



Population Heterogeneity and the Nonprofit Sector in the United States : Global versus Local Spatial Approaches

Marie Lebreton, Katia Melnik

► To cite this version:

Marie Lebreton, Katia Melnik. Population Heterogeneity and the Nonprofit Sector in the United States : Global versus Local Spatial Approaches. 2006. halshs-00410518

HAL Id: halshs-00410518

<https://shs.hal.science/halshs-00410518>

Preprint submitted on 21 Aug 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GREQAM

**Groupeement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III**

**Document de Travail
n° 2006-25**

POPULATION HETEROGENEITY AND THE NONPROFIT SECTOR IN THE UNITED STATES : GLOBAL VERSUS LOCAL SPATIAL APPROACHES

**Marie LEBRETON
Katia MELNIK**

May 2006

DT-GREQAM

Population Heterogeneity and the Nonprofit Sector in the United States: Global versus Local Spatial Approaches

Marie LEBRETON* and Katia MELNIK†

June 21, 2006

Abstract

In this paper we study the relationship between the relative size of the U.S. nonprofit sector and population heterogeneity, at the county level, by adopting and extending the model of Alesina, Baquir and Easterly (1999). The relative size of the voluntary sector is assessed as a share of voluntary donations in the total public good provided locally via public expenditures and private contributions.

We demonstrate empirically that the relative size of the nonprofit sector in each county depends not only on its population heterogeneity, but also on its neighbors' average relative size of the nonprofit sector and average population heterogeneity. Moreover, this relationship seems to be unstable across counties as the signs and magnitudes of neighborhood effects vary with geographical location.

JEL classification: C21, D71, H41.

Keywords: Public good, Nonprofit sector, Parameter heterogeneity, Spatial models

1 Introduction

Since the seminal work of Weisbrod (1977, 1986), voluntary organizations are recognized as alternative mechanisms of private provision of collective-type goods. Though criticized (e.g. DiMaggio and Anheier, 1990), Weisbrod's approach remains the dominant theoretical perspective in the nonprofit field.

*Marie LEBRETON, GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13326 Marseille Cedex 2, France. Email: egen_us@yahoo.com

†Corresponding author: Katia MELNIK, GREQAM, Vieille Charité, 2 rue de la Charité, 13002 Marseille, France. Email: melnik@ehess.univ-mrs.fr. The authors thank summer school participants at CORE (Louvain-la-Neuve) for comments and suggestions. Further, the authors gratefully acknowledge Jean-Benoît Zimmermann and Anne Péguin-Feissolle for very helpful suggestions and comments.

He pointed out the role of voluntary nonprofit sector as a “*major provider of collective-consumption goods which enter positively utility functions of a limited set of persons representing a demand insufficient to organize governmental or market provision of excludable goods*”. Thus the voluntary sector is considered as an extra-governmental provider of collective consumption goods which represents an alternative for unsatisfied demanders. The unsatisfied demands arise from the fact that in a majority voting political system, political decisions about provision of local public goods tend to satisfy the demand of the median voter. Government finances the public provision of public goods by compulsory taxation. In a democratic society with heterogeneous demand, voters who are not satisfied by the public provision of collective good can either move to another jurisdiction or form a voluntary organization to supplement the government’s provision. The broadly quoted Weisbrod’s theory predicts that “*the relative size of the voluntary sector in a field [e.g. education or health] can be expected as a function of the heterogeneity of population demands. Thus, for any given level of governmentally provided output, the market share of the voluntary sector in the provision of collective goods will vary directly with the heterogeneity among individual demand schedules for these goods*”(p.61). While Weisbrod did not propose any explicit model, he clearly augured that “*if two political units differ in the degree of heterogeneity of their populations, the more homogenous unit will, ceteris paribus, have a lower level of voluntary sector provision of collective-type goods or their private-good substitutes*” (idem).

In the same perspective, James (1993) presented an empirical cross-country study of educational services provided both by public and nonprofit sectors. She assumed that the relative size of the nonprofit sector in education was a function of “excess demand”, “differentiated demand” for education, cultural and income heterogeneity of population, and public policies towards education (public educational spending and public subsidies to private education). The excess demand is supposed to stem from people who are unsatisfied by the publicly provided level of education or excluded from public provision because of low public spending. The differentiated demand flows from people with preferences for specific features which result from cultural, religious or other affiliations. She studied the relative size of the nonprofit sector in the field of education with the percentage of enrollments in private nonprofit schools used as proxy.¹ She finds that cultural (and particularly religious) heterogeneity of populations is an important determinant of the larger size of the nonprofit educational sector.

Lassen (2006) provides a cross-country study explaining the size of the informal sector by the degree of country’s ethnic heterogeneity. He considers the

¹The private sector is considered here as a synonym of the nonprofit one. According to James (1993), in many countries “nonprofit status is legally required for educational institutions”.

informal sector as opposed to the public one and supposes that in ethnically heterogeneous societies, ethnic communities may constitute an alternative for public provision of some public goods. He finds a positive effect of population ethnic heterogeneity on the size of the informal sector. While using regional dummy variables, he does not find any significant regional pattern in the studied relationship.

However, the effects of populations' heterogeneity on the relative size of voluntary sector are not straightforward. On the one hand, it is a source of economic and other individual incentives to deliver a public good, which is under-provided or provided in an unsatisfactory manner by state or market. On the other hand, previous literature indicated that heterogeneity may also be a source of inefficiencies in the provision of local public goods (Alesina et al, 1999, Miguel and Gugerty, 2004) and charitable giving in the U.S.A. (Okten and Osili, 2005). In these studies, the economic and social heterogeneities of population assessed through income inequality and ethnic or religious fractionalization (Alesina et al, 2003) negatively influence the provision of public goods. Moreover, an empirical study of Alesina and La Ferrara, (2000) on the effects of population heterogeneity on membership in groups suggests that income and ethnic heterogeneities diminish the propensity to participate in social activities (measured as membership in groups), particularly for "face-to-face" groups, where direct interactions among people are important. The idea underlying these studies is that heterogeneity in agents' characteristics may be translated into heterogeneous preferences towards public good. The question we raise in this paper concerns precisely the effects of heterogeneity in population characteristics on the portion of local public good generated through private donations.

Methodologically, we investigate the effects of population heterogeneity on the relative size of the U.S. nonprofit sector using three different approaches. The first one, the standard linear regression model estimated by OLS assumes that the relationship under study is stationary over space and that there is no spatial dependence in the observations. As in the empirical papers mentioned above, it is implicitly assumed that observations are spatially independent, i.e. there are no interactions between neighboring counties.

However, spatial dependence may arise when we deal with located observations because of measurement errors for observations or because some unobserved economic and social phenomena present a spatial structure leading to complex interactions. Therefore, the second estimation approach incorporates explicitly spatial dependence through the inclusion of the spatially lagged dependent and independent variables. The maximum likelihood estimation of the resulting model called a Spatial Durbin model (SDM) allows us to demonstrate the existence of geographical spillovers among U.S. counties as the relative size of the nonprofit sector in a county depends not only

on the values of its own independent variables, but also on the average values of the dependent and independent variables in the neighboring counties. In order to test the robustness of the SDM results, a Matrix Exponential Spatial Structure (MESS) model is also estimated. The MESS results support the SDM ones.

