Endogenous Prices in a Riemannian Geometry Framework

François Gardes
Paris School of Economics, Université Paris I Panthéon Sorbonne*

Preliminary version 26 April 2018

Abstract

Economic agents decide often under different price conditions, and moreover the individual price systems may depend on their previous choices. Such a situation appears for instance for consumers' choices under constraints or whenever a non-monetary constraint is added to the monetary budget constraint. In case where no market exist which regulates these prices and make them converge toward a common price for all agents, observed differences appear between the social distribution of consumer expenditures and their change over time which is modelled here using Riemannian geometry. Social distribution is measured along the geodesics of Riemannian surfaces, while changes over time correspond to movements along the tangents of these Riemannian surfaces. The Riemannian curvature of the consumption space is shown to be non-null for the Polish consumers surveyed in a four years Polish panel. This implies that usual econometric methods based on a unique metric over the whole consumption space are inadequate to estimate geodesics on the Riemannian surface. In order to propose an alternative, we define a synthetic time axis in the space of the variables which are observed in cross-section. Considering the relative position of two individuals along this time dimension allows us to estimate equations of geodesics. Also, an instrumentation using this synthetic time axis is proved to be very efficient compared to usual instrumentation for dynamic models on panel data.

Introduction

The difference between cross-section and time-series estimates, recognized early in the literature, is not well accepted by the profession, as it implies abstaining from making dynamic inferences from cross-section estimation (differences between agents observed in the same period does not provide the same information as changes over time for the same individual or the same population). This difference has been suggested to result from aggregation biases, from the time when panels of individuals were not available.

*Maison des Sciences Economiques, 106–112 Boulevard de l'Hôpital, 75647, Paris cedex 13, France.
This paper has benefited remarks from seminar participants at University of Québec in Montréal, Caen, Paris X, Clermont-Ferrand, Strasbourg Universities and JMA congress, and from J.L. Arcand, S. Bazen, J. Boelaert, F. Laisney, P. Merrigan and V. Simmonet. The author thanks Professor Gorecki and Christophe Starzec for allowing him to use the Polish panel data set and Nicholas Sowels for revising the text.

Specification biases (whenever the estimates performed on surveys do not take into account dynamic behaviour, such as habits or addiction in consumption functions) and different effects of errors in variables in the two dimensions have also been advanced as potential explanations. With panel data, it is possible to take these difficulties into account and to show that differences still appear between estimates in the temporal and spatial dimensions (Gardes et al., 2005). Whence cross-section estimates differ from time-series, the first one can be biased by the existence of endogenous permanent latent variables, the effect of which disappear in the time dimension. For instance, the relative position of a household in the income distribution, supposed to remain constant through time, may imply specific constraints or social interaction which determine its economic choices. Comparing two households with different relative incomes thus imply different effects of this permanent latent variable while the observation of the same household between two periods reveals only the effect of a change of its current income conditional to a constant relative income.

This article considers the intrinsic geometric properties of a dataset, such as a survey of family budgets, as a tool to get some information on the hidden effect of latent variables on the model. Indeed, consider a curve on a surface (x, z) inR^{n+1}, with x a set n explanatory variables and z the explained variable. The curve $x = g(z)$ can be considered as immersed in the smooth surface in R^{n+1} corresponding to all observation of (x, z) given by the dataset. This surface is thus a hypersurface of dimension n in the space of dimension (n+1) which corresponds to an economic model defined by the equation relating x to z. The intrinsic geometry of the surface contains all geometric characteristics which could be observed from within the surface – thus without any information which is not given by the dataset, such as the influence of latent variables. For instance, Gauss proved in his Theorema Egrenium, that, for n-dimensional hypersurface inR^{n+1}, the product of the largest and the smallest among the principal curvature of the surface (named Gaussian curvature) can be measured from within the surface: this Gaussian curvature is an intrinsic property of the surface. Therefore, although the principal curvatures are not intrinsic, a particular combination of them is intrinsic.

In this article, surfaces corresponding to a dataset are considered as Riemannian manifolds and the tangent planes at each point of the surface are linked to the model conditional to a fixed value of all latent variables which determine z at this point (i.e. for the corresponding values of variables x at this point). In a 3-dimensional space (for instance consumption x determined by income per unit of consumptionz_1 and age of the family head z_2), tangent space at a point (x, z) are defined as corresponding to the variation of (x, z) conditional to a normal to the tangent plane equipollent to the normal at the surface in M (conditional in a (n+1)-dimensional space, to the fixity of the basis over the tangent plane). Subsequently, this normal vector will be assimilated to a vector of shadow prices of x_1 and x_2, including the influence of (potentially endogenous) latent variables for each x. Therefore, the tangent plane indicates changes in z unrelated to the influence of latent variables (at least independently from the influence of latent variables for a move from point M over the surface). For instance, consider panel data and the difference between the estimate obtained on cross-section (by a between transform for instance) and the estimate in the time-series dimension (by a within transform). This last estimate can be considered as the one corresponding to changes in x conditionally to no change permanent latent variables W (such as cohort or education effects which can be supposed to be constant over time). The time-series estimate thus indicates the gradient of z over x conditionally to constant W and are no more biased by potential endogeneity of these variables.

Note that in a Riemannian space all variables play the same role as concerns the curvature of the space. The method in this article measures the particular curvature which is created by the effect of a correlation between the residual in a model and the explanatory variables, this correlation being due to endogenous latent variables which explain part of the residual. It differs from the curvature of the algebraic structure of the model $x = g(z)$ which can be calculated by usual method for a 2 or 3-dimensional space (and tensor analysis for larger dimensions, see for instance Jeanperrin, 2000) and which has for instance been analyzed in the calculation of the rank of a demand system (Lewbel, 1990). The Riemannian curvature adresses to the particular model g which is analysed and indicates the remaining change in the hypersurface corresponding to this model, which makes it non-euclidian (meaning that a unique distance cannot be applied to the whole surface).

I consider that spatial relationships correspond to geodesics in a Riemannian space, while temporal relationships are modelled as movements along tangents to these surfaces. Surprisingly, Riemannian geometry has been only little applied to economic problems (it has, however, recently appeared in theoretical statistics). I show that tensor algebra allows us to analyze the difference in estimations in both dimensions, and to compute the curvature of the Riemannian space. When this curvature is non-zero (i. e. when the integrability conditions, which make it possible to define a common Euclidean metric for all points of the space, do not hold), the space is no longer Euclidean, and the shortest route between two points are not lines but geodesics, which may bring about new econometric problems. Agents situated on a Riemannian space are supposed to follow optimal pathes (along their life-cycle) which are geodesics over the surface. Thus observing the geodesics gives some information on the time structure of cross-sectional data. Geodesics are defined by a system of differential equations the estimation of which allows to recover some of the geometric properties of the Riemannian space. Therefore, considering these equations may permit to estimate dynamic models using only cross-sectional data, as shown in the last section.

Section 1 presents the Riemannian specification of the economic model. Section 2 proposes a method to evaluate the Riemannian curvature of Family Budget surveys, with an application to a Polish panel which leads us to discuss the application of usual econometric methods on cross-sections. Section 3 presents various economic situations to which such a Riemannian analysis can be applied . Section 4 discusses the estimation of an economic model in the cross-section dimension while taking into account the effect of permanent latent variables. All of these models are derived for households expenditures data but can easily be applied to other types of statistics.

1 The Riemannian Geometry of consumption

1.1 Metric and tangent surfaces

Consider the set of n variables consisting of the n_1 variables at choice (for instance expenditures) x_i^h and the n_2 characteristics z_k^h of an economic agent h. This set is at least of dimension n_2 whenever the x_i depend on the z_k. The variables x and z define a point M in $\mathbb{R}^n$, with $n = n_1 + n_2$. A cross-section consists in N such points for all agents surveyed, a time-series in the dynamic configuration of T points for each agent. We suppose that z_i is independent from z_j for all $j \neq i$ and that the characteristics are exogenous ($\frac{\partial z_k}{\partial x_{k'}} = 0$, $k \neq k'$). Suppose N is large and the observations (x, z) describe a smooth surface over $\mathbb{R}^n$. At each point, we define a tangent space by means of the gradient of x at this point. As this gradient changes from one point M to another infinitely close point M+dM the derivative corresponding to these two close points on the surface will differ from the gradient at each point. Thus, the gradient at a point is related to the time variations of the variables, while the derivative (named absolute derivative) over the surface describes the cross-section variations. Riemannian geometry consists in connecting together the tangent surfaces corresponding to close points by linear transformations, so that the Riemannian surface can be locally represented by an Euclidean space, the metric of this space defining the scalar product and the distance between two points on the tangent surface. If the integrability conditions are satisfied, these Euclidean metrics can be embedded in an Euclidean metric for the whole Riemannian space, which becomes an Euclidean space. In this case the Riemannian curvature is zero. Conversely, if the curvature differs from zero, the space cannot be considered as Euclidean and one cannot define an Euclidean metric over the whole space.

