and $0 < \beta < 1$, α being the share of dwelling, and β , the share of non-exchangeable goods in the budget. As for A_i, it is a standardized measure (score) of the amenities of MSA i and the unknown coefficient κ, the valuation of those amenities. The L_i^{θ} term captures the unobservable agglomeration

---

[3] Behrens et al. (2011b) quantify such unobserved amenities of the great American cities by means of a structural general equilibrium model. They find that those amenities are (i) quantitatively important, and (ii) increasing with the size of the urban population.

externalities (access to product diversity, social interactions, etc.) that are not capitalized in the wages and urban costs. It will be assumed that the prices of the non-tradable goods do not vary in a significant way throughout the considered space (in any case, those variations are relatively weak compared with the dwelling budget). Therefore, $t_i = t$ for all location i. The ratios of indirect utility are given by:

$$\frac{\mathrm{e}^{\kappa A_i}\frac{w_i}{h_i^{\alpha}}L_i^{\theta}}{\mathrm{e}^{\kappa A_j}\frac{w_j}{h_j^{\alpha}}L_j^{\theta}} = \left(\frac{n_i}{n_j}\right)^{\beta}\frac{V_i}{V_j}.$$

In the United States, the US Department of Agriculture provides amenities scores and we used them as a proxy for the A_i values. It is also easy to find data about α for the United States. According to figures of the US Bureau of Labor Statistics, shelter represented, in 2007, about 34% of the average American budget (then $\alpha = 0.34$). The left side of equation (4) being known, it was applied to the counterfactual populations L_i stemming from component TP. Since there exist no satisfactory data allowing for measuring the n_i values (there is no price indices at the level of the American states),[4] in order to compute the $\tilde{V}_i/\tilde{V}_j$ values, we had to make the strong assumption that $n_i \equiv n$ for all i.

The differences in indirect utility being estimated, it is now possible to determine, in the context of a general equilibrium model, the NEG corrections to be made to the TP population projections. In order to do so, if necessary, it is possible to recuperate the absolute levels $\tilde{V}_i$ even if it requires normalizing one of those V values through the choice of a measure unit.

Determining the NEG corrections stems from the equilibrium conditions. At equilibrium, each agent chooses the location that provides the highest indirect utility while taking into account the real wage differences ($w_i h_i^{-\alpha}$), the amenities differences (A_i) and the other "unobservable agglomeration benefits" (L_i). Moreover, migrating from i to j costs something. The loss of utility associated with those migration costs is approximated by means of the so-defined net utility an agent h residing at i can get by migrating to j:

$$V_{ij}^h \equiv V_j^h d_{ij}^{-\gamma}$$

where d_{ij} is the distance between i and j. For technical reasons, an increasing monotonous transformation of those preferences is necessary. More precisely, utility is expressed in log form. The econometrician cannot observe utility. Hence, it is stated that the following is observed:

$$\ln U_{ij}^h = \ln V_{ij}^h + \varepsilon_i^h,$$

where $\varepsilon_i^h \to$ i.i.d. Gumbel $(0, \sigma^2\beta^2/6)$ is a double-exponential Gumbel distribution of the agents' preferences for their present MSA of residence. According to McFadden (1977), Anderson et al. (1992) (for example), the probability of an agent choosing to move to j is the following:

$$\mathbb{P}_{ij} = \Pr\left(U_{ij} > \max_{k\neq j} U_{ik}\right) = \frac{\mathrm{e}^{\ln V_{ij}/\sigma}}{\sum_k \mathrm{e}^{\ln V_{ik}/\sigma}} = \frac{V_{ij}^{1/\sigma}}{\sum_k V_{ik}^{1/\sigma}}.$$

Let us note that parameter σ can be interpreted as a preference heterogeneity parameter (Tabuchi and Thisse, 2002; Murata, 2003); If σ is very large, the differences in real wages, amenities, and other agglomeration benefits do not play a role in the agents migration decisions, while they are strongly influenced by those variables if σ is small. When $\sigma \to 0$, the agents are homogeneous, and their migration decisions are entirely determined by the differences in real wages and amenities.