The possible spatial non stationarity of the relationship under study is also ignored and results in global modelling. However, the spatial non stationarity of the relationship across counties may result from misspecification in the model (e.g omitted variables, a wrong functional form) or from intrinsically different behaviors of local governments. As pointed out by Durlauf (2002), Durlauf and Fafchamps (2003), if the distribution of a given error depends on its associated geographical area, then a model allowing for parameter heterogeneity is appropriate to fit the data. In other words, parameter estimates may differ across geographical units (Brock and Durlauf, 2001). Although ignoring this aspect may lead to misleading conclusions concerning studied relationships between variables, especially in cross-country analysis, most empirical works do not allow for parameter heterogeneity. For instance, Salamon et al (2000) looked at the relationship between the size of nonprofit sector (measured as a percentage of total nonagricultural employment) and the population's religious heterogeneity in different countries. Based on the estimations of a linear regression, they concluded that no significant relationship was detected. According to their data, *"...countries with similar levels of fractionalization, such as Colombia and Ireland, or the Netherlands and the U.K., vary considerably in the size of their nonprofit sectors"* (p. 10.) However, does it really make sense to suppose that an increase in ethnic heterogeneity in Latin America and in Europe would have the same effect on the size of the nonprofit sector?

To take into account the possible spatial non stationarity, the assumption of spatial stationarity is alleviated. We use the Spatial Autoregressive Local Estimate (SALE) model developed by Pace and LeSage (2002) as it accommodates simultaneously for spatial dependence and parameter heterogeneity. The SALE results detect spatial variation in the parameter estimates meaning that the relationship under study is not stable across counties.

The paper is organized as follows. The theoretical foundations are presented in section one. Section two describes the data used for analysis. Section three is devoted to the different econometric approaches adopted in the paper. Results are displayed in section four. Section five concludes.

2 Theoretical model

To formalize the link between the relative size of the voluntary sector and population heterogeneity, we adopt and extend the model of Alesina et al (1999.)

2.1 The model of Alesina, Baqir and Easterly (1999)

Consider a jurisdiction (a county, for instance) of size 1, in which political decisions about the size and the type of local public good g are taken by majority-vote rule. The individual preferences are written as

$$U_i = g^a(1 - l_i) + c, \quad (1)$$

with

$$\begin{aligned} 0 < a < 1, \\ 0 \leq l_i \leq 1, \end{aligned}$$

where g is the local public good which can be located on an "ideological line" $[0,1]$ of individual preferences concerning different types of public good; l_i is the distance between the individual i 's preferred public good and the delivered one; c denotes the private consumption. Income is considered as equal for everybody, and private consumption is equal to disposable revenues:

$$c = y - t \quad (2)$$

In (2), y is the pre-tax income and t is the lump-sum tax equal for everyone by assumption. As the size of population is normalized at 1, the aggregate and per capita variables are the same, so g per capita equals the total size of g . Then the public budget constraint is

$$g = t, \quad (3)$$

Now individual preferences can be rewritten as follows:

$$U_i = g^a(1 - l_i) + y - g, \quad (4)$$

Here it is supposed this jurisdiction decides by majority rule first on the amount of taxation and thus on the size of the public good, and second on the type of the public good. According to the median voter theorem, for any positive amount of the public good g , the type of public good chosen for the provision will be the most preferred by the median voter. Thus the optimal size of the public good is given by the solution of equation (5).

$$\max U_i = g^a(1 - \tilde{l}_i) + y - g, \quad (5)$$

where \tilde{l}_i is the distance of individual i from the type preferred by the median voter. The equation above incorporates the fact that when the decision on the size of g is taken, the agents know that the type of the publicly provided local public good will be the one most preferred by the median voter. The solution of (5) gives the individual i 's most preferred size of public good:

$$g_i^* = [a(1 - \tilde{l}_i)]^{1/(1-a)}, \quad (6)$$

Now, following Alesina et al (1999) note \tilde{l}^m as the median distance from the type most preferred by the median voter or "the median distance from the median". If the agents know that the type of the public good chosen for the public provision will be the one most preferred by the median voter, then the amount of public good provided in equilibrium will be given by

$$g^* = [a(1 - \tilde{l}^m)]^{1/(1-a)}, \quad (7)$$

polarization of preferences. voter, Equation (7) implies that in equilibrium, the amount of public good of the type preferred by the median voter is decreasing in \tilde{l}^m .

2.2 Heterogeneity effects on individual utility

On the basis of the model above, it is possible to express the individual utility resulting from g^* by substituting g^* in (4).

$$U_i(g^*) = g^* \left[\frac{1 - \tilde{l}_i}{a(1 - \tilde{l}^m)} - 1 \right] + y, \quad (8)$$

The individual utility drawn from the optimal size and type of public good g^* is decreasing with the distance to the median voter \tilde{l}_i . People with \tilde{l}_i smaller than \tilde{l}^m are relatively close to the type preferred by median voter, while people with \tilde{l}_i greater than \tilde{l}^m are relatively distant, hence less satisfied by the public provision.

The median distance \tilde{l}^m can be proxied by some measures of populations' cultural, ethnic, linguistic or religious heterogeneities. Think for instance of preferences for education or art that can be strongly related to cultural or ethnic backgrounds.

According to this model, when an important fraction of population is at a great distance from the median voter, hence the median distance from the median voter preferred type is large, and the type of public good is far from corresponding to the preferences of a large share of population. In this case this population would prefer to contribute less to this public good and to keep taxes low. On the other hand, this category of population would devote more resources to private consumption. The wealthiest people can replace collective-type good by its private substitutes. Moreover, in some jurisdictions with heterogeneous population, public resources may be directed to some specific groups of population via some preferential treatments or "patronage" (Alesina et al, 1999).

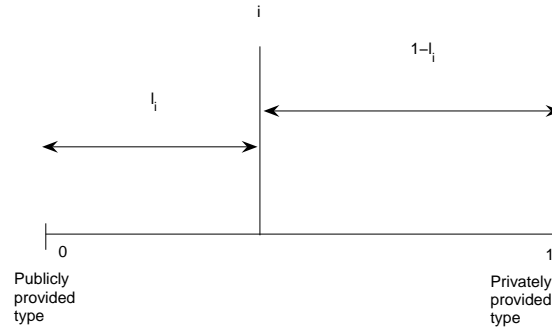
In this paper we assume that people who are unsatisfied by the public provision of collective-type goods may create voluntary organizations or contribute to them in order to provide collective-type goods or services that correspond better to their preferences and needs (Weisbrod, 1986.)

2.3 The case with voluntary contributions

Giving to voluntary organizations targets some individual preferences more perfectly than the governmental provision of public good. Individuals may prefer to redistribute a part of their income towards charities (or nonprofit organizations) responding to the needs or causes they feel particularly concerned about. Therefore, donating is a private act of income redistribution as well as a means of provision of some public goods. Let us consider private donations to voluntary organizations as an example of monetary resources identified with collective-type goods provided via private redistribution. We also call the latter privately provided local public good, which is "local" in the sense that it may correspond more or less well to the preferences of individuals according to its locations in ideological, geographical or other spaces.