We consider that (x, z) pertain to a n-dimensional manifold V_n, associated with a domain E^* of $\mathbb{R}^n$ (the atlas at point (x, z)) by a C^p-diffeomorphism ($p \geq 2$):

$$m = (y^1, \ldots, y^n) \in \mathrm{E}_n^* \subset \mathbb{R}^n \;\Rightarrow\; M = f(m) \in V_n \tag{1}$$

Here y^i are the coordinates of M on the chart E^*. This means that every neighbourhood of a point M_0 in V_n can be represented by the system of n coordinates $\{y^i\}_{i=1}^n$. V_n is said continuously differentiable at order $p \geq 2$ if other systems of coordinates, $\{u\}$, are related to $\{y\}$ by a C^2-diffeomorphism $u^i = g(y)$. A natural basis $\{e_i\}$ corresponds to these coordinates for point M_0 such that $OM_0 = \sum_i y^i e_i = y^i e_i$, using the Einstein convention (summation runs over indexes which appear as both a subscript and a superscript). So $\frac{\partial M}{\partial y^i} = e_i$. As E_n^* is Euclidean, we can define its metric at each point M_0 by the quadratic form :

$$\mathrm{d}s^2 = g_{ij} \mathrm{d}y^i \mathrm{d}y^j \tag{2}$$

with $g_{ij} = e_i e_j$. The g_{ij} are arbitrary continuously differentiable (at order p) functions of the y^i. So far, V_n is a topological manifold covered by compatible C^p coordinate charts, with a "Riemannian metric" g at each point which is any smooth positive definite matrix[1].

[1]Note that this is not really a more general setting than $\mathbb{R}^n$, as proved by the Nash theorem—every

We can now define the tangent surface at m_0 (the corresponding point of $M_0 \in V_n$ in E^*) by the equation:

$$m_0 m = \left[(y^i - y_0^i) + \Psi^i(y^i - y_0^i)\right] e_i$$

with Ψ^i a second order function with respect to $(y^i - y_0^i)$. This equation defines the first-order representation in the neighbourhood of M_0. Hence the system of coordinates y^i also applies to point m in the neighbourhood of M_0. It can be shown (**lichnerowicz1987**) that any change in the system of coordinates from $\{y\}$ to another system $\{y'\}$, with $A_i^k = \frac{\partial y^i}{\partial y'^k}$, changes the metrics by the formula $g_{ij} = A_i^k A_j^l g_{kl}$, so that the metric does not truly depend on the system of coordinates: it is an intrinsic notion. A second-order representation of the tangent surface can also be defined in M_0 so that (**lichnerowicz1987**) the Euclidean metric on E^* and the Riemannian metric on V_n have the same coefficients with the same derivatives (they are said to be connected metrics for $y^i = y_0^i$).

From point M in V_n to another point $M + \mathrm{d}M$ (with $\mathrm{d}M = \{\mathrm{d}y^i\}$) which is infinitely close to M, the basis $\{e_i\}$ changes according to formula:

$$e_i + \mathrm{d}e_i = e_i + \omega_i^j e_j \tag{3}$$

with ω_i^j defined as a linear function of the changes $\mathrm{d}y^i$ of the coordinates of M:

$$\omega_i^j = \Gamma_{ki}^j \mathrm{d}y^k. \tag{4}$$

This change is related to the change of the metric, since $g_{ij} = e_i e_j$. Thus, the metric in each point writes: $g_{ij} = \frac{\partial x/\partial y^i}{\partial x/\partial y^j}$.

If the metric g satisfies the *integrability conditions*, there exists a system of coordinates in E_n^* such that the metric in E_n^* takes the form of equation (3) on all points of E_n^*. This means that g is a continuous function on E_n^*. These conditions, which apply to the differential equations relating the two points and the associated basis ($\mathrm{d}M = \mathrm{d}y^i e_i$ and $\mathrm{d}e_i = \omega_i^j e_j$: (**lichnerowicz1987**),) require the symmetry of the second derivatives of M and e_i, and can be written: $\Gamma_{ki}^j = \Gamma_{ik}^j$ for all i, k on all points $\{y^i\}$. They allow us to calculate at each point the n^3-scalar Γ_{ki}^j knowing the $\frac{n(n+1)}{2}$ values g_{ij} on each point.

1.2 The Levy-Civita linear connection

Consider now the change from M to $M + \mathrm{d}M$. The natural basis $\{e_i\}$ assigned to point M changes to $\{e_i + \mathrm{d}e_i\}$ on point $M + \mathrm{d}M$. Thus, any vector $v = v^i e_i$ is differentiated as

$$\mathrm{d}v = \mathrm{d}v^i e_i + v^i\,\mathrm{d}e_i = \mathrm{d}v^i e_i + v^h \omega_h^i e_i$$

abstract Riemannian manifold can be *isometrically* embedded in some $\mathbb{R}^m$—but the calculus is usually easier in V_n.

So, the components of the vector dv write $\nabla v^j = \mathrm{d}v^i + \omega^i_h v^h$ and are denoted the *absolute differential* of v^i. The corresponding partial derivatives[1] are:

$$\nabla_k v^i = \partial_k v^i + \Gamma^i_{kh} v^h. \tag{5}$$

This defines the Levy-Civita connection between the tangent spaces.

We identify the *absolute derivative*, which includes the change of coordinates along the tangent surface and the change of the natural basis, with the cross-section marginal propensities on the Riemannian surface. These marginal propensities correspond to the overall change between two points M and M + dM on the surface, while the derivatives $\partial_k v^i$ are associated with the change of the coordinate v^i along the tangent surface (i.e. for a constant basis). The estimation of the cross-section and time-series parameters for each point of the surface thus allows us to calculate $\Gamma^i_{kh} v^h$. Supposing that these parameters are constant over a neighbourhood of M_0 allows us to calculate the coefficients Γ^i_{kh} as the parameters of the coordinates $v^h(M)$ for M in this neighbourhood.

In the application, we consider $v = (x, z)$ as containing the expenditures x_i of households on n commodities, and the characteristics z_k of the household, such as its income (or total expenditure), family size (measured in units of consumption), location, individual prices and so on.

Consider now Cartan's *quasi-parallelogram*: an initial change from M_0 to $M_1 = M_0 + \mathrm{d}M_0$ is followed by a second change from $M_0 + \mathrm{d}M_0$ to $M'_1 = [M_0 + \mathrm{d}M_0 + \delta(M_0 + \mathrm{d}M_0)]$. The two derivatives d and δ can be commuted:

$$M_0 \;\Rightarrow\; M_2 = M_0 + \delta M_0 \;\Rightarrow\; M'_2 = M_0 + \delta M_0 + \mathrm{d}(M_0 + \delta M_0).$$

Cartan shows that points M'_1 and M'_2 coincide if the Christoffel symbols are symmetrical:

$$\mathrm{d}\delta m_0 - \delta \mathrm{d} m_0 = (\Gamma^h_{ki} - \Gamma^h_{ik})\,\mathrm{d}y^k\,\delta y^i e_h = 0 \iff \Gamma^h_{ki} = \Gamma^h_{ik},$$

which is the integrability condition for *points*, with m_0 being the point corresponding to M_0 on the related Euclidean surface.

As M varies on the Riemannian surface, the basis also changes and is submitted to a rotation which is measured by the Riemann—Christoffel Tensor of Curvature. Suppose the two differentiations defined by Cartan are used one after another. The difference of the change of the basis according to the order of differentiation can be computed as: $\mathrm{d}\delta e_i - \delta \mathrm{d} e_i = R^h_{irs}\,\mathrm{d}y^r\,\delta y^s e_h$, and this difference is null if the Tensor of Curvature disappears:

$$\mathrm{d}\delta e_i - \delta \mathrm{d} e_i = R^h_{irs}\,\mathrm{d}y^r\,\delta y^s e_h = 0 \iff R^h_{irs} = \partial_r \Gamma^h_{is} + \Gamma^h_{rl}\Gamma^l_{is} - \partial_s \Gamma^h_{ir} - \Gamma^h_{rl}\Gamma^l_{ir} = 0. \tag{6}$$

R^h_{irs} is called the Riemann—Christoffel Tensor of Curvature. This condition defines integrability for vectors, i. e. basis, and the absence of torsion in the Riemannian space.

[1] The absolute (or covariant) derivatives are the component of a tensor, i. e. change according to certain formulas when the basis changes, while the ∂_k cannot be considered as tensors. Thus, only the ∇_k correspond to the change from M to $M + \mathrm{d}M$ on the Riemannian surface.

The Riemann—Christoffel Tensor of Curvature can also be defined in terms of the covariant derivatives of the Christoffel symbols using the Ricci identities, $R^h_{irs}v^s = \nabla_{rs}v^h - \nabla_{sr}v^h$, so that:

$$R^h_{irs} = \nabla_r \Gamma^h_{is} - \nabla_s \Gamma^h_{ir}. \tag{7}$$

The test of the Riemannian versus the Euclidean structure of the space relies on the nullity of the scalar Riemann curvature (which is obtained by contraction[1] of the curvature tensor): the space is Euclidean only if the scalar Riemann curvature is zero at each point.

1.3 Computation of the Christoffel symbols

We can equate the cross-section marginal propensities with the covariant derivatives $\nabla_k v^i$ and the between-period propensities with the time derivatives $\partial_k v^i$; we therefore compute the Christoffel symbols as $\nabla_k v^i - \partial_k v^i = \Gamma^i_{kh} v^h$. The covariant derivatives of these $\nabla_s \Gamma^i_{kh} = \Gamma^i_{kh;s}$ can be obtained by estimating the equation: $\Gamma^i_{kh} = \Gamma^i_{kh;s} \cdot v^s + \gamma^i_{kh} + \varepsilon^i_{kh}$. The system of these two equations reduces to:

$$\nabla_k v^i - \partial_k v^i = \Gamma^i_{kh;s} \cdot v^h v^s + \gamma^i_{kh} v^h + \varepsilon^i_k. \tag{8}$$

Another method of estimating the differences between the absolute and partial differentials of v^i would be to use the formulas defining geodesics in Riemannian space: along geodesics acceleration is null, so that (**morgan1998**) and (**lichnerowicz1987**):

$$\frac{\mathrm{d}^2 v^i}{\mathrm{d}t^2} + \Gamma^i_{kh} \frac{\mathrm{d}v^h}{\mathrm{d}t} \frac{\mathrm{d}v^k}{\mathrm{d}t} = 0. \tag{9}$$

1.4 Consequences

The Riemannian structure of the consumption space can then be described as follows:

1. $\Gamma^i_{kh} \neq 0 \iff$ estimated cross-section consumption behaviour differs from time-series behaviour[2].
2. $\Gamma^i_{kh} = \Gamma^i_{hk} \iff$ path independence of consumption. For instance, two consecutive changes can affect a household: first an increase of its income (both adults work on the market), then the birth of a child (only one spouse working on the market), or the reverse two changes. The two sequences, operated during the same lengh of time, may result in different socio-economic situations for similar households.
3. $R^h_{irs} \neq 0 \iff$ the differential equations $\mathrm{d}e_i = \omega^j_i e_j$ are integrable, i. e. there is no torsion for the basis affected to the different points in the space. In that case, the same metric can be applied to the whole surface (corresponding to a survey).
4. The sign of R indicates the sign of Riemannian scalar curvature at each point.