The unknown parameters of the NEG component must now be estimated. According to our NEG submodel, the migratory flows from i to j in a given period are given by $X_{ij} = \mathbb{P}_{ij} L_i$. Hence, we get:

---

[4] The ACCRA indices could have been used, but they bear significant methodological problems.

$$\begin{aligned}\ln(X_{ij}/L_i) &= \ln \mathbb{P}_{ij} = \frac{1}{\sigma}\ln V_{ij} - \ln\left(\sum_k V_{ik}^{1/\sigma}\right) \\ &= -\frac{\gamma}{\sigma}\ln d_{ij} + \frac{\kappa}{\sigma}A_j + \frac{1}{\sigma}\ln\left(\frac{w_j}{h_j^{\alpha}}\right) + \frac{\theta}{\sigma}\ln L_j + \text{oFE}_i\end{aligned}$$

where we take advantage of the fact that the sum in k does not vary between destinations depending on the initial location choice i. Because of the log-linear specification of the model, this last term amounts to a fix origin effect. In the United States, the data required for estimating this equation (especially the intranational migration flows X_{ij}) exist. Having estimated this equation allows to compute the choice probabilities (i.e., migration probabilities) for any counterfactual population distribution, holding the structural parameters constant. An error-correction is applied to make the initial choice probabilities as consistent as possible with the observed migration patterns between 1990 and 2000. Finally, the so obtained choice probabilities are used in the 'topodyn/NEG' procedure to yield migration flows that lead to a 'NEG corrected' population distribution taking a given counterfactual TP distribution as input. Having estimated the per capita income elasticity with respect to population, we can also easily adjust our measure of productivity changes.

5 The data challenge

The specification of the NEG component stems strongly from the constraints imposed by the availability of data. The first application of the GRIEG model involved the whole world divided in 2,397 urban regions, four years of observation (1980, 1990, 2000 and 2007), and two projection horizons (2030 and 2060). The data required by the NEG component were available in the United States, but not elsewhere. So the NEG component was used only in the context of North and Central America to generate what we called the 'topoNEG' projections.

The main problems faced in the application of the GRIEG model have been related to data. They came from two sources. First, they were linked to the fact that there is no universal standard for determining what is a city, an agglomeration or a metropolitan region. Definitions vary and the borders of each entity fluctuate through time, which makes generating disaggregated world projections extremely difficult. Second, data about other variables than population and production are unreliable or non-existent in most countries, and, when they exist, they do for different disaggregation levels (counties, metropolitan regions, states or national). As previously noted, in the NEG sub-model as in any sub-model, the level of breaking-up cannot exceed the level of the less disaggregated variable.

To deal with the first set of problems, the following choices were made. The world was divided into 2,397 urban regions, the center of each urban region corresponding to an agglomeration or metropolitan area of 100,000 inhabitants or more. The population and production data of 1980, 1990, 2000 and 2007 have been computed in such a way that:

- throughout the world, the sum of the population or production of all the urban regions belonging to a given country is equal to the population or production of that country;
- inside a country, the population of the country that exceeds the total population of the 'agglomerations or metropolitan areas' of 100,000 inhabitants or more was distributed amongst those 'agglomerations/metropolitan areas' according to the populations those 'agglomerations/metropolitan areas' had in 1980; this option was chosen in order to eliminate the distortions stemming from a distribution of that "excess" population according to the populations those agglomerations/metropolitan areas had in the other years (1990, 2000 or 2007); those distortions caused some slowly growing urban regions to end up declining, which was unacceptable;
- when population data were not available directly for 1980, 1990, 2000 and 2007, but existed for previous or following years, the estimates were made consistent with a mathematical inference obtained through a non-linear extrapolation based on average annual growth rates.

The population data mainly stem from the Nations Statistics Division (`http://unstats.un.org/unsd/default.htm`).