Let us locate the chosen public good type on the "ideological line" $[0,1]$ at 0 and the type of privately provided public good at 1, according to the idea of a strong complementarity between these two local public goods. Suppose the population's characteristics are distributed following a normal unimodal law. This can be obtained by the expression of a given individual preference from its proper characteristic through a twice-differentiable function from \mathbb{R} into $[0,1]$ giving 0 for the mean and 1 for the extremes². Figure (1) illustrates a location of i between the different types of publicly and privately provided public goods.

Figure 1: The individual i 's location on the "ideological line".



Following Weisbrod's suggestion, we assume that individuals who are less satisfied with the public provision of public good may decide to create

²Take for instance the function $g(x) = 1 - e^{-x^2}$

or to contribute to a nonprofit. We incorporate this idea into the model by introducing voluntary contributions into the basic expression of individual utility. Incorporating the voluntary contributions into equation (4) yields equation (9):

$$U_i = g^{*a}(1 - l_i) + d^b l_i + y - g^* - d_i(1 - \gamma), \quad (9)$$

$$0 < b < 1,$$

where d_i is the individual i 's contribution to the differentiated public good, b is the parameter of technology of the privately provided local public good and γ is the marginal tax rate. As equation (9) shows, the model describes the context of a jurisdiction, where voluntary donations are deduced from imposable revenues.

$$d = \sum_i d_i, \quad (10)$$

Equation (9) assumes that technologies a and b of publicly and privately provided local public goods are independent and may differ.

If the type and the amount of public good are the most preferred by the median voter, the individual's preferred amount of contribution is given by the solution of:

$$\max U_i = g^{*a}(1 - \tilde{l}_i) + d^b \tilde{l}_i + y - g^* - d_i(1 - \gamma), \quad (11)$$

where g^* is given by (7). The solution of (11) for d_i is

$$d_i^* = \left[\frac{1 - \gamma}{b \tilde{l}_i} \right]^{\frac{1}{b-1}}, \quad (12)$$

Now the amount of giving may be defined by

$$d^* = \left[\frac{1 - \gamma}{b \tilde{l}^m} \right]^{\frac{1}{b-1}}, \quad (13)$$

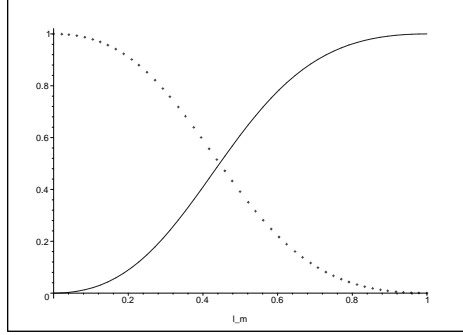
From equation (13) it follows that the equilibrium amount of giving that we associate with the differentiated public good is increasing in \tilde{l}^m , the median distance from the type preferred by the median voter.

The total supply of public good can be written as the sum of publicly and privately provided public goods $G = g + d$, where d is the differentiated public good corresponding to the part of expenditures targeting the preferences of those who donate better, and g is the part of pure public good corresponding to the median voter's most preferred type and not or very imperfectly to the demand of the people whose preferences are located far from the median voter's preferences.

As Figure (2) shows, the increasing heterogeneity in population's characteristics measured by \tilde{l}^m leads to the raise of the share of privately provided

public good in total public good G (noted $R_d = d/G$ and represented by the continuous line), and to the decrease of the portion of its public provision (noted $R_g = g/G$ and represented by the dropped line).

Figure 2: R_g (dotted line) and R_d (continuous line) as functions of \tilde{l}^m , when $a=b=0.5$ and $\gamma = 0.2$



The rest of the paper provides empirical evidences of the relationship between the relative size of privately provided public good and population's heterogeneity.

3 The data

The following proposal from an advertisement well illustrates why voluntary contributions can be considered as a means of supporting cause an individual feels concerned about: *"When choosing a charity, it is important to decide what is most important to you. Most likely, it will be something that has personally affected you. For example, consider your Alma Mater, medical research for a disease that a loved one has endured, the training fund for an athlete you admire, or a local community initiative"*.

In the United States the voluntary sector is represented by two main types of nonprofit organizations. The first one embodies the organizations registered in section 501(c)3 of the Internal Revenue Code (this category includes cultural organizations, foundations, nonprofit schools and universities, day-care centers, charities, and community improvement organizations among others.) These organizations are tax exempt, and they can receive tax-deductible charitable donations from individuals, but cannot engage in any political or commercial activities. The second type is represented by the organizations falling under the smaller 501(c)4 category such as some social welfare organizations, organizations performing lobbying activities on behalf of specific issues, social clubs and others. Generally, contributions to this type of nonprofit are not tax-deductible, except for some volunteer fire companies and similar organizations, as well as some war veteran's organizations. One cannot deduct his or her donations to labor unions, political candidates, social clubs, homeowner associations, most of the foreign

charities, and organizations performing social activities for the pleasure or recreation of members.

In the U.S. only itemized donations reduce taxable income.³ The tax saving from the act of giving usually equals the amount of deduction (which is normally the total contributed except in some cases) times the marginal tax rate, namely the top rate for the person's income level. Giving to federal, state, and local government is also tax-deductible if the contribution is for public purposes. However, voluntary organizations are more likely to attract private monetary and labor donations than governments (Rose-Ackerman, 1996.) According to the Internal Revenue Service (IRS), in 1998 the direct contributions represented more than 50 percent of total contributions, gifts and grants received by 501(c)3 organizations. Americans gave 160 billion dollars to charities in 1998, and donations rose up to 212 billion in 2001. According to IRS, in 1999 about 35.5 million taxpayers made deductible charitable contributions totaling nearly 125.8 billion dollars. Contributions from individuals account for more than 75 percent and about 82 percent of total giving is itemized for federal income tax returns.⁴ The data concerning the giving are drawn from the National Center for Charitable Statistics. They indicate the amount of total itemized households' contributions by U.S. county. Counties seem to represent a relevant level of analysis as they deal with provision of a large fraction of local public goods. They support the provision of local public goods in the fields of public welfare, health and hospitals, environment and housing (parks, recreation, community development) among others. The provision of local public goods is financed by local taxes (e.g. property tax, 23,3 percent of general revenue), intergovernmental transfers (35.5 percent in 1996-97, coming mostly from states), and charges. According to the U.S. Census Bureau,⁵ in 1996-97, the U.S. counties' expenditures on public welfare, educational services, hospitals, health, and parks and recreation equaled respectively 13.5 percent, 13.2 percent, 9.8 percent, 7.3 percent, and 2 percent of total expenditures.

For the analysis provided below, it is important to mention that U.S. counties kept their borders constant over years. The dependent variable which is supposed to indicate the share of impure public good in total public good is constituted as follows: $Y = R_d = \frac{d}{d+g}$, where d is the amount of total households' contributions in each county reported in 1998, g is the amount of public expenditures in each county in 1996-97 which comes from the U.S. Census Bureau, 1997 Census of Governments.