[1] Contraction consists in summing up a tensor over the same index which is in both high and low position: $R_{ij} = R^h_{ihj}$.

[2] Note that the difference between cross-section and time-series parameters depends on the agent's location on the Riemannian surface, i. e. that the shadow prices are not the same for all agents.

In Riemannian space, the metric is attached to each point by the equation which defines locally the distance between two arbitrary close points M and $M' = M + \mathrm{d}M$: $\mathrm{d}(M, M') = g_{ij}\, \mathrm{d}v^i\, \mathrm{d}v^j$. Christoffel symbols are related to the metric g by the formulÃ (**delachet1969calcul**) and (**lichnerowicz1987elements**):

$$g_{ih}\Gamma^h_{kj} + g_{jh}\Gamma^h_{ki} = \partial_k g_{ij}, \quad \text{with} \quad \partial_k g_{ij} = \frac{\partial g_{ij}}{\partial v^k}$$

Christoffel symbols thus indicate second-order derivatives and can be used to calculate the curvature of the space. These formula allow us to calculate Γ in terms of the metric:

$$\Gamma^h_{ki} = g^{jh}\Gamma_{kji} = \frac{1}{2}(\partial_k g_{ij} + \partial_i g_{gk} - \partial_j g_{ki}) \tag{10}$$

Therefore, the curvature of the space is defined by the metric which is attached to each point, and which can be represented as a square matrix over indices (i, j) of the different variables. This metric describes in a sense how a particular situation M is related to its close neighbours. It defines the situations, in terms of variables v, in which the individual will move when the determinants of its consumption change. The set of these situations is the geodesic passing through M. Thus, the metric, which plays the role of gravitation in physics (describing the properties of the space by the gravity attached to each of its point) can be interpreted as the full cost associated with each situation (see Section 3.2).

2 The Riemannian curvature of the consumption space

2.1. Theory

Classic methods

The curvature of a Riemannian surface is usually recovered by two alternative methods: either the metric g_{ij} is calculated using the relationship between a particular basis on the surface (for instance polar coordinates) and the euclidian basis of the space. In that case, the Christoffel symbols can be calculated on each point by equation (10); another method calculates directly the curvature of the equation which generates the surface. For instance for a surface in a three dimensional space given by the equation $z=g(x_1,x_2)$, the total curvature is given in terms of the two first order derivatives p, q over x_1 and x_2 and the three second order derivatives r, t and s over (x_1,x_1), (x_2, x_2) and (x_1,x_2): $K=\frac{rt-s^2}{(1+p^2+q^2)^2}$. For a Working specification of the budget share w_i in terms of log-income per Unit of consumption $\frac{y}{n}$ and log-price:

$$w_i = \alpha_i + \beta_i log(\frac{y}{n}) + \gamma_i log(\pi_i) + \varepsilon_i \tag{11}$$

the total curvature writes $K = \frac{-n^2\gamma_i(\beta_i+\gamma_i)}{\left[1+n^2(w_i+\beta_i)^2+\frac{\gamma_i^2 y^2}{\pi_i^2}\right]^2}$ which is positive (respectively negative) when $\beta_i < -\gamma_i$(respectively $>$). The estimation of food expenditure on the Polish panel used in this article gives the following estimates for the average values of income, price, food budget share and family size: 1581.03 zlotis, 1.15, 0.39 and 2.48. The estimates of the parameters are: $\beta_i = -0.168$; $\gamma_i = -0.036$ which gives a total curvature $K = +1.39x10^{-5}$.

Empirical curvature

I propose in this article an alternative original method based on equations (5) and (8). Consider that the survey given by n variables (here the partial expenditure for some good i, its relative price, the household's income and various other covariates such as the age of the head) corresponds to a surface where the explained variable (here the partial expenditure) depends on the other variables. This gives rise to a (n-1) dimension hypersurface the curvature of which can be related to the Christoffel symbols which describe the geometric intrinsic properties of the surface. The relation between these Christoffel symbols and the Lévy-Civita connection is used to evaluate empirically these parameters and subsequently the Riemannian curvature.

The derivatives $\Gamma^i_{kh;s}$ of Christoffel symbols are computed as indicated in sub-section 1.3 by estimating equation (7) . $\nabla_k v^i$ is the derivative on income (i.e. on log-income per UC in a linear Working specification such as equation (11)) estimated in the Between dimension. $\partial_k v^i$ is the corresponding parameter after a within transformation. These parameters are supposed to change from one point on the surface to another. Several estimations of these derivatives could be obtained by grouping a panel into cells defined by exogenous variables (such as education level, location and age groups). Under some hypothesis on the grouping method, the absolute derivative $\nabla_k v^i$ could be calculated on the cross-section dimension within cells, while the time derivative $\partial_k v^i$ is estimated between the different observations of individual in the same cell and in different periods. A similar method could be applied to a pseudo-panel of repeated cross-sections. These parameters could also be recovered at each point of the space by a local estimation using a non-parametric method. Another method is used here. We estimate a system of equations based on the Working form (or another model such as the linear or double-log) on a four years Polish panel presented in Appendix: the first four equations relate the budget share of good i in period t for household h w_{ith} to lny=log($\frac{y_{ht}}{n_{ht}}$) the logarithm of total expenditure per unit of consumption,Z_{ht} a vector of covariates describing the households' characteristics, $ln\pi_{ht}$ the logarithmic full price changing from one household to another (the definition of which is presented explained in Appendix D).

This set of equation is completed by three equations in difference corresponding to equation (8):

$$\partial_k v^i = \nabla_k v^i - (\Gamma^i_{kh;s} \cdot v^h v^s + \gamma^i_{kh} v^h) + \varepsilon^i_k. \tag{12}$$

which is applied to the data as:

$w_{ith} - w_{i,t-1,h} = dw = \alpha^0_{it} + \beta_i dlny_{ht} - \{\gamma_{211} w^2_{ith} + \gamma_{212} w_{ith} lny_{ht} + \gamma_{213} w_{ith} ln\pi_{ht} + \gamma_{222} lny^2_{ht} + \gamma_{223} lny_{ht} ln$
$\{\gamma_{311} w^2_{ith} + \gamma_{312} w_{ith} lny_{ht} + \gamma_{313} w_{ith} ln\pi_{ht} + \gamma_{322} lny^2_{ht} + \gamma_{323} lny_{ht} ln\pi_{ht} + \gamma_{333} ln\pi^2_{ht}\} dln\pi_{ht} + \eta_{ith}$

where γ_{rsv} is an estimate of $\Gamma^i_{rs;v}$ for good i.

The Riemann curvature is computed by equation (6) :

$$R^h_{irs} = \partial_r \Gamma^h_{is} + \Gamma^h_{hl}\Gamma^l_{is} - \partial_s \Gamma^h_{ir} - \Gamma^l_{ih}\Gamma^h_{sl} = \nabla_r \Gamma^h_{is} - \nabla_s \Gamma^h_{ir} =$$
$$= (\gamma^i_{ks;h}\nabla_r v^h + \gamma^i_{ks;r}) - (\gamma^i_{kr;h}\nabla_s v^h + \gamma^i_{kr;s}) = \gamma^i_{ks;h}\nabla_r v^h - \gamma^i_{kr;h}\nabla_s v^h$$

for $h \neq s$ and r, as the γ are symmetrical in (r, s).

The Riemann-Christoffel tensor R^h_{irs}describes the curvature properties of the Riemannian surface. This tensor depends on four variables i, k, h, j and can be contracted to the Ricci tensor by equalizing k with h: $R_{ij} = R^k_{ikj}$. This tensor can be contracted equalizing i and j to give rise to the scalar Riemannian curvature: R=g$_{ij}$R$_{ij}$. The $\frac{n(n+1)}{2}$ weights g$_{ij}$ defining the metric depend on the Christoffel symbols by means of $\frac{n^2(n+1)}{2}$ differential equations which cannot be easily resolved. Therefore, this scalar curvature R cannot be calculated except by calibrating the g$_{ij}$. In case of a surface of two dimensions in a three dimensional space (a variable x, for instance consumption, depending on two determinant variablesz_1,z_2, for instance income per unit of consumtion and the commodity full price), the two dimensional Ricci tensor R_{ij} occurs only two times: for k=x$_1$, i=x$_1$ and j=x$_2$ (or its negative inversing i and j) and for k=x$_2$, i=x$_2$ and j=x$_1$ (or its negative inversing i and j). All other alternatives (such as $k = x_1$, $i = x_2$ and $j = x_2$) are nul. The two tensors R_{223} and R_{332} thus measure the principal curvatures of the surface (its maximum and minimum curvatures) which allow to calculate the average and total Gaussian curvature as R_{223} R_{332} and the average curvature as $\frac{R_{223}+R_{332}}{2}$.