The gross domestic products (GDP) of the urban regions were computed by multiplying their population by the relevant regional or national per capita "comparable" GDP expressed in constant dollars US from 2005 and purchasing power parity (PPP). The used "comparable" GDP data stem from the Penn World Tables' real GDP per capita chain series produced by the University of Pennsylvania (Heston et al., 2009). Since some statistics are missing in that database for some countries, especially for year 1980, the following source was also used to complete our database: United Nations Statistics Division — National Accounts — GDP statistics.

In order to find a long-run solution to the problems stemming from the fact that the existing international statistics about urban areas are inconsistent and unreliable from one country to another, from one province/state to another, and, even, from one city to another, while they are as inconsistent and unreliable from one year to another, we dare propose such a universal urban metric system that is not based on transportation behaviors or political structures. That system would be built the following way. Earth would be covered with identical equilateral ('spherical') triangles having the following dimensions: 1 kilometer: distance between a triangle center and one of its sides; 2 kilometers: distance between a triangle center and one of its vertices; 2 kilometers: distance between two neighboring triangle centers; 3 kilometers: height of a triangle; 3.4641016 kilometers: length of a triangle side; and 5.1961524 square kilometers: area of a triangle. According to Euler's relation, it is impossible to cover the terrestrial sphere only with such triangles; however, it is possible to minimize the number of irregular tiles, and to manage to have a maximum number of irregular tiles located in lakes, seas or oceans.

The universal urban metric system would include a "basement" and 10 urban "floors" or "levels"; the basement and the 10 urban levels correspond to the following "radius of agglomeration": radius of the basement: $10 \cdot 2^0 = 10$ kilometers; radius of level 1: $10 \cdot 2^1 = 20$ kilometers; radius of level 2: $10 \cdot 2^2 = 40$ kilometers; radius of level 3: $10 \cdot 2^3 = 80$ kilometers; radius of level 4: $10 \cdot 2^4 = 160$ kilometers; radius of level 5: $10 \cdot 2^5 = 320$ kilometers; radius of level 6: $10 \cdot 2^6 = 640$ kilometers; radius of level 7: $10 \cdot 2^7 = 1,280$ kilometers; radius of level 8: $10 \cdot 2^8 = 2,560$ kilometers; radius of level 9: $10 \cdot 2^9 = 5,120$ kilometers; radius of level 10: $10 \cdot 2^{10} = 10,240$ kilometers. The Earth circumference being 40,000 kilometers long, the largest urban region of level 10 ends up covering the whole Earth at level 11, which makes that level and the following ones irrelevant.

The basic "districts" of the basement are obtained in the following way. The most populated triangle of the world annexes all the triangles whose centers are located within 10 kilometers from its own center to form its own district, and all the annexed triangles cannot be part of any other district; the center of the most populated triangle becomes the center of the district. The most populated triangle among the remaining triangles annexes all the remaining triangles whose centers are located within 10 kilometers from its center to form its district; the center of this most populated triangle becomes the center of that second district. And so on...

The "urban regions of the level 1" are obtained in the following way. The most populated district of the world annexes all the districts whose centers are located within 20 kilometers from its own center to form its own urban region of level 1, and all the annexed districts cannot be part of any other urban region of level 1; the center of the most populated district becomes the center of that first urban region of level 1. The most populated district among the remaining districts annexes all the remaining districts whose centers are located within 20 kilometers from its center to form its urban region of level 1; the center of this most populated district becomes the center of that second urban region of level 1. And so on...

The "urban regions of the level 2" are obtained in the following way. The most populated urban region of level 1 of the world annexes all the urban regions of level 1 whose centers are located within 40 kilometers from its own center to form its own urban region of level 2, and all the annexed urban regions of level 1 cannot be part of any other urban region of level 2; the center of the most populated urban region of level 1 becomes the center of that first urban region of level 2. The most populated urban region of level 1 among the remaining urban region of level 1 annexes all the remaining urban region of level 1 whose centers are located within 40 kilometers from its center to form its urban region of level 2. And so on... Urban regions of the level 3, 4, 5, 6, 7, 8, 9, and 10 are obtained in the same systematic way. Of course, the higher a level is, the smaller the number of its urban regions is, and the more populous those regions are.