As we are testing the local effects of population heterogeneity on the relative size of the voluntary sector, our main explanatory variable we are interested in is the index measuring population heterogeneity. We employ the broadly-

³The donations are deduced from the federal personal income tax.

⁴In this study we do not take into account the non-itemized charitable contributions.

⁵Source:<http://www.census.gov/govs/estimate/97countysummary.html>

used measure of heterogeneity named "fractionalization" which indicates the probability for two randomly selected individuals in a county to be affiliated to different ethnic groups (Alesina et al (1999), Alesina et al (2003), Alesina et al, (2002.)) Alesina and al, 1999 provide a detailed explanation of ethnic heterogeneity as a proxy of preferences heterogeneity of population. It is generally calculated as follows:

$$FRACT_j = 1 - \sum (Ethnic_{ij})^2, \quad (12)$$

where $Ethnic_{ij}$ denotes the share of population in a county j self-identified as of ethnic group i . Following the classification used by the U.S. Census i = White, Black, American Indian or Alaska, Asian, Hawaiian or Pacific, Other, two or more races. As noted Alesina et al, 1999, the category Other represents essentially Hispanic people. The index of population ethnic heterogeneity was calculated from the U.S. Census 1998 data. A fractionalization index equal to 0 for a county means that all households affiliate themselves to the same ethnic group, while a fractionalization index value of 1 represents the maximum heterogeneity. When the index of fractionalization tends to 0, that means that there is a dominant ethnic group. When it tends to 1, several more or less equal groups are present in the county. In this study, the index of ethnic heterogeneity lies in the range [0.01 0.687]. To assess linguistic heterogeneity, we use the share of population which is more than 5 years old and speaks another language at home. We measure income inequality at the county level (variable "incineg") as a ratio of the mean household income to the median household income in a county (Alesina and al, 1999, p. 1259). In order to account for other differences in preferences, we introduce the variables of income per capita (variable "income"), the share of people living below the level of poverty (variable "poverty"), the percentage of population of 65 years and older (variable "age"), as well as the percentage of population older than 25 with a bachelors degree or higher (variable "education").

We know that a rise in wealth tends to increase the demand of public goods. However, with still growing wealth, the demand of certain public goods decreases and leads to the substitution of public goods by their private analogues (e.g. municipal versus private swimming pools). In other words, the demand for such collective goods is a peaked function of income.

As it was stressed in the previous literature, the life cycle hypothesis predicts that persons over 65 years tend to spend a larger part of their current income on immediate consumption than younger people, and to demand a larger quantity of public goods than younger persons with the same income and tax share (Bergstrom and Goodman, 1973.) The level of education can affect the public provision of public goods as more educated counties can choose better public policies, demand more education for children, or monitor provision of public goods (Alesina and al, 1999.) The size and type

of population may affect the share of giving in the total public expenditures in different manners. For example, large urban jurisdictions are often considered as suffering from inefficiencies in the provision of public goods and as more unwieldy than small ones. Therefore, in some jurisdictions people may contribute relatively more to nonprofits in the amount of local expenditures. Moreover, in larger urban jurisdictions the amount of giving may be positively affected by the greater number of organizations asking for voluntary contributions. On the other hand, as jurisdictions with small population can exert more social control, one can speculate that voluntary income redistribution could be higher in smaller jurisdictions. We assess the type of population through the log of urban population in each county. The size of population is assessed through the log of population of each county reported in the US Census 1997. The set of the explanatory variables used in the model is written as :

$X = [\iota, \text{frac}, \text{incineg}, \text{otherlanguage}, \log \text{ population}, \log \text{ urban population}, \text{poverty}, \text{education}, \text{age}, \text{income per capita}]$, with ι denoting a vector of constant term ⁶.

The location of the 3111 counties is determined using the 1990 census information on latitude and longitude. 48033, We have decided to exclude from the analysis, observations related to the District of Columbia (zip code 11001) because of its special status. Alaska and Hawaii have also been excluded from the analysis as they do not share a common border with other counties. Associated with the initial 23 missing values in the dataset, we get a sample of 3083 observations.

4 The spatial econometric approach

4.1 Global models with fixed spatial weight matrix

4.1.1 Assuming homoscedastic errors

When observations are geographical units as countries, spatial dependence and heterogeneity may arise among observations. Spatial dependence is defined as the lack of independence between observations over space (Anselin, 1988). For instance, positive spatial dependence occurs when similar locations in space exhibit similar values which leads to apparent clusters in space. To accommodate spatial autocorrelation in the disturbances, the spatial error model (SEM) has been developed. For a dataset with n geographical

⁶Collinearity had not been detected in the data according to the BKW (1980) diagnostics.

units, it is written as

$$\begin{aligned} y &= X\beta + \epsilon \\ \epsilon &= \lambda W\epsilon + u \\ u &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{14}$$

where y is the $(n \times 1)$ vector of dependent variable, X is the $(n \times (k + 1))$ matrix of explanatory variables including a constant term, W is the $(n \times n)$ matrix of spatial weights used to define the structure of the correlations between the disturbances and λ , the parameter of interest denotes the strength of this spatial dependence.

The weighting scheme used to design W is usually based on contiguity between counties. In each row of W , A positive value is assigned to counties that are close 'enough' to county i to be considered as its neighbors and a null value otherwise, with $i = 1, \dots, n$. The diagonal of this matrix contains zero element as a county cannot be used to predict itself.

The SEM log-likelihood function is given in (15).

$$\begin{aligned} \ln L &= -(n/2)\ln(2\pi) - (n/2)\ln\sigma^2 + \ln |I - \lambda W| \\ &\quad - (1/2\sigma^2)(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta). \end{aligned} \tag{15}$$

The first order conditions give the generalized least squares estimate for β and σ^2 , conditional on λ :

$$\hat{\beta}_{ML} = [(X - \lambda WX)'(X - \lambda WX)]^{-1}(X - \lambda WX)'(y - \lambda Wy). \tag{16}$$

$$\hat{\sigma}_{ML}^2 = (e - \lambda We)'(e - \lambda We)/n, \tag{17}$$

with $e = y - X\beta_{ML}$. Introducing (16) and (17) in (15) leads to the concentrated log-likelihood function for the SEM model.

$$\ln L_C = C - (n/2)\ln[(e'(I - \lambda W)'(I - \lambda W)e)/n] + \ln |I - \lambda W|. \tag{18}$$

Pre-multiplying (14) by $(I - \lambda W)$ leads to (19).

$$y = \lambda Wy + X\beta - \lambda WX\beta + u. \tag{19}$$

This model may be written as what is known as the spatial Durbin model (see Anselin L. and Bera A.K.(1998)) given in (20) if the restriction $-\lambda\beta = \gamma$ is imposed in (19).

$$y = \lambda Wy + X\beta + W\tilde{X}\gamma + u \tag{20}$$

where \tilde{X} denotes the matrix of independent variables without the constant term (i.e. each row of X is defined as $x_i = [1, \tilde{x}_i]$). In the spatial Durbin

model (SDM), the dependent variable is explained firstly by a set of independent variables, secondly by the average of the dependent variable (Wy) in neighboring counties and thirdly by the average of the independent variables in neighboring counties (WX). The concentrated log-likelihood function for the SDM model is formulated in (21) with $e_0 = y - X\beta - W\tilde{X}\gamma$ and $e_1 = Wy - X\beta - W\tilde{X}\gamma$.

$$\ln(L) = C - \ln |I - \lambda W| - (N/2) \ln(e_0'e_0 - 2\lambda e_1'e_0 + \lambda^2 e_1'e_1). \quad (21)$$

The Burridge common factor allows us to test this restriction. If this constraint cannot be rejected, a model with spatial dependence in the disturbances is estimated, otherwise a SDM is estimated.