R_{223} and R'_{12} can finally be written:

$R_{223}(i, h, t) = \gamma_{213}\nabla_{ly} w_{ith} + \gamma_{233}\nabla_{ly} ln\pi_{ht} - \gamma_{211}\nabla_{l\pi} w_{iht} - \gamma_{211}\nabla_{l\pi} lny$

$R_{332}(i, h, t) = \gamma_{321}\nabla_{l\pi} w_{ith} + \gamma_{322}\nabla_{l\pi} lny_{ht} - \gamma_{331}\nabla_{ly} w_{iht} - \gamma_{333}\nabla_{ly} ln\pi_{ht}$

2.2. Application

$\nabla_k v^i$ and $\partial_k v^i$ have been estimated on the Polish panel (1997–2000, see Appendix A for a description of the data) containing data over 3052 households. Prices are defined as full prices calculated by means of an estimation of the opportunity cost of time using the model exposed in Gardes (2017) described in Appendix. This model is based on a household domestic production for each commodity (such as eating or ransportation) using a direct utility approach. A Cobb-Douglas specification is adopted both for the direct utility function depending on the quantities of domestic activities and for the domestic production functions of these activities (which depend on the monetary expenditure and time use for each activity). All functions can be estimated locally, either on subpopulations or by a non-parametric method, so that te assumption of unitary elasticity of substitution applies only to a neighborhood of each observation. This allows to define an indirect utility function depending on a proxy of total expenditure and total domestic time.The opportunity cost of time is the ratio of the marginal utilities of money and time and can be recovered by means of the first order conditions of the optimization of the direct utility (equation (A4) in Appendix B). Then, full prices are calculated supposing either that the two inputs of the domestic productions (money and time) are complements, as in the case examined by Becker (1965), or substitutes, as supposed by Becker and Michael (1973) and Gronau (1977). Both assumptions give rise to similar estimates of the price effects (see Alpman and Gardes, 2016 and Gardes, 2017). The definition of these two variants of the full prices is presented in Appendix.

The absolute and tangent partial derivatives $\nabla_k v^i$ and $\partial_k v^i$(identified with the cross-section and time-series derivatives) depend on the specification of the function between the explained variable (partial expenditure) and the set of explanatotory variables. Therefore, the estimated empirical curvature indicates wether there remains some unexplained endogeneity in the cross-section estimation, i.e. an effect of endogenous permanent latent variables over the cross-section dimension, which disapear for variations through time. In that case, the specification is not sufficient to take care of this remaining endogeneity. The comparison between the curvatures found for different specifications (for instance a linear and a Working) thus gives rise to a non-parametric specification test, in the sense that a specification takes care more completly of the endogenous biases created by permanent latent variables. The importance of the remaining curvature (once the curvature linked to the specification has been taken into account) indicate the existence of an endogeneity bias when estimating on the surface (i.e. using the dataset given by the survey) instead of estimating on tangent planes where latent variables are constant.

The estimates of the Riemann curvatures for food expenditures explained by a Working specification are presented in Table 1. The estimates for other expenditures and sub-population defined by age classes or family composition are presented in Appendix. Also, estimates corresponding to the linear model and the double-logarithmic specification can be found in Appendix.

Food	All households	1st level of education	2nd level of education	3rd level of education
R_{223}	0.0437	-0.1146	-0.0749	-0.0136
s.e.	(0.0093)	(0.0275)	(0.0162)	(0.0609)
R_{332}	0.0418	-0.0231	0.0299	0.2946
s.e.	(0.0541)	(0.1217)	(0.0840)	(0.2756)
$\mathrm{R}_{12}\mathrm{R}R_{332}$	18.27x10^{-4}	26.47x10^{-4}	-22.40x10^{-4}	-40.07x10^{-4}
$\frac{R_{223}+R_{332}}{2}$	0.0427	-0.0689	-0.0225	0.1405
s.e.	0.0274	0.0624	0.0428	0.0706
N	3,052	915	1,870	267

Table 1. Curvature for food expenditure

Note: standard error in parentheses. Bootstrap for Gaussian curvature.

Specification of the surface: $w_{iht} = \alpha_i + \beta_i \ln m_{ht} + \gamma_i \ln \pi_{iht} + Z_{ht}\delta_i + \varepsilon_{iht}$.

Explanatory variables: lnm_{ht}= logarithmic total expenditure per Consumption Unit, $\ln\pi_{iht}$= logarithmic full price, Z_{ht}= log age of the head and survey dummies.

Data: The Polish panel covering 3,052 households for the period 1997–2000.

Table 1 presents the estimated curvatures corresponding to the endogeneity bias on the income effect (R_{223}) or the price effect ($R^{'}_{12}$) in the Working specification. The Riemannian curvatures are normalized, being divided by the average budget share of the commodity (since the specification is written in that term as concerns the explained variable for the Working model). Estimation over the whole population in all four waves (Table 1) generates a Riemannian curvature significantly different from zero regarding the derivative k with respect to total expenditure for all consumptions R_{12}, except Clothing (which corersponds to 8% of total expenditure). The estimation for sub-populations (Table A1 in Appendix) by age group (20-34, 35-55, over 56), three levels of education or family types (singles, two adult without children, families with children) indicates also significant curvatures with a majority of declining value for age classes and family size, while it tends to increase with the education level. The estimates on sub-populations are less significant than for the whole population, which could be explained by the greater homogeneity of these sub-populations and a weaker effect of latent variables between similar households. The curvatures differ between sub-populations, as concerns their sign as well as their level, which is normal as each expenditure is impacted by specific endogeneity biases due to specific latent variables.

The curvature corresponding to the k-derivative with respect to the full prices R'_{12} is not significantly different from zero except for Clothing and Other expenditures. This indicates that endogeneity biases are less significant as concerns the price effect. The results show also more heterogeneity of conditions of choice for large families and more educated families.

The curvature for the linear model is significant for almost all commodities (21 over 24 at 5%); generally greater than for the double-log model (17 only significantly differentfrom zero) which means that the assumption of a constant income elasticity fits better the dataset than supposing the constancy of the marginal propensity.

Different signs and levels linear compared to double-log, which shows that these model differ as conern their proximity to the dataset.

Note that the level of the curvature for the linear model ($2.2x10^{-5}$) is comparable with the curvature of the surface induced by the Working model (-1.4x10^{-5}) (but with an opposite sign).

2.3. Discussion

The consequence is that the metric changes from one point to another: on a Riemannian space, the differential metric writes $\mathrm{d}s^2 = g_{ij}\,\mathrm{d}v^i\mathrm{d}v^j$, with $\mathrm{d}v^i$ and $\mathrm{d}v^j$ the changes in variables v^i and v^j. Indeed, minimising the Riemannian distance over the whole space corresponds to calculate the usual estimate by Maximum Likelihood or least squares, constraining the optimization by the change of the shadow prices from one point to another: the metric at each point changes according to the non-monetary resources and the constraints, which characterize this location in Riemannian space. The curvature of the Riemannian space thus indicates the heterogeneity of these conditions of choice over the population. Therefore it is not possible to recover consumption functions by minimising a unique distance over the whole space: the distance between some point M and its estimate M_1 depends on the coefficients g_{ij} which are attached to this point, if the two points are sufficiently close so that the same metric applies to them. For another point M', it will be necessary to take into account the change of the metric when computing the distance from its estimate M'_1. Moreover, if points M and M' are not sufficiently close, the metric changes continuously on every path between them, so that it is necessary to determine the minimum path between them in the Riemannian space, i. e.

a geodesic between the two points, by integrating ds between the two points.

A solution would be to project the Riemannian space on an Euclidian space, where a unique distance is defined. It is well known that every Riemannian manifold of dimension n can be embedded in an Euclidian space of greater dimension. The first embedding theorems were proved by Janet and Cartan, the Euclidian space necessitating at least $\frac{n(n+1)}{2}$ dimensions (see **ivey2003cartan**), but the general theorem was proved by extraordinary methods by John Nash (1956): his Theorem 3 realizes the imbedding of a non-compact n-manifolds in $\frac{n(n+1)(3n+11)}{2}$ dimensions. We propose a simple projection of the Riemannian consumption space into an Euclidian space by defining a unique metrics defined by the position of the household in its life cycle.

Indeed, another way of estimating the curvature of the space would be to estimate the equation of geodesics on cross-sections: whenever the curvature is non-null, the geodesics are not straight lines, because the basis changes along the geodesic. Thus, a new econometrics on Riemannian space must be defined by considering geodesics. In cross-sections, geodesics represent the shortest paths from one point to another. So, an agent can be supposed to follow a geodesic over her life cycle. The definition of a time dimension on cross-section, discussed in the subsequent section, may permit the estimation of Christoffel symbols along geodesics. Considering households, one solution consists in taking into account the age of the family head or, for firms, the number of years the firm has been active in the market as measuring the passage of time in the survey. Thus, the priority of one variable over another is defined by the priority of the agents over which they are measured. For instance, the past and future consumption of some household h in a dynamic model (representing habit and addiction effects) can be instrumented by the expenditures of similar households (by age, education, location, family structure) one year younger or older in the survey[1]. One drawback is that age is correlated with cohort effects in cross-sections. We propose in section 5 a formula to correct this cohort effect and find that such instrumentation is very effective compared to usual instruments such as past and future prices.

3 Endogenous prices

pb: latent var representable by prices? Condition.