Had such an urban metric system existed, the present research would have proceeded in the following way. First, the level x whose number of urban regions would have been, in 2007, approximately equal to 3000 would have been selected, and all the statistics of population and production for 1980, 1990, 2000 and 2007 would have been collected on the basis of the limits of those urban regions in 2007. Second, the GRIEG model would have been applied to the world divided into those urban regions in order to estimate the 2030 and 2060 threshold values of population and production for each continent. Third, the GRIEG model would then have been applied to each continent using the same urban regions to generate the population and production projections for those urban regions of level x. Finally, the GRIEG model could then have been applied to a particular urban region of level x to generate projections for its urban regions of level $(x-1)$ or $(x-2)$ according to the "Russian dolls" procedure (explained below), which consists in moving from a macro application to micro ones.

The second set of data problems relates to the unreliability or non-existence in most countries of disaggregated data about other variables than population and production. This led us to restrict the application of the NEG component to the United States. Even then, building up the data bank for the NEG component was complex. The used variables and the source of their corresponding data are the following:

- MSAs (Metropolitan Standard Areas) and CBSAs (Core Based Statistical Areas) populations (Census 1990, Table P001; Census 2000, Table PCT001);
- MSAs and CBSAs' number of 18-or-more years old persons with at least an "Associate Degree" (Census 1990, Table P060; Census 2000, Table PCT065);
- MSAs and CBSAs' average annual income in U.S. dollars (Census 1990, Table P114A; Census 2000, Table PCT130);
- MSAs and CBSAs' median value of the properties occupied by their owners (Census 1990, Table H061A; Census 2000, Table HCT066);
- MSAs and CBSAs' median rents paid by the tenants (Census 1990, Table H043A; Census 2000, Table HCT052);
- the migratory flows between the various counties (`http://www.census.gov/population/www/cen2000/ctytoctyflow/index.html`);
- the natural amenities index developed by the US Department of Agriculture (`http://www.ers.usda.gov/data/naturalamenities/`).

To deal with those variables, further data were required like:

- a table of correspondence between the US counties and the MSAs (`http://www.census.gov/econ/cbp/download/00_data/index.htm`);
- the geographical coordinates of the US counties (`http://www.census.gov/geo/www/gazetteer/gazette.html`).

It must be noted that, since the topodynamic parts of the GRIEG model used a partition of the United States territory into 154 urban regions while the NEG part of the model referred to the 253 US MSAs, the results of the NEG component of the model based on the 253 MSAs had to be transferred to the 154 US urban regions.

6 Comparing projections

Behrens et al. (2011a) presents the GRIEG 1980–2007, 1990–2007 and 2000–2007 projections at horizons 2030 and 2060 in terms of population, production and GDP per capita for 100 of the 2,397 urban regions of the world. Seven versions of the model were tested. They differ from one another depending on whether or not they include: — 'topodyn' corrections; — 'topoNEG' corrections; — or thresholds preventing the population or production of any urban region to decrease below 40% of its population or production observed in 2007. The projections were generated following a step-by-step approach, which we called the "Russian dolls" approach. In a first step, the "*grami*" parameters were optimized in the context of the whole world, and projections of population and production were generated for the 2,397 urban regions of the world. Those projections were used to determine the target populations and productions corresponding to the three bases of our projections (namely, periods 1980–2007, 1990–2007 and 2000–2007), as well as to horizons 2030 and 2060, for Africa, North and Central America, South America, Eurasia, Australia, and New Zealand. Once those targets have been estimated, in a second step, "*grami*" parameters were optimized in the context of each of these continents (or 'countries' in the case of Australia and New Zealand), and new population and production projections were generated for their urban regions.

Everywhere, except in Eurasia, the version of the GRIEG model with 'topodyn' corrections but no 'topoNEG' corrections nor thresholds yielded convincing results in terms of coherence with the observed evolutions (especially in terms of the smooth evolution of the standard deviation ellipses) and likelihood (which remains subjective). In the case of Eurasia, the version with 'topodyn' corrections and thresholds, but no 'topoNEG' corrections, often yielded more realistic projections. This is due to the fact that Eurasia is marked by an extreme contrast between regions that are emerging very fast (China, Korea, India, amongst others), rich regions with rather small growth rates (Western Europe, Japan), and poor regions that remain lagging. Introducing thresholds is appropriate in such a context.