However as United States counties do not have the same size, heteroscedasticity in the error term may occur. Moreover as noted by LeSage (1997) "enclave effects"⁷ may lead to fat tailed errors or t distributed errors. Thus LeSage proposed to estimate global spatial models with the Bayesian method in order to correct this failure.

4.1.2 Assuming heteroscedastic errors

LeSage J.P (1997) proposed to introduce an additional fixed but unknown parameter noted r in the global spatial model in order to accommodate outliers and observations with large variances. For instance, the heteroscedastic spatial Durbin model may be written as:

$$\begin{aligned} y &= \lambda Wy + X\beta + W\tilde{X}\gamma + \epsilon \\ \epsilon &\sim N(0, \sigma^2 V) \\ V &= \text{diag}(v_1, v_2, \dots, v_n) \end{aligned} \quad (22)$$

A chi-squared prior distribution for the terms in V is suggested by LeSage $\pi(r/v_i) = IID\chi^2(r)$ as it reduces to introducing a single parameter noted r ; a normal-gamma conjugate prior for β and γ , and σ^2 when the sample size is large which may be formulated as follows for $\theta = [\beta \ \gamma]$: $\pi(\theta) \sim N(c, T)$, with c and T taking large values, and $\pi(1/\sigma^2) \sim \Gamma(d, v)$. Finally a uniform prior for λ is set as $\pi(\lambda) = U[0, 1]$.

4.2 Global models with flexible spatial weight matrix

Note that not only the estimation and inference results presented in the preceding section are conditional on the specification of the spatial weight matrix, but also this specification takes place before estimating the model

⁷When the relationship under study is different in a particular set of counties compared to the entire territory under study, this is known as "enclave effects".

whose aim is to detect and estimate the spatial interactions designed in W. LeSage and Pace (2000a) proposed a Matrix Exponential Spatial Structure model in which a flexible form for the spatial weight matrix is allowed and is an integral part of the spatial model estimation. They also developed a version of the MESS model (2000b) robust to heteroscedastic disturbances and outliers in which the spatial weight matrix will be a linear combination of individual nearest neighbor matrices with declining coefficients as we move from the first nearest neighbors to the second and so on. This flexible weight matrix takes the form:

$$W = \sum_{i=1}^q \left(\frac{\rho^i N_i}{\sum_{i=1}^q \rho_i} \right) \quad (23)$$

where N_i represents the spatial matrix associated with the i^{th} nearest neighbors in space with $i = 1, \dots, q$. q is the maximum number of nearest neighbors to take into account, these can be first order neighbors (i.e. direct neighbors), second order neighbors (neighbors of neighbors) or higher order neighbors of observations. ρ is the decay parameter which is specific to each individual matrix N .

The Bayesian MESS model given in (25) relies on the spatial weight matrix exponential transformation of the dependent variable specified in (26)

$$Sy = Z\theta + \epsilon, \quad (24)$$

$$\epsilon \sim N(0, \sigma^2 I_n) \quad (25)$$

with S the $(n \times n)$ positive, definite matrix:

$$S = e^{\alpha W} \quad (26)$$

with W defined in (23) and $Z = [XW\tilde{X} \dots W^{l-1}\tilde{X}]$, the matrix of explanatory variables with independent variables noted \tilde{X} (excluding a constant term) and lagged independent variables up to order $(l-1)$. LeSage and Pace (2000) recommended to set $l = 7$ to ensure the convergence in the Taylor series expansion. Note that the 'traditional' spatial dependence parameter, noted before λ can be recovered as $\lambda = 1 - e^\alpha$.

Robust estimates to non constant variance in the error terms can be obtained through a Bayesian heteroscedastic model with the same prior for θ, σ^2, v_i as in the Bayesian heteroscedastic spatial Durbin model. Non informative priors on α, ρ and q i.e. the parameters entering the spatial weight matrix are assigned using uniform distributions. More particularly, $\pi(\alpha) = U[-\infty, 0]$, $\pi(\rho) = U(0, 1)$ and $\pi(q) = U^D[1, q_{max}]$.

4.3 Local model: SALE

When we deal with spatial dataset, the hypothesis of spatial stationarity defined as the stability over space of the relationship under study (Anselin,

1988) is often not met. In this case, local models like the nonparametric locally weighted regression (LWR) model developed by McMillen (1996) are more appropriate than global ones as they allow for parameter heterogeneity. However the widely used LWR method gives results that are conditional on a single value of the bandwidth appearing in the weighting function. Changing this value produces different estimated values for the parameters and consequently different conclusions regarding the stationarity of the relationship under study. Moreover the LWR model has been constructed in such a way that the smaller is the sub-sample size, the less frequently the spatial dependence is likely to occur. Thus the LWR parameter estimates can be used to test the variability over space of the relationship under study. However, a locally weighted regression on a smaller sub-sample size leads to more volatile parameter estimates which may question the significance of the detected regional patterns in the parameter estimates. Furthermore, as the absence of spatial dependence in the LWR residuals cannot be guaranteed in small sub-sample size, estimating a local model with a spatially lagged dependent variable as additional explanatory variable and then testing the significance of the spatial dependence parameter seems appropriate. Thus to deal simultaneously with spatial dependence and parameter heterogeneity, Pace and LeSage (2002) propose a Spatial Autoregressive Local Estimation (SALE). It takes the following form:

$$M(i)y = M(i)Wy\lambda_i + M(i)X\beta_i + M(i)W\tilde{X}\gamma_i + M(i)\epsilon. \quad (27)$$

The vectors y and ϵ , the matrices \tilde{X} and W are the same as those previously defined with (14). M is a $(n \times n)$ diagonal matrix in which a value of one is assigned to the m -nearest neighbors to observation i and a zero value for the others. M allows to extract a sub-sample on which the regressions are conducted. $M(i)Wy$ is the spatially lagged dependent variable and its associated parameter, noted λ_i , measures the extent to which the dependent variable in i , noted $M(i)y$ may be explained by the average of the ones of its nearest neighbors in space, $M(i)Wy$. Similarly, γ_i denotes the effect of the average of the independent variables in the neighborhood of i on y_i . As pointed out by Pace and LeSage (2002), the inclusion of the spatially lagged dependent variable in the model may firstly solve the spatial dependence problem arising from large sub-sample size and secondly may decrease the sensitivity of the parameter estimates to the bandwidth resulting in more stable estimates.

Contrary to the LWR model, the SALE model may be estimated for any sub-sample size, simply by modifying the matrix M . This provides a way to investigate the occurrence of spatial dependence as the number of neighbors is no longer constrained. In other words, this method enable us to see how the spatial dependence parameter changes with the sub-sample size from small to full sample i.e. from SALE (respectively LSDM) estimates (reflecting parameter heterogeneity) to SAR (respectively SDM). estimates

(reflecting parameter homogeneity).