The effect of latent variables which produce endogenous biases could be represented by the cost they imply for the explained variable: for instance, a rationing scheme can be taken into account by the virtual price corresponding to the constarint. This implies that the economic price may differ between households, even if monetary prices are the same, which in fact is not usually the case. Prices are known to vary according to the agent location, sometimes its age (special tarifications) or other socio-economic characteristics of the agent (taxes according to the income class). Also, Barten (1964) explains by the public nature of some consumptions, why, in a simple model of the allocation of goods in the family, the relative price of individual consumptions (such as milk) compared to a

[1]Note that it is generally necessary when estimating dynamic models to instrument past values of the variables, even when these past values are observed.

public consumption such as heating, varies according to the demographic structure of the household (being greater for larger families). These economic costs are named here endogenous prices. Two important examples of endogenous prices are presented hereafter.

3.1 The shadow price of a constraint

Neary and Roberts (1980) compares demand theory without or under rationing by decomposing the set of commodities into a subset of goods that can be freely chosen and a second subset of rationed goods. They show that, under usual condition for the existence of implicit functions (quasi-concavity of the utility function and differentiability at the optimum), the restricted demand functions can be considered as unrestricted choices with virtual prices defined as those which would induce an unrationed agent to choose the same bundle set as the rationed agent [1]. Thus, unconstrained estimation of a system of demand adding virtual prices to the set of market prices can be substituted to the estimation under the constraints corresponding to these virtual prices. Barten and Bettendorf (1995) and Gardes and Starzec (2014) provide applications of this method to the computation of virtual prices of the rationing of housing in Belgian for the Interwar period and Poland before its liberalization.

3.2 Full price in the domestic production model

Becker (1965) considers a set of final goods which are the arguments of the consumer direct utility of the consumer. In order to simplify the analysis, Becker states that a separate activity i produces the final good in quantity Zi using a unique market good (an hypothesis which can be generalized easily) in quantity xi and time ti per unit of activity. The market good and time are supposed to follow a Leontief technology: The full price π_i for one unit of the final good (activity) i can be written: $p_i x_i + \omega t_i$ with an opportunity cost of time ω which is usually taken as the agentâs market wage rate net of taxes. The unit full price of the market good x_i is therefore: $p_i + \omega\tau_i$. It is assumed here that the agent's opportunity cost ω differs from her net wage, so that the full budget constraint is:

$$\sum_i (p_i x_i + t_i) = y^f - w(T - t_w) = y^f - w \sum_i \tau_i x_i \tag{13}$$

In this formula, the full income is corrected by means of a function of the domestic production time which represents the difference between the market (w) and the personal valuation (ω) of that time. An important difficulty in this generalization of the domestic production model lies in the valuation of time. A Cobb-Douglas specification for both the utility function and the domestic production functions allows estimating locally (for each household) the opportunity cost of time by means of the first order conditions for the substitution between time and monetary resources used for the domestic production (estimation performed on data grouped into 40 homogenous cells to take care of possible errors of measurement, see Gardes, 2014). The average of the estimated opportunity cost

[1] Barten and Betteldorf (1995) show how virtual prices can be computed by means of the estimation of a complete demand system.

is 6.74 euros (with a standard-error of 1.5), close to the minimum wage rate and about two third of the householdsâ average wage rate. It is positively indexed on the householdâs net wage (with an elasticity of 0.85) and on income (conditional to net wage: elasticity of 0.19). In the application, the value of time will be defined either by the householdâs average wage net of taxes, the minimum wage rate or this estimated value. Note that both full budget shares and the full income and full prices depend on the opportunity cost which is choosen, which involves a systematic endogeneity in the demand equation. The empirical application shows that the resulting price elasticities do not depend much on the opportunity cost choosen for these variables.

3.3 Definition of full prices

The full expenditure for one unit unit of activity i is the sum of its monetary cost and the opportunity cost of time used for that activity. It depends on the households' characteristics, both through the household's opportunity cost of time and its domestic production technology. Two definition of full prices are proposed in Appendix: first under the assumption of substituability between the factors of production of the activity, time and money, second under complementarity (as in Becker's original model). With these definitions, it is possible to proxy the changes in the full prices, observing only monetary and full expenditures.

Two hypotheses were necessary to derive full prices from monetary and time expenditures: first, domestic production functions are supposed to be Leontief functions (note that the parameters of these functions are specific for each household); second, absence of joint production is supposed, which may be more easily verified for broad categories of activities such as housing and food (see Pollack-Wachter, 1976, for a discussion). Full price are computed in Gardes (2015) by matching a French Family Budget with a Time Use survey using either an exogenous opportunity cost of time or an estimate. The estimated price elasticities compare well with the estimates by other methods. The method affords also elasticities of consumption with respect of time and the opportunity cost of time.

3.4 Empirical evidence

Full prices are shown to depend on macroeconomic conditions, since the opportunity cost of time diminishes during recessiosns (see Alpman and Gardes' analysis of the great recession in the US). An estimation on French data (Gardes, 2016) shows that full prices depend on the opportunity cost of time and on the amount of time consecrated to each activity. Their variation (see Table A4 in Appendix) is quite large across the population (for instance, the full price for transport under substitutability varies from 1.003 to 6.66, with a mean 1.94 and a standard error 0.54). The OCT and time use may both depend on the household's level of being, its situation in its life cycle (family size, age of the head and the spouse), its location… A regression analysis of the logarithm of full prices on these variables shows that all prices decrease with income, with an average elasticity -0.30 (from -0.13 for transport to -0.15 for food and housing, -0.4 for clothing and other commodities and -0.75 for leisure). They increase with age (except for other expenditures) with an average elasticity 0.4, and decrease with the proportion of children in the family.

4 Shadow Prices

Differences between cross-section and time-series estimates of demand functions are commonly observed in recent empirical work: for instance, Gardes et al. (2005) analyse the bias in income and total expenditure food elasticities estimated on panel or pseudo-panel data caused by measurement error and unobserved heterogeneity. This bias suggests that unobserved heterogeneity imparts a downward bias to cross-section estimates of income elasticities of at-home food expenditure and an upward bias to estimates of income elasticities of away-from-home food expenditure. Moreover, the magnitude of the differences in elasticity estimates across methods of estimation is roughly similar in U.S. and Polish expenditure data: for instance, despite some differences between the estimations: the relative income elasticity of food at home is around 0.2 based on a number of different methods with PSID data (1984–1987), while the time-series estimates (within or first differences) are 0.4. The same order between income elasticities in the cross-section or time-series dimensions is found for Polish households observed in a four years panel (1987–1990), with higher elasticities as can be expected for a much poorer country. On the contrary, the cross-section elasticity for food away from home is estimated as much higher than the time-series one. Similar results have been obtained on pseudo-panels of French and Canadian surveys (Cardoso et al. ;Gardes et al., 1996). Generally speaking, such endogeneity biases are shown to exist in the cross-section estimates for half of the commodities. A recent empirical analysis by Christensen (2014) on a spanish panel data as well as estimations performed previously on pseudo-panles (see Gardes et al,) show that the cross-section income effect is generally significantly greater for most expenditures on services, while changes in expenditure on housing over time are more strongly related to income changes than are the differences between two households in the same survey.

One way to explain these differences is to consider the relationship between the relative position of the agent in the income distribution and his non-monetary resources (such as time) or the presence of constraints (such as subsistence constraints) on choice, which both can be related to virtual prices.

Suppose that the *monetary price* p_m and a *shadow price* π corresponding to *non-monetary resources* and to *constraints faced by the households* are combined together into a complete price. Expressed in logarithmic form, we have: $p_c = p_m + \pi$.

Consider two estimations of the same equation:

$$x_{iht} = \boldsymbol{Z}_{ht}\beta_i + p_{cht}\gamma_i + u_{iht} \tag{14}$$

for good i ($i = 1, \ldots, n$), individual h ($h = 1, \ldots, H$) in period t ($t = 1, \ldots, T$), with $\boldsymbol{Z}_{ht} = (\boldsymbol{Z}_{1ht}, \boldsymbol{Z}_{2ht})$. These estimations are carried out on cross-section and time-series data using the same dataset.

Set $u_{iht} = \alpha_{ih} + \varepsilon_{iht}$, where α_{ih} is the specific effect which contains all of the permanent components of the residual for individual h and good i. As discussed by **mundlak1978pooling** the cross-section estimates can be biased by a correlation between the explanatory variables $\boldsymbol{Z}_{ht}$ and the specific effect. This can result from latent permanent variables (such as an event during childhood, parents' characteristics, or permanent wealth) which are related to some of the explanatory cross-section variables $\boldsymbol{Z}_{ht}$: for instance, the relative income position of the household can be related to its wealth or its genetic inheritance. Thus, the correlation δ_i between the time average of the vector of the explanatory variables, $\boldsymbol{Z}_{ht} = \left\{z^k_{1ht}\right\}_{k=1}^{K_1}$, transformed by the Between matrix:

$$\boldsymbol{B}\boldsymbol{Z}_{ht} = \left\{\frac{1}{T}\sum_t z^k_{ht}\right\}_{k=1}^{K_1},$$

and the specific effect α_{ih}, $\alpha_{ih} = \boldsymbol{B}\boldsymbol{Z}_{ht}\delta_i + \eta_{ih}$, will be added to the parameter β_i of these variables in the time average estimation: $\boldsymbol{B}x_{iht} = \boldsymbol{B}\boldsymbol{Z}_{ht}(\beta_i + \delta_i) + \eta_{ih} + \boldsymbol{B}\varepsilon_{iht}$, so that the between estimates are biased. The difference between the cross-section and the time-series estimates amounts to δ_i.