The case of the United States allowed comparing the results of the version with 'topodyn' corrections, but no 'topoNEG' corrections nor thresholds, with the version with 'topodyn' and 'topoNEG' corrections, but no thresholds. Interestingly enough, the long-run projections (at horizon 2060) of both versions do not differ much. Table 3 presents a measure of the similarity of the two sets of projections. The value of that index was computed with respect to the shares of the various urban regions in the US population, in the US production and in the per capita product. The index S of similarity for the population is given by:

$$S = 1 - \frac{1}{2}\sum_{i=1}^{n}\left|\frac{P_{Di}}{\sum_{j=1}^{n} P_{Dj}} - \frac{P_{NEG,i}}{\sum_{j=1}^{n} P_{NEG,j}}\right|$$

where:

P_{Di}: 'topodyn' population projection for urban region i;
$P_{NEG,i}$: 'topoNEG' population projection for urban region i;
n: number of urban regions.

Index S varies between 0 and 1. It takes the value 1 in the case of perfect coincidence, and the value 0 if the whole projected US population is concentrated in a single urban region in both the 'topodyn' and 'topoNEG' projections, but in a different region. Then the two distributions are as dissimilar as possible. The index S for the production and the per capita product are defined similarly.

Table 3: Index of similarity of the 'topodyn' and 'topoNEG' projections for the US

Year	Projection	Index S for the share of the US population	Index S for the share of the US production	Index S for the share of the US per capita product
2030	1980–2007	91,45%	96,23%	91,97%
	1990–2007	94,70%	92,09%	93,73%
	2000–2007	93,12%	92,82%	93,91%
2060	1980–2007	82,57%	83,81%	85,37%
	1990–2007	89,84%	80,82%	84,43%
	2000–2007	85,28%	84,98%	88,98%

Analyzing the content of Table 3 allows making the following observations. First, the levels of similarity between the 'topodyn' and 'topoNEG' projections are quite high considering the theoretical differences between the two versions of the GRIEG model, and considering the fact that horizon 2030 is separated from year 2007 (which is the projection starting point) by 23 years, and horizon 2060, by 53 years, which is enormous in the context of projections. The levels of similarity decline between 2030 and 2060, which is normal, but the decline is not as important as one could have expected.

Similarly, the levels of similarity are generally lower in the case of production than in the case of population, which is normal considering that ‘topodyn’ production projections are entirely independent of the ‘topodyn’ population projections while ‘topoNEG’ production projections directly depend on ‘topoNEG’ population projections. However, the differences between the degrees of similarity of the population and production distributions are remarkably small.

The levels of similarity obtained in the case of the per capita products are stunning, since per capita products are influenced by both the population and production projections. This is all the more surprising since ‘topodyn’ per capita product projections are the result of the comparison of population and production projections, which are totally independent from one another in the ‘topodyn’ case. It must be said, that, in the numerous past applications of the topodynamic model, similar coherence and realism of the ‘topodyn’ per capita product projections were observed.

This leads to conclude, first, that, most of the time, the ‘topodyn’ and ‘topoNEG’ projections do not contradict each other. Second, their theoretical complementary nature and their empirical similarities help to precise what the long-run trends are. Third, the ‘topodyn’ projections generated for areas where the NEG component cannot be implemented for a lack of required data can be trusted.

In order to better compare the ‘topodyn’ and ‘topoNEG’ projections, it is interesting to look at the average standard deviations of the 1980–2007, 1990–2007 and 2000–2007 projections of both types, as presented in Table 4. They are expressed as a percentage of the “mean of the means”; for example, in the case of the population projections, the average standard deviation of the 1980–2007, 1990–2007 and 2000–2007 ‘topodyn’ projections is expressed as a percentage of the mean of the three mean ‘topodyn’ projected populations corresponding to the 1980–2007, 1990–2007 and 2000–2007 ‘topodyn’ projections.