Even if both methods seem similar, they are methodologically different. In the LWR framework, the weighting scheme places a uniformity condition on the parameter of spatially neighboring observations. This is not the case in the SALE method. Analyzing the beta convergence parameter estimates for each observation of the sample, LeSage and LeGallo (2003) investigate the empirical supporting existence of the local convergence concept in Europe.

The SALE estimates are the ones which maximize the following profile likelihood function (see Anselin, 1988, Pace and Barry, 1997):

$$\ln L(\lambda) = C + \ln |I - \lambda W| - (n/2) \ln(SSE(\lambda)), \quad (28)$$

where C denotes a constant and $SSE(\lambda)$, the sum of squared errors.

5 The results

5.1 Global aspatial and spatial models

A linear regression model is estimated with ordinary least squares. Table 1 gives the estimation and inferences results of the standard linear regression model. All the coefficients are significant on at least 5 %. The OLS adjusted measure of fit is 0.3107. Values in brackets are the t-probability coming from White (1980) variance matrix robust to heteroscedasticity. The positive and significant coefficient of fractionalization is consistent with the theoretical model. Parameter associated with income inequality has a weak negative effect, as well as the share of people in a county speaking other language. The log of urban population has the strongest influence on the relative size of voluntary sector. As expected, age has a negative but small significant effect, while the size of county assessed through the log of population demonstrates a small positive effect. Finally, income per capita has the smallest positive global effect.

The OLS errors are non normally distributed according to the Jarque-Bera test. They are also heteroscedastic as the null hypothesis of the Koenker (1981) test robust to normal errors can be rejected at the significant level of 5%. Moreover, two tests against unspecified alternative for the spatial dependence process have been carried out. The spatial dependence in the disturbances is tested using a contiguity matrix found by Delaunay triangulation. The Moran I-statistic, adapted to the OLS residuals by Cliff and Ord (1981) indicates spatial dependence in the OLS residuals. But, according to Fingleton (1999), this test statistic may detect not only spatial dependence but also spatial non stationarity. It is also not reliable when heteroscedasticity occurs. Thus in order to evaluate the significance of both effects, tests against specified alternatives have been realized. They have been noted

LMLAG and LMERR. The decision rule elaborated by Anselin and Florax (1995) enables us to choose a spatial error specification as the RLMERR statistic is significant whereas the RLMLAG is not ⁸. Bidirectional tests have also been done. Clearly, spatial dependence and heteroscedasticity are the joint source of misspecification in the model. Thus a model with spatial dependence in the error terms has been estimated by maximum likelihood. The estimated results of the SEM model are given in the second column of Table 1. The SEM model produces a positive and significant spatial autocorrelation parameter. The introduction of the latter rises the adjusted measure of fit by 50 percent. According to the LMLAG* test introducing a spatially lagged dependent variable as additional explanatory variable to the SEM model is not necessary. However, the Burridge common factor test indicates that the spatial error model cannot be rewritten as a spatial Durbin model (SDM) and thus the latter has to be estimated. The results of the maximum likelihood estimation of the spatial Durbin model are in the second column of Table 2. The coefficients of variables "fractionalization" and "age" become insignificant, while the coefficient of variable "other language" becomes positive. The size and type of population keep their strength, significance, and positive signs. The coefficient of the variable "poverty" remains significant and negative. The coefficient related to education is still positive and significant, consistent with stylized facts.

The greatest coefficient is associated with the measure of urbanization. This result is consistent with the idea according to which the social interactions among people contribute to increase the relative size of the voluntary sector. This could also mean that urban population is more likely to donate because of different factors (e.g. people are better informed, the greater number of voluntary organizations asking for donations). In fact, it is generally recognized that in urban areas the public expenditures are often more important than in smaller and rural jurisdictions. However, our result suggests that in urban areas the increase in public expenditures g is less than the increase in d .

It is interesting to note that the coefficient of the spatially lagged variable "fractionalization" is significant and positive. In other words, when the degree of population's ethnic diversity increases in county's direct neighborhood, the relative size of the nonprofit sector in all counties increases to an extent proportional to the geographical proximity. In contrast, the impact of the language heterogeneity and income inequality of neighboring counties are negative and significant.

As heteroscedasticity may occur (e.g. as counties are highly heterogeneous

⁸The decision rule states that if the p-value of the LMLAG statistic (respectively LMERROR) is less than the one for the LMERROR statistic (respectively LMLAG) and that RLMLAG (respectively RLERROR) is statistically significant but not LMERROR (respectively RLMLAG), then the appropriate model is the spatial autoregressive model (SAR) (respectively the spatial error model (SEM)).

in size), we present in Table 2 the results of the Bayesian estimation of a SDM model with an heteroscedastic prior (denoted $r=4$) for the variance of the error terms. The Bayesian estimation of heteroscedastic model leads to small changes compared to the homoscedastic model: the parameter of "other language" becomes insignificant, while the one associated with "age" becomes significant and negative. One striking feature is that the measures of ethnic, language and income heterogeneity seem to influence the relative size of the nonprofit sector only through spatial interactions with neighboring units. The Bayesian framework allows us to compare the specifications of the SDM model. More particularly, the posterior probability of each of the three models may be calculated and models may be discriminated through the calculus of posterior odds ratio: the SDM model estimated by maximum likelihood, a SDM model with a homoscedastic prior and a SDM model with an heteroscedastic prior. As the posterior probability equals one for the last mentioned model, only the results of the Bayesian heteroscedastic spatial Durbin model are included in the Table. Moreover, the Bayesian homoscedastic model estimation gives the same results as the maximum likelihood estimation as a large dataset is used in this paper.

Finally a MESS model and a Bayesian heteroscedastic MESS model have been estimated in order to evaluate the robustness of the SDM results to a more flexible form for the spatial weight matrix. The corresponding results are in Table 2. Let us recall that the MESS model does not presume that the design of the spatial interactions between counties is given *a priori*. Moreover, inferences may be drawn about the parameters entering the spatial weight matrix given in (23). The (homoscedastic) MESS model better fits the data than the (homoscedastic) SDM model as the adjusted measure of fit raises from 0.3532 to 0.4443 when a matrix exponential transformation of the dependent variable is used. A positive and significant spatial dependence parameter is estimated in the MESS model, $\lambda = 1 - e^\alpha = 0.5527$. However, the extent of this detected spatial spillover is larger than in the SDM model. In fact, the decay parameter equal to 0.887 implies what LeSage called a "half-life" of seven neighbors meaning that the 7th nearest neighbor exerts less than 1/2 the influence of the first nearest neighbors. Furthermore, the posterior mean for the neighbor parameter is 21.58 with a standard deviation of 7.69. Thus the magnitude of spatial influence seems to be quite small. When non constant variance in the error terms is allowed, the strength and scope of the spatial influence are larger than when the heteroscedastic nature of the errors is ignored. The spatial dependence parameter equals 0.77 and a 'half-life' of eleven neighbors is estimated. Figure 3 plots the posterior means of the individual variance. The pattern of higher variance over observations 2323 to 2364 and 2861 to 2897 supports the heteroscedastic

MESS model.⁹ The observations mentioned in figure 3 may be considered as outliers.