Let us now assume that the shadow price π_{iht} of good i for household h in period t, depends on variables $\boldsymbol{Z}_{1ht}$, which also appear in the consumption function for good i:

$$x_{iht} = g_i(p_{ht}, \boldsymbol{Z}_{ht}, \boldsymbol{S}_{ht}) + u_{iht}$$

with p_{ht} the vector of prices p_{jht} containing (if it exists) a shadow, unknown component π_{jht}, and $\boldsymbol{S}_{ht}$ the vector of all other determinants.

We now assume that only the monetary component of prices change over time (the shadow component being related to permanent variables), while the different agents observed in the cross-section survey are characterized by different non-observed shadow prices (corresponding to individual non-monetary resources and constraints). Equation (14) writes on time-series (for instance in first differences between periods):

$$x_{iht} = \boldsymbol{Z}_{ht}\beta_i + p_{miht}\gamma_i + u_{iht}$$

while on cross-section it is, supposing the price effect γ_i and monetary prices are the same on both dimensions:

$$x_{iht} = \boldsymbol{Z}_{ht}\beta_i' + u_{iht}' = \boldsymbol{Z}_{ht}\beta_i + \pi_{iht}\gamma_i + u_{iht}$$

with obvious notations. Thus, the difference between the two estimations is:

$$\boldsymbol{Z}_{ht}\delta_{1i} = \boldsymbol{Z}_{ht}\theta_1\gamma_i' + (\boldsymbol{S}_{ht}\theta_2 + \lambda_{ih} + \mu_{iht})\gamma_i$$

which allows to calculate the set of parameters θ_1 after calibrating the price effect measured by γ_i.

The marginal propensity to consume with respect to $\boldsymbol{Z}_{1ht}$, when considering the effect of the shadow prices π_{jht} on consumption, can be written as:

$$\frac{\mathrm{d}x_{iht}}{\mathrm{d}\boldsymbol{Z}_{ht}} = \frac{\mathrm{d}g_i}{\mathrm{d}\boldsymbol{Z}_{ht}} + \sum_j \left(\frac{\mathrm{d}g_i}{\mathrm{d}\pi_{jht}}\right)\left(\frac{\mathrm{d}\pi_{jht}}{\mathrm{d}\boldsymbol{Z}_{ht}}\right).$$

The second term will differ between cross-section and time-series because of the correlation of the shadow price with the endogenous variables $\boldsymbol{Z}_{ht}$. So, comparing two different households surveyed in the same period, this bias adds to the direct unbiased consumption propensity with respect to $\boldsymbol{Z}_{ht}$, as estimated on time-series data. For instance, the influence of the household head's age cohort or income may differ in cross-section and time-series estimations if the shadow prices depend on cohort effects or on the relative income position of the agent (note that the same effect may occur with respect to monetary prices).

The term $\sum_j \frac{\mathrm{d}g_i}{\mathrm{d}\pi_{jht}} \frac{\mathrm{d}\pi_{jht}}{\mathrm{d}\boldsymbol{Z}_{ht}}$ above can be used to reveal the variation of shadow prices over $\boldsymbol{Z}_{1ht}$, $\frac{\mathrm{d}\pi_{jht}}{\mathrm{d}\boldsymbol{Z}_{ht}}$, since it can be computed by resolving a system of n linear equations after having independently estimated the price marginal propensities $\frac{\mathrm{d}g_i}{\mathrm{d}\pi_j} = \gamma_{ij}$. We can also consider only the *direct effect* of the variables $\boldsymbol{Z}_{ht}$ through the price of good i, $\gamma_{ii}\frac{\mathrm{d}\pi_i}{\mathrm{d}\boldsymbol{Z}}$, so that:

$$\frac{\mathrm{d}\pi_i}{\mathrm{d}\boldsymbol{Z}} = \frac{\beta_i^{\text{c.s.}} - \beta_i^{\text{t.s.}}}{\gamma_{ii}}. \tag{15}$$

The price effect γ_{ii} is supposed to be the same for monetary and shadow prices. Thus, the change in the shadow price between two periods can be written as: $\mathrm{d}\ln\pi_{iht} = \sum_k \left(\frac{\mathrm{d}\pi_i}{\mathrm{d}z^k}\right) \mathrm{d}z^k_{ht}$. Under homogeneity (of degree m) of shadow prices over variables $\boldsymbol{Z}_{1ht}$, the shadow prices can be computed as $\ln\pi_{ih} = m\sum_k \frac{\mathrm{d}\pi_i}{\mathrm{d}z^k} z^k_{ht}$. However, this homogeneity assumption is quite strong, and we will prefer to compute only the change in the log shadow prices[1].

The income elasticities of the shadow prices of food at home and food away from home expenditures are computed in Table **??** for the PSID and the Polish panel (using equation 15, and assuming that direct price elasticities are minus one half of the corresponding income elasticities). These estimated parameters are remarkably similar in both countries: positive and smaller than one for food at home, so that the full price of food at home is greater for richer than for poorer households. One interpretation is that rich people are time constrained and have a larger opportunity cost for the additional time spent on food at home compared to food away. This difference may be thought to be greater in the USA than in Poland. On the contrary, the income elasticity of the full price for food away from home is negative, and of the same magnitude in both countries.

In our analysis, the relation between cross-section and time-series estimates, modelled by shadow prices, is supposed to be linear over all of the distribution. It is however likely that the derivatives $\frac{\mathrm{d}\pi_{jht}}{\mathrm{d}\boldsymbol{Z}_{1ht}}$ also depend on individual characteristics. This is for instance the case in Barten's model (**barten1964family**) where relative prices depend on the family structure. This local dependency requires a geometric characterization of the consumer space.

[1] This model is presented more thoroughly and applied to rationality tests in **diaye2001world**

Empirical evidence

Boelaert-Gardes-Langlois (L'Actualité Economique, 2018) apply this method to a pseudo-panel of seven canadien Households Expenditures surveys from 1978 to 2008. The shadow prices are shown to change along the income distribution, increasing with relative income for Food at Home, Alcohol, Tobacco and Private Transport, decreasing for Public Transport, Health expenditures and Education. The recent widening of income inequality in Canada implies that these virtual prices are more inequal in the whole population, which make the conditions of the economic choices motre inequal between rich and poor households (a supplementary inequality between them). The virtual prices are also shown to be correlated with full prices (Table 5 in Boelaert et al. article), which are one component of the economic cost.

Conclusion

We have shown that the consumption space has a Riemannian structure, which allows us to connect the time dimension of consumption laws to the social distribution of consumption expenditure in the population. Riemannian curvature of consumption means that there exists, for the social distribution of consumption, path-dependency with respect to the order of the changes in the different variables influencing consumption choices. For instance, comparing a couple of two adults to a family with children in cross-section allows us to compute an equivalence scale which may depend not only on income levels, but also on income changes, in the sense that an increase in family size $\mathrm{d}s$ for a family of size s and income $y_1 = y_0 + \mathrm{d}y$ may not give rise to the same levels of expenditure as an increase in income $\mathrm{d}y$ for a family with size $s + \mathrm{d}s$ and income y_0 (whenever the condition for point integrability, i. e. the symmetry of Christoffel symbols, does not hold). Thus, considering the Riemannian structure of the expenditure space may be useful for Lewbel's problem of identification in equivalence scale models[1].

Second, the impossibility of defining a unique metric for the whole population means that usual econometric estimations of consumption laws on cross-section data are misleading. This is due to the fact that local conditions of choice, which correspond to local shadow prices, are not taken into account in the estimation. In a sense, the Riemannian curvature of the consumption space can be related to social heterogeneity, and the change of the basis from one point to another in this space indicates the variation in the various constraints and non-monetary resources which influence consumer choice by means of the associated shadow prices. The conditions of choice depend on the situation of the agent, i. e. his or her location in Riemannian space. Barten's (1964) discussion of the change in relative monetary prices due to changes in family composition can be considered as an example of this relationship.

We have shown that the substitution between time and the socio-economic determinants of household behaviour makes it possible to estimate dynamic models on cross-sectional individual data. The position of individuals on the synthetic time axis also

[1] I am grateful to Alain Trognon for this suggestion.

provides a natural distance between them, which can be related to the differences which are observed between their expenditure patterns.

Third, preference interdependencies are generally estimated by adding the average quantities consumed by similar households $x_{H,t}$ (see **montmarquette2002large** for a discussion of this specification). This model is highly implausible, though it is generally estimated as significant. In the model we propose in Section 5, this variable corresponds to the instrumented past or future expenditures of the household, so that the so-called interdependence effect can be interpreted as a habit or addiction effect of past and future expenditures. These effects may be inversely related to the distance between similar consumers, which is measured by the time distance between the household and its past and future expenditures[1].

Finally, the geometric structure of a survey could be considered to estimate dynamic models on a cross-section. For a survey on households, it could consists in taking into account the age of the family head or, for firms, the number of years the firm has been active in the market, as measuring the passage of time in the survey. For instance, the past and future consumption of some household h in a dynamic model (representing habit and addiction effects) can be instrumented by the expenditures of similar households (by education, location, family structure) one year younger or older in the survey. One drawback is that age is correlated with cohort effects in cross-sections. To correct this cohort effect, Gardes (2018) uses the fact that in theories such as the life cycle, both the dependent variable and its determinants are related to the time dimension. For instance, considering savings, income per unit of consumption increases early in the life cycle, then falls as household size increases, and finally increases to a constant level at the end of the life cycle, while the savings rate also varies systematically over the life cycle. Therefore, income and age can substitute *locally* as concerns their influence on savings. For example, a rich household aged 35 may have the same saving behaviour as a poor household aged 50, so that income compensates for age in determining savings. Similar relations as those considered with the head's age can be found for other household's characteristics (such as its demographic composition), so that a combination of all may be more efficient to date each household on the time axis. Two methods to correct for cohort effects are presented in Gardes (2018). These corrections are applied to the estimation of a dynamic model of consumption and to the estimation of Christoffel symbols using the equation of geodesics (which relates the second order derivatives over time of some variable at choice with a linear combination of first order derivatives with coeffcients depending on the Christoffel symbols). Finally, the estimation of Christoffel symbols can be used to correct for the endogeneity biases of cross-section estimates.