Table 4: Average standard deviations of the 1980–2007, 1990–2007 and 2000–2007 ‘topodyn’ and ‘topoNEG’ projections for the US expressed as a percentage of the mean of the means of their respective projections

Variable	Horizon	Average standard deviations	
		‘topodyn’ projections	“topoNEG’ projections
Population	2030	2,61%	5,53%
	2060	5,61%	10,48%
Production	2030	13,28%	6,78%
	2060	28,26%	6,06%
Per capita product	2030	14,78%	12,39%
	2060	30,55%	16,45%

As can be seen, as far as population or production are concerned, sometimes, the ‘topodyn’ projections’ average standard deviation is smaller then the corresponding ‘topoNEG’ one, and, other times, the opposite is observed. However, in the case of the per capita product, the ‘topoNEG’ average standard deviations appear smaller than their ‘topodyn’ counterparts. This is not surprising, since ‘topodyn’ population and production projections are independent from one another while ‘topoNEG’ production projections derive from the population ones. Moreover, the NEG component assumes that some relations between urban costs, wages, productivity and other economic variables go on unchecked up to 2060, which is not the case with ‘topodyn’ projections.

7 Long-run relevance of the basic assumptions of the GRIEG model

This leads to wonder to what extent some assumptions of the GRIEG model can remain realistic in a medium or long run. The ‘topodyn’ part of the model is based on the conviction that topodynamic inertia is a strong basis for generating projections. This is probably the case, but it is also true that populations and production interact and that developing regional differences in per capita income play a role in the demo-economic evolution. This is why the introduction of the NEG component presents a real interest. However, as it has been stressed, the NEG sub-model requires disaggregated data about various variables that do not exist in most countries. Moreover, that component is based on many assumptions regarding the long-run stability of the various elasticities in the model and the multiple relations between migrations, urban costs, wages, productivities, amenities, etc. Assuming that those factors will not change for the next fifty years may seem excessive, and one can suspect that the farther we go into the future, the less unquestionable those assumptions get.

8 Conclusion

Marrying the topodynamic and NEG approaches and optimizing the topodynamic parts of the model with NOMAD has no precedence. In that sense, the GRIEG model is absolutely original. Despite the huge problems stemming from the inadequacy of the international system of urban statistics, the results presented in Behrens et al. (2011a) appear consistent with the past evolutions, realistic and enlightening. Moreover, the model has been conceived to be as user-friendly as possible. It opens up new perspectives in model building.

The next steps should consist in finding a way to conceive a NEG component that could be used in the case of countries whose statistical system is not as sophisticated as the American one. Also, in our opinion, the international urban statistical system should undergo a thorough revision. This paper has made a proposal in that respect. It may be daring, but it suggests that a radical change could be the only way to get out of the present unsatisfactory situation.

Lastly, the ‘topodyn’ projections generated for areas where the NEG component cannot be implemented for a lack of required data can be trusted, as suggested by the comparison of the ‘topodyn’ and ‘topoNEG’ projections for the United States. In fact, the improvement brought by the introduction of the NEG component and the hard-to-find data it requires is not indisputable as far as the comparison we have made of the ’with and without NEG component’ versions of the GRIEG model is concerned.

References

Abramson M A, Audet C, Couture G, Dennis Jr. J E, Le Digabel S, “The NOMAD project” Sofware available at `http://www.gerad.ca/nomad`.

Abramson M A, Audet C, Dennis Jr. J E, Le Digabel S, 2009, “OrthoMADS: A deterministic MADS instance with orthogonal directions” *SIAM Journal on Optimization* **20**(2) 948–966

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

Audet C., Dennis Jr. J E, 2006, “Mesh adaptive direct search algorithms for constrained optimization” *SIAM Journal on Optimization* **17**(1) 188–217

Behrens K, Bolduc D, Brlek S, Hansen P, Peeters D, Tellier L N, Thisse J F, 2011a, “The world in 2030 and 2060: Demo-economic projections based on the GRIEG space-economic model”, Technical report, GERAD, HEC Montréal

Behrens K, Mion G, Murata Y, Südekum J, 2011b, “Spatial frictions” Université du Québec à Montréal, mimeo

Chen P C, Hansen P, Jaumard B, Tuy H, 1992, “Weber’s problem with attraction and repulsion” *Journal of Regional Science* 467–486

Clarke F, 1983, *Optimization and Nonsmooth Analysis* (Wiley, New York). Reissued in 1990 by SIAM Publications, Philadelphia, as Vol. 5 in the series Classics in Applied Mathematics.