5.2 Local spatial model

The estimated parameters of the parameterized form of the spatial weight matrix introduced in the heteroscedastic BMESS model have been used in the SALE model to define the spatial interdependence between the observations. Moreover, using a sub-sample of 456 observations leads to the lowest cross-validation error at each observation, as depicted in figure 4, thus regression will be run under this sub-sample size.

According to figure 3(a), a strongest interdependency in nonprofit sector size in western counties (with highest values in Arizona and New Mexico) and some eastern counties (states Michigan, Ohio, Kentucky, Tennessee) is detected. The higher interdependency in the west may be explained by the fact that these counties present a higher than average urbanization and face a low density that confront them with higher costs of provision of certain local public goods (transportation, water supply and other services). Thus, they are more likely to cooperate at local and state levels.

The strongest variations in the spatial dependence parameter are localized in the central counties. This finding is linked to the existence of some unobserved variables such as common cultural features, history, or area specificities¹⁰. Figures 3(b) to 3(g) display the local parameters for the population's ethnic, linguistic, and income heterogeneities and for these spatially lagged variables. The local parameters associated with fractionalization present an increasing trend from negative in the west to positive in the east, as illustrated in figure 3(b). In other words, population heterogeneity has no identical effect across observations. Figure 3(c) shows the local estimates related to linguistic heterogeneity. The parameter estimates present a rather homogenous feature across observations with highest but still weak values in the middle-east and in the north-west. Negative values are concentrated in the center of the country. The local parameters associated with income inequality are indicated in figure 3(d) . Positive coefficients are mainly localized in Florida, Georgia and partly in Texas.

Concerning the local spatial interactions between counties, the positive coefficients of the lagged variable of fractionalization exhibit a single-centered pattern localized in Colorado, Kansas and Nebraska as depicted in figure 3(e). The coefficients decrease with distance from this center until they take negative values (in Wisconsin and North Dakota, for instance). The esti-

⁹The heteroscedastic models present weaker values of measure of fit, than the homoscedastic ones, as their primary aim is not to adjust well the model to the data, but to be robust for heteroscedasticity.

¹⁰Note that the most important spatial interdependencies correspond to the western area, which is also the one composed of the largest counties

mated parameters for linguistic heterogeneity in the neighborhood of each county (displayed in figure 3(f)) are only positive in the north part of the center of the country (namely, Minnesota, North and South Dakota, Iowa), while they are negative elsewhere. Finally, figure 3(g) demonstrates the positive coefficients associated with income inequality in neighboring counties are clustered partly in Texas, Pennsylvania and Wisconsin, while the negative parameters are concentrated in the extreme North-East. The aspatial global estimation methods usually employed in the empirical studies mask these spatial variations. This fact may lead to misleading conclusions.

6 Conclusion

In this paper the effects of population heterogeneity on the relative size of the nonprofit sector in the United States using county level data are investigated. Assuming that the relation under study is stationary over space, the spatial interdependence between U.S. counties are taken into account via a spatial (global) Durbin model. We show that in a global spatial setting, the relative size of the nonprofit sector in each county depends positively among others on the county's linguistic heterogeneity and on the relative size of the nonprofit sector in the neighboring counties. According to our results, the relative size of the voluntary sector is more important in urban, more educated and richer counties. However, these results are only global in the sense that they ignore local spillovers that may occur. In a local setting, population heterogeneity and more particularly ethnic diversity and linguistic heterogeneity may affect the relative size of the nonprofit sector, but to a different extent and in a different manner according to the location. Future research may consist in developing a theoretical background to model the spillovers detected in this study. Besides, a Bayesian spatial local estimation robust to heteroscedastic errors should be conducted in order to make valid inferences on the estimates.

Appendix

Table 1: OLS, SEM

| estimations | OLS | SEM | tests | OLS | SEM |
|------------------|----------------------|-----------------------|--------|--------------------------|--------------------------|
| constant | 0.0676 (0.0088) | −0.0629 (0.0152) | JB | $3.372e^{+5}$ (0.000) | - |
| frac | 0.0643 (0.0000) | 0.0294 (0.0017) | BP | $2.629e^{+3}$ (0.000) | $2.639e^{+3}$ (0.273) |
| incineg | −0.0177 (0.0001) | −0.0117 (0.0012) | KB | 109.57 (0.000) | () |
| otherlang | −0.0010 (0.0000) | −0.0004 (0.0078) | Moran | 26.32 (0.000) | - |
| logpop | 0.0069 (0.0000) | 0.0082 (0.0000) | LMERR | 679.03 (0.000) | - |
| logurban | 0.7702 (0.0047) | 0.9061 (0.0014) | RLMERR | 151.37 (0.000) | - |
| poverty | −0.0016 (0.0031) | −0.0013 (0.0000) | LMLAG | 527.69 (0.000) | - |
| education | −0.0004 (0.0436) | 0.00067 (0.0012) | RLMLAG | 0.0308 (0.8607) | - |
| age | −0.0017 (0.0000) | −0.00076 (0.01155) | LM1 | 3308 (0.000) | - |
| income | 0.000004 (0.0043) | 0.000003 (0.0000) | LM2 | 3308 (0.000) | - |
| $\hat{\lambda}$ | - | 0.581 (0.0000) | LMLAG* | — | $4.356e^{-10}$ (1) |
| $\overline{R^2}$ | 0.3107 | 0.4628 | CF | — | $-8.25026e^{+7}$ (0) |
| $\hat{\sigma}^2$ | 0.0026 | 0.002 | - | - | - |

Notes: $\overline{R^2}$ is the adjusted measure of fit. JB is the Jarque-Bera normality test. BP denotes the Breusch-Pagan heteroscedasticity test and KP a robust to non normal error heteroscedasticity test. Moran is the Moran I-statistic. LMERROR and its robust version named RLMERROR are Lagrange Multiplier (LM) tests for spatially autocorrelated errors. LMLAG and its robust version RLMLAG are tests for omitted spatially lagged dependent variable. LM1 is a joint test for spatial dependence and heterogeneity (Anselin, 1988, p71). It is the sum of BP and LMBI (and LMBI=RLMLAG+LMERROR or alternatively LMBI=LMLAG+RLMERR). LM2 is the joint test for heteroscedasticity and spatial autocorrelation in the errors (Anselin, 1988, p121). It is the sum of BP and LMERR. CF is the Burrige (1981) common factor test.