Appendix A: The Polish Data (1997-2000)

Household budget surveys have been conducted in Poland for many years. In the period analysed, the annual total sample size was about 30 thousand households, which

[1]A test of this interpretation would be to estimate the parameters of the autoregressive operator A over the synthetic time scale, in a model: $x_{ht} = A(L)x_{Ht} + \boldsymbol{W}_{ht}\beta + \varepsilon_{ht}$.

represent approximately 0.3% of all households in Poland. The data were collected by a rotation method on a quarterly basis. The master sample consists of households and persons living in randomly selected dwellings. This was generated by a two-stage, and in the second stage, two-phase sampling procedure. The full description of the master sample generating procedure is given by Kordos and Kubiczek (1991).

On every annual sample it is possible to identify households participating in the surveys during four consecutive years. For the four years panel from 1997 to 2000, 3052 households remain in the data set after deleting a few number of households with missing values. The available information is as detailed as in the cross-section surveys: the usual socio-economic characteristics of households and individuals, as well as information on income and expenditures. A large part of this panel containing demographic and income variables is included in the comparable international data base of panels in the framework of the PACO project (Luxembourg) and is publicly available. The four surveys from 1997 to 2000 have been matched with the 2000 Time use Polish survey by Rubin's method, which assigns an adult observed in the Time use survey to each adult of a household. The usual matching method by regression has two shortcomings: the reduction of the variance of the imputed values and the conditional independency assumption of the variable which are imputed with those which are observed in the dataset. Both problems can be solved to a great extent by the statistical matching procedure proposed by Rubin (1986). This procedure allows its user to assume a partial correlation value between the two variables that are jointly unobserved. This method conserves the individual distribution of the matched times and from this point of view is better than the usual regression method which diminishes drastically the variance of the matched variables compared to their distribution in the survey where they are observed. The matching procedure is discussed in Alpman (2016), Alpman and GArdes (2015) and Alpman, Gardes and Thiombiano (2017).

The 1997-2000 panel corresponds to the post transition high economic growth period with relatively low inflation, decreasing unemployment and generally improved socio-economic situation in the context of almost totally liberalized economy.

Appendix B: Estimation of the opportunity cost of time

To estimate the shadow price of time, I assume (Gardes, 2017) that the consumer combines time with monetary expenditures to produce activities that generate utility in a model where the market work time is valued by the consumer's wage rate while the remaining time (e.g., time allocated to leisure or non-market work) is valued by the shadow price of time that may differ from the wage rate. It is assumed that the consumer's utility function is given by u(Q) = Q i aiQ i i where ai is a positive parameter and Qi is the quantity of the activity i produced by the combination of monetary and time inputs denoted mi and ti, respectively: $Q_i = b_i m_i^{\alpha_i} t_i^{\beta_i}$ where $m_i = x_i p_i$ with x_i the quantity of the market goods i, p_i its price, and b_i a positive parameter. The domestic production Q is assumed to depend on expenditures m_i (rather than quantities x_i) because the dataset informs only expenditures. This approach yields consistent results when all households face the same prices.Expenditures are divided by a yearly price index and the regressions include State and year dummies.

The choice of the Cobb-Douglas forms allows the parameters to be identiable. Note that the estimation is made locally, either over sub-population or locally, for each household, using a non-parametric method similar to the Lowess estimator. The Cobb-Douglas specification implies simply constant substitution between time and monetary resources only in the neighborhood of each individual's equilibrium point. Combining the utility and the production functions allows to write the utility in terms of inputs:

$u(m,t) = (\prod_i a_i^{\gamma_i} b_i^{\gamma_i})(\prod_i m_i^{\frac{\alpha_i\gamma_i}{\sum \alpha_i\gamma_i}})^{\sum \alpha_i\gamma_i}(\prod_i t_i^{\frac{\beta_i\gamma_i}{\sum \beta_i\gamma_i}})^{\sum \beta_i\gamma_i} = Am'^{\sum \alpha_i\gamma_i} t'^{\sum \beta_i\gamma_i}$

where m' and m' are geometric weighted means of the monetary and time inputs and $A = \prod a_i^{\gamma_i} b_i^{\gamma_i}$. In this framework, the consumer is subject to an income constraint, $\Sigma m_i = wt_w + V$, and to a time constraint, $\Sigma t_i + t_w = T$, where t_w is the time allocated to market work, V is other income and Y and T total income and available time (excluding sleeping time which is not included in the time uses t_i). Utility maximization implies that the shadow price of time, denoted ω, is given by $\omega = \frac{\frac{\partial u}{\partial t'}\frac{\partial t'}{\partial(\Sigma t_i)}}{\frac{\partial u}{\partial m'}\frac{\partial m'}{\partial Y}} = \frac{m'\Sigma\beta_i\gamma_i\frac{\partial t'}{\partial(\Sigma t_i)}}{t'\Sigma\alpha_i\gamma_i\frac{\partial m'}{\partial Y}}$ (A1). The shadow price of time differs from the market wage rate when, for instance, there exists some market imperfections, transaction costs, and constraints on the labor market or in the home sector. The shadow price of time can be estimated provided that estimates of α_i, β_i and γ_i are available. Under the assumed functional forms, the optimal combination of the inputs yields $\alpha_i$and β_i: $\alpha_i = \frac{m_i}{\omega t_i + m_i}$ and $\beta_i = \frac{\omega t_i}{\omega t_i + m_i}$. The optimal allocation of an input across the activities i and j implies:

$\frac{\gamma_i}{\gamma_j} = \frac{\beta_j t_i}{\beta_i t_j} = \frac{a_j mi}{a_i m_j}$ and the system of $(n-1)$ equations: $m_i\gamma_j = m_j\gamma_i + \omega\gamma_i t_j - \omega\gamma_j t_i$ which allow to estimate an average opportunitry cost of time $\omega$and all parameters γ.

This system is overidentied and the unknown parameters, that is, ω and each γ_i, can be estimated as a system of $(n-1)$ equations under the constraint $\sum_i \gamma_i = 1$. Using the shadow price of time estimated from the system, we obtain the estimates of α_i and β_i for each individual using equation (4) that allow to derive estimates of the shadow price of time at the individual level by equation (A1).

Appendix C: Full prices in a Becker's domestic production framwork

Full prices under substitutability assumption

Following Gronau (1977) , the full expenditure can be written as the sum of its monetary and time components:

$\pi_i^1 z_i = p_i x_i + \omega t_i$

with and p the full and the monetary prices corresponding to the quantities z and x of the activity and of the corresponding market good, for activity i (note that prices and the opportunity cost of time depend on household and time, which indices are removed).

The full price is the derivative of the full expenditure over z, which writes for the Cobb-Douglas specification of the domestic production functions:

$\mathrm{p}_i^{f1} = p_i\frac{\partial x_i}{\partial z_i} + \omega\frac{\partial t_i}{\partial z_i}$

The optimization program gives rise to the first order condition: $\frac{t_1}{x_1} = \frac{p_1}{\omega}\frac{\beta_1}{\alpha_1}$ (note that all variables, including and , are measured or estimated at the household level). Writing the quantity of the activity z$_i$ in terms, either of t or x, gives:

$t_i = \frac{1}{a_i} z_i (\frac{p_i \beta_i}{\alpha_i})^{\alpha_i}$and $x_i = \frac{1}{a_i} z_i (\frac{\omega\alpha_i}{p_i\beta_i})^{\beta_i}$

So that the full price becomes:

$p_i^{f_1} = \frac{1}{a_i} p_i^{\alpha_i} \omega^{\beta_i} \left\{ \left(\frac{\beta_i}{\alpha_i}\right)^{\alpha_i} + \left(\frac{\alpha_i}{\beta_i}\right)^{\beta_i} \right\}$

This derivation of , and at the individual level allows identifying the full price for each household (a_i being supposed constant across the population).

Full prices under complementarity assumption

Becker's full price for one unit of activity i can be written: p$p_i^{f_2} = p_i + \omega\tau_i$ with τ_i the time use necessary to produce one unit of the activity i. Suppose (as in Becker's model) that a Leontief technology allows the quantities of the two factors to be proportional to the activity:

$x_i = \xi_i z_i$ and $t_i = \tau_i z_i$with τ_i=$\frac{\theta_i}{\xi_i}$

This case corresponds to an assumption of complementarity between the two factors in the domestic technology, which allows calculating a proxy for the full price of activity i by the ratio of full expenditure over its monetary component:

$\pi_i = \frac{(p_i + \omega\tau_i)xi}{p_i x_i} = \frac{p_i + \omega\tau_i}{pi} = 1 + \frac{\omega\tau_i}{pi} = \frac{1}{p_i} p_i^{f_2}$

Note that under the assumption of a common monetary price pi for all households in a survey made in period t, this ratio contains all the information on the differences of full prices between households deriving from their opportunity cost for time _h and the home production technology represented by the coefficient of production _i. If the monetary price p changes between households or periods, the full price can be computed as the product of this proxy π_{ih} with p_i: $p_i^{f_2} = p_i\pi_i$. With these definitions, it is possible to measure the full prices, observing only monetary and full expenditures.