Combes P P, Duranton G, Gobillon L, 2008, “Spatial wage disparities: Sorting matters!” *Journal of Urban Economics* **63**(2) 723–742

Combes P P, Mayer T, Thisse J F, 2006, *Économie géographique. L’intégration des régions et des nations* (Economica, Paris)

Fujita M, Krugman P, Venables A J, 1999, *The Spatial Economy. Cities, Regions and International Trade* (The MIT Press, Cambridge MA)

Fujita M, Thisse J F, 2002, *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth* (Cambridge University Press). French translation: *Économie des villes et de la localisation*, 2003 (De Boek, Louvain-la-Neuve)

Heston A, Summers R, Aten B, 2009, “Penn world table version 6.3” Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania

Jalal G, Krarup J, 2003, "Geometrical solution to the Fermat problem with arbitrary weights" *Annals of Operations Research* **123** 67–104

Krugman P R, 1991, "Increasing returns and economic geography" *Journal of Political Economy* **99** 483–499

Le Digabel S, 2011, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm" *ACM Transactions on Mathematical Software* **37**(4) 44:1–44:15

Glaeser E L, Kolko J, Saiz A, 2001, "Consumer city" *Journal of Economic Geography* **1**(1) 27–50

McFadden D, 1977, "Modelling the choice of residential location" Cowles Foundation Discussion Papers 477, Cowles Foundation for Research in Economics, Yale University

Melo P, Graham D J, Noland R B, 2009, "A meta-analysis of estimates of urban agglomeration economies" *Regional Science and Urban Economics* **39** 332–342

Murata Y, 2003, "Product diversity, taste heterogeneity, and geographic distribution of economic activities: Market vs. non-market interactions" *Journal of Urban Economics* **53** 126–144

Ottaviano G, Takatoshi T, Thisse J F, 2002, "Agglomeration and trade revisited" *International Economic Review* **43** 409–436

Ottaviano G, Thisse J F, 2004, "Agglomeration and economic geography", in *Handbook of Regional and Urban Economics*, volume 4 Eds J V Herderson, J F Thisse, (North-Holland, Amsterdam)

Ottaviano G, Thisse J F, 2005, "New economic geography: what about the N?" *Environment and Planning A* **37** 1707–1725

Rosenthal S S, Strange W C, 2004, "Evidence on the nature and sources of agglomeration economies", in *Handbook of Regional and Urban Economics*, volume 4 Eds J V Herderson, J F Thisse, (North-Holland, Amsterdam) pp 2119–2171

Tabuchi T, Thisse J F, 2002, "Taste heterogeneity, labor mobility and economic geography" *Journal of Development Economics* **69** 155–177

Tellier L N, 1985, *Économie spatiale: rationalité économique de l'espace habité* (Gaëtan Morin, Chicoutimi)

Tellier L N, 1992, "From the Weber problem to a "topodynamic" approach to locational systems" *Environment and Planning A* **24** 793–806

Tellier L N, 1995, "Projecting the evolution of the North American urban system and laying the foundations of a topodynamic theory of space polarization" *Environment and Planning A* **27** 1109–1131

Tellier L N, Polanski B, 1989, "The Weber problem: Frequency of different solution types and extension to repulsive forces and dynamic processes" *Journal of Regional Science* **29**(3) 387–405

Weber A, 1909, *Über den Standort der Industrien* (J.C.B. Mohr, Tübingen) — English translation: *The Theory of the Location of Industries*, 1929 (Chicago University Press, Chicago)