Table 2: *SDM (ML), BSDM(r=4), BMESS(r=200), BMESS(r=4)*

| estimations | SDM | BSDM | BMESS(r=200) | BMESS(r=4) |
|------------------|-----------------------|-----------------------|----------------------|-----------------------|
| constant | 0.0964 (0.0007) | 0.0643 (0.0004) | 0.125 (0.0000) | 0.0071 (0.3695) |
| frac | -0.005 (0.6413) | 0.0025 (0.6993) | -0.005 (0.334) | -0.00587 (0.1975) |
| incineg | -0.00796 (0.028) | -0.00323 (0.0977) | -0.0087 (0.0125) | -0.0018 (0.1885) |
| otherlang | 0.0006 (0.003) | 0.00018 (0.165) | 0.0007 (0.0000) | 0.0001 (0.1555) |
| logpop | 0.0079 (0.0000) | 0.0062 (0.0000) | 0.0086 (0.0000) | 0.0072 (0.0000) |
| logurban | 0.83 (0.0034) | 0.522 (0.0009) | 0.9142 (0.0005) | 0.5888 (0.0000) |
| poverty | -0.0012 (0.0005) | -0.00059 (0.0029) | -0.0013 (0.0005) | -0.0004 (0.0375) |
| education | 0.00174 (0.0000) | 0.0004 (0.0076) | 0.0017 (0.0000) | 0.0004 (0.006) |
| age | -0.00014 (0.649) | -0.0010 (0.0000) | 0.00008 (0.391) | -0.001 (0.0000) |
| income | 0.000003 (0.0000) | 0.000006 (0.0000) | 0.00003 (0.0000) | 0.000007 (0.0000) |
| Wfrac | 0.04167 (0.0008) | 0.0212 (0.0074) | 0.0395 (0.002) | 0.0073 (0.21) |
| Wincineg | -0.0074 (0.307) | -0.0095 (0.022) | 0.00174 (0.4365) | -0.005 (0.2245) |
| Wotherlang | -0.0012 (0.0000) | -0.0008 (0.0000) | -0.00127 (0.0000) | -0.0004 (0.011) |
| Wlogpop | -0.00488 (0.0061) | -0.0027 (0.0175) | -0.0069 (0.002) | -0.0054 (0.0015) |
| Wlogurban | -0.962 (0.1117) | -0.5752 (0.1155) | -1.1517 (0.0975) | -0.8897 (0.0925) |
| Wpoverty | 0.0002 (0.6766) | 0.00047 (0.11036) | -0.00009 (0.439) | 0.001 (0.0045) |
| Weducation | -0.0022 (0.0000) | -0.0009 (0.0000) | -0.0023 (0.0000) | -0.0006 (0.004) |
| Wage | -0.0013 (0.0072) | -0.0002 (0.493) | -0.0019 (0.0000) | 0.0008 (0.0315) |
| Wincome | -0.000001 (0.0000) | -0.000004 (0.0000) | -0.000001 (0.042) | -0.000005 (0.0000) |
| alpha | - | - | -0.8047 (0.00000) | -1.472 (0.0000) |
| rho | - | - | 0.8878 (0.00000) | 0.9349 (0.0000) |
| q | - | - | 21.59 (0.00000) | 29.87 (0.0000) |
| λ | 0.528 (0.0000) | 0.46 (0.0000) | 0.5527 (0.0000) | 0.7705 (0.0000) |
| R^2 | 0.3570 | 0.2994 | 0.446 | 0.4 |
| $\hat{\sigma}^2$ | 0.0020 | 0.0027 | 0.0021 | 0.0005 |

Figure 3: Distribution of the means for 2000 v_i estimates

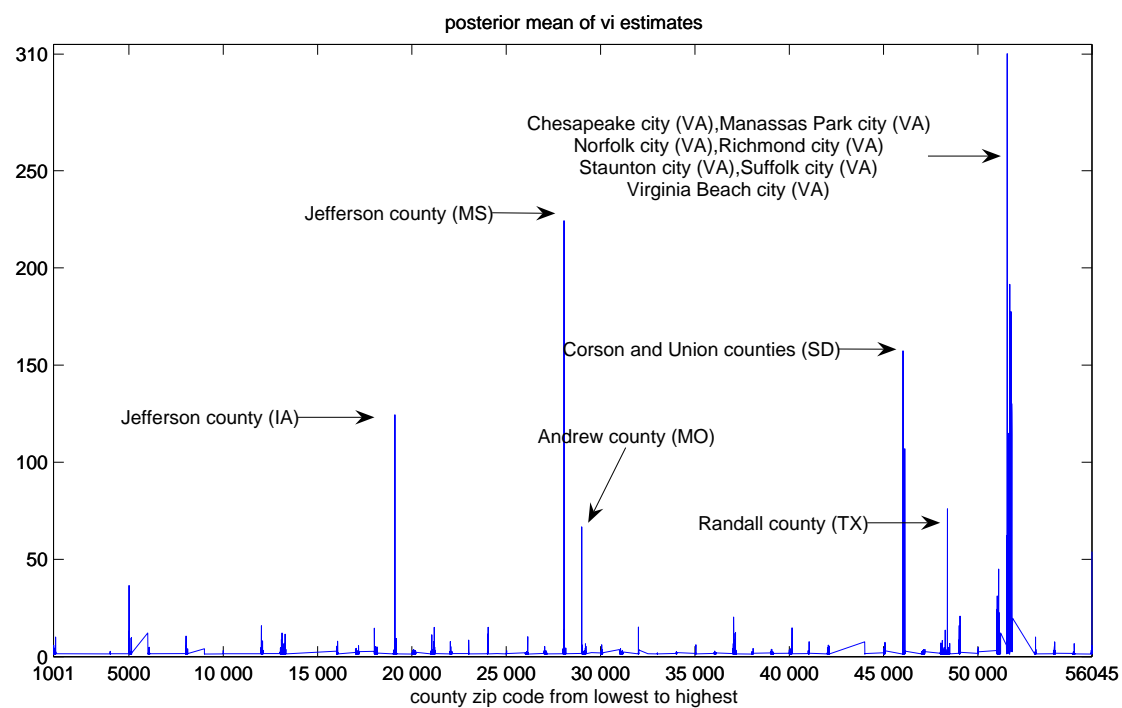


Figure 4: Prediction error on center of area for various sum-sample size

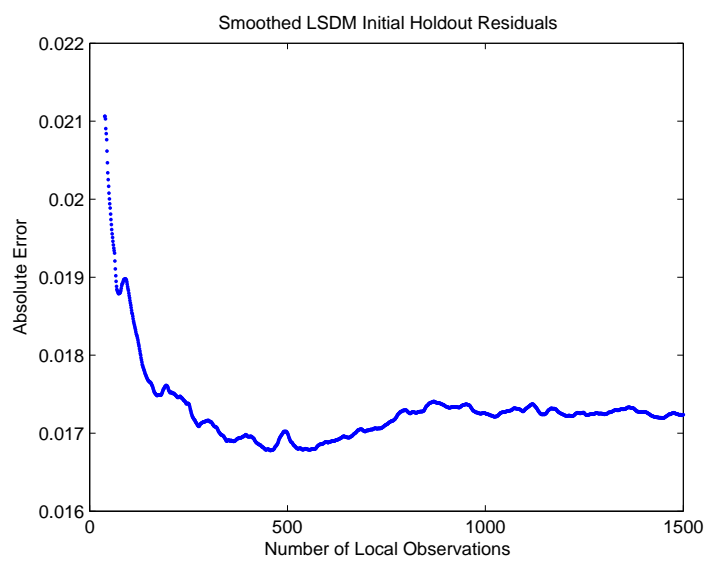