The first definition of prices corresponds to a complete substitution between the two factors in the model (which is used in section 1 to estimate the opportunity cost of time), since the Cobb-Douglas domestic production functions are characterized by a unitary elasticity of substitution between the two factors. Its relies on the estimation of three parameters: , and . On the other hand, the second definition supposes no substitution between the two domestic production factors but it may give a more robust measure of the full prices since it depends only on the estimation of the households' opportunity cost for time . Both definitions of the full prices will be used in the estimation. However, there exists a simple relation between these two definitions of the full prices. Using equations (9) we obtain:

$p_i^{f_1} = \frac{1}{a_i} p_i^{\alpha_i} \left(\frac{m_i}{\omega_{ht} t_i}\right)^{\beta_i} \left\{1 + \frac{\omega t_i}{pi}\right\}$

So that their logarithmic transforms differ only by _i logm_i/t_i on a cross-section: $log(p_i^{f_1}) = constant + \beta_i log(\frac{m_i}{ti}) + log(\pi_i)$ with prices p_i set to one.

Two hypotheses were necessary to derive full prices from monetary and time expenditures: first, the domestic production functions are supposed to be Leontief functions with constant production coefficients (for Becker's prices) or Cobb-Douglas functions (for Gronau's prices); second no joint production exists, which may be more easily verified for broad categories of activities such as housing and food.

Appendix D: Empirical curvature: results

	Food	Housing	Transport	Clothing	
R_{223}, Whole survey	0.0437	0.0238	0.0931	-0.0096	
s.e.	(0.0093)	0.0123	0.0286	0.0074	
R_{332}, Whole survey	0.0418	-0.3643	-0.9361	0.0298	
s.e.	(0.0541)	0.2176	0.4717	0.0109	
$R_{223}R_{332}$, Whole survey	18.27x10^{-4}	-86.70x10^{-4}	-439.15x10^{-4}	-2.86x10^{-4}	
$\frac{R_{223}+R_{332}}{2}$ Whole survey	0.0427	-0.1703	-0.4213	-0.0048	
s.e.	0.0274				
Age 20-34: R_{223}/R_{332}	-0.0967*/-0.0902	-0.0020/-0.2347	-0.0064/0.0061	0.01461*/-0.0379	-0.
Age 35-54	0.0620*/0.0181	0.0796*/-0.4818	0.0135*/0.0029	0.0051/0.0059	-0
Age 55-	0.0139/0.0599	0.0139 /0.0599	0.0338*/0.0032	0.0061/0.0864*	-0.0
Education 1	-0.1146*/-0.0231	-0.0731/-0.0548	0.0351*/0.0022	0.0059/0.0583*	-0.
Education 2	-0.0749*/0.0299	0.0731*/0.4845	0.0165*/0.0083	-0.0192/0.0263	0
Education 3	-0.0136/0.02946	0.1370*/-0.0582	-0.0887*/-0.0039	-0.0117/0.0225	-0
Singles	0.1416*/0.3134	0.0418*/-0.4342	0.1841*/0.0382	0.0355/0.1763	-0.1
2 adults	-0.0037/0.1314	0.0611*/-0.4636	0.0166*/0.0039	-0.0142/-0080	-0.1
Family with children	0.0443* 0.0197	0.0273 -0.4344	-0.0255* 0.0220	-0.0411* 0.0114*	-0.

Table A1. Gaussian Riemannian curvature, Working specification: $\mathbf{R}_{12}\mathbf{R}_{12}^{'}$, Polish Panel, 1997–2000.

* significant at 5%; $\mathrm{R}_{12}\mathrm{R}_{12}^{'}$ multiplied by 10^4

Table A2. Gaussian Riemannian curvature, Working specification: $\mathbf{R}_{12}\mathbf{R}_{12}^{'}$, Polish Panel, 1997–2000.

* significant at 5%; ** significant at 1%

References

Alpman, Anil (2016) \Implementing Rubin's Alternative Multiple-Imputation Method for Statistical Matching in Stata." Stata Journal, 16, 717739.

Alpman, A. et Gardes, F., 2015, Time Use during the Great Recession - A Comment, Cahiers de recherche du CES, 2015.12.

		Food	Housing	Transport	Clothing	Leisure	Ot
Linear	R_{223}	0.000129**	-0.000790**	-0.001628	-0.000145**	0.000111**	-0.00
	R_{332}	0.173236**	-0.449871**	-0.023963**	0.000969**	0.039849**	0.000
	$R_{223}.R_{332}$	2.23x10^{-5}	35.52x10^{-5}	3.90x10^{-5}	-0.014x10^{-5}	0.443x10^{-5}	-8.31
	$\frac{R_{223}+R_{332}}{2}$	0.08668	-0.22533	-0.01280	0.00041	0.01998	0.0
Log-linear	R_{223}	-0.0001079	0.001183**	0.000980**	-0.000601	0.001938**	0.000
	R_{332}	0.001474**	0.002588**	0.000785**	0.003575**	0.002757**	0.000
	$R_{223}.R_{332}$	-0.00158	0.3625x10^{-5}	-0.0769x10^{-5}	0.21474x10^{-5}	0.5343x10^{-5}	0.044
	$\frac{R_{223}+R_{332}}{2}$	0.00074	0.00189	0.00088	0.00149	0.00235	0.0
Working	R_{223}	0.0437**	0.0238*	0.0931**	-0.0096*	-0.0280**	-0.0
	R_{332}	0.0418	-0.3643*	-0.9361*	0.0298**	0.0159	-0.0
	$R_{223}R_{332}$	18.27x10^{-4}	-86.70x10^{-4}	-439.15x10^{-4}	-2.86x10^{-4}	-4.45x10^{-4}	1.34
	$\frac{R_{223}+R_{332}}{2}$	0.0427	-0.1703	-0.4213	-0.0048	-0.0061	-0.

Alpman, Anil, Francois Gardes, and Noel Thiombiano (2017) \Statistical Matching for Combining Time-Use Surveys with Consumer Expenditure Surveys: An Evaluation on Real Data." CES Working Paper, 2017.24.

Angrist, Joshua D. and Alan B. Krueger (1999). Empirical strategies in labor economics, vol. 3A. Ed. by O. C. Ashenfelter and D. E. Card. Vol. 3. North-Holland, pp. 1277–1397.

Barten, Antonius Petrus (1964). "Family composition, prices and expenditure patterns". In: Economic analysis for national economic planning. Ed. by Gordon Mills Peter E. Hart and John K. Whitaker. London: Butterworths.

Becker G.S. and Michael. "On the New Theory of Consumer Behavior." Swedish Journal of Economics, 75.4 (1973): 378-396.

Bettendorf Léon, Barten Anton. Rationnement dans les systèmes de demande : calcul des prix virtuels. Économie & prévision. N 121, 1995-5.

Boelaert, J., Gardes, F., Langlois, S., 2017, Consommation marchande et contraintes non monétaires au Canada (1969-2008), L'Actualité Economique-Canadian Review of Economic Analysis.

Christensen, B., 2014, Heterogeneity in Consumer Demands and the Income Effect: Evidence from Panel Data, The Scandinavian Journal of Economics, 116, 2, 335-355.

Cardoso, N., Gardes, F., 1996b, Estimation de lois de consommation sur un pseudo-panel d'enquêtes de l'INSEE (1979, 1984, 1989) , Economie et Prévision, 5 : 111–125.

Collet, R., Gardes, F., Starzec, C., 2011, Does Addiction Matters in Transportation, w.p. CES, Université Paris I.

Delachet, André (1969). Le Calcul Tensoriel. Vol. 1336. Presses universitaires de France.

Diaye, Marc-Arthur, François Gardes, and Christophe Starzec (2004). Non-parametric Tests of Demand Theory and Rational Behavior. Annals of Economics and Statistics.

Gardes, F., 2016, The estimation of price elasticities and the value of time in a domestic framework: an application on French micro-data, under revision in Annals of

Economics and Statistics.

Gardes, F., 2018, Time as a social distance, w.p. PSE, University Paris I, May.

Gardes, F., G.J. Duncan, P. Gaubert, C. Starzec, 2005, A Comparison of Consumption Laws Estimated on American and Polish Panel and Pseudo-Panel Data, Journal of Business and Economic Statistics, April.

Gardes, F., Langlois, S., Richaudeau, D., 1996, Cross-section versus Time Series Income Elasticities of Canadian Consumption, Economics Letters, 51, 169-175.

Gardes, F. and Starzec, C., 2014, Evidence on Addiction Effects from Households Expenditure Surveys: the Case of the Polish Panel, w.p. CES, Université Paris I.

Gronau R. "Leisure, Home Production, and Work – The theory of the Allocation of Time revisited." Journal of Political Economy 85 (1977): 1099-1123.

Jeanperrin, Calcul Tensoriel,

Heckman, James L. (1999). Causal parameters and policy analysis in economcs: A twentieth century retrospective. Tech. rep. National Bureau of Economic Research.

Hicks, John Richard (1979). Causality in economics. Basic Books New York.

Ivey, Thomas A. and J. M. Landsberg (2003). "Cartan for beginners". In: Graduate Studies in Mathematics 61.

Kordos, J. and A. Kubiczek (1991). "Methodological Problems in the Household Budget Surveys in Poland". In: GUS, Warsaw.

Lichnerowicz, André (1987). Eléments de calcul tensoriel. Jacques Gabay.

Morgan, Frank (1998). Riemannian Geometry: A Beginners Guide. AK Peters/CRC Press.

Mundlak, Yair (1978). "On the pooling of time series and cross section data". In: Econometrica: Journal of the Econometric Society, pp. 69–85.

Neary, J.P., Roberts, K.W.S., 1980, The Theory of Household Behaviour under rationing, European Economic Review, 19, pp. 25-42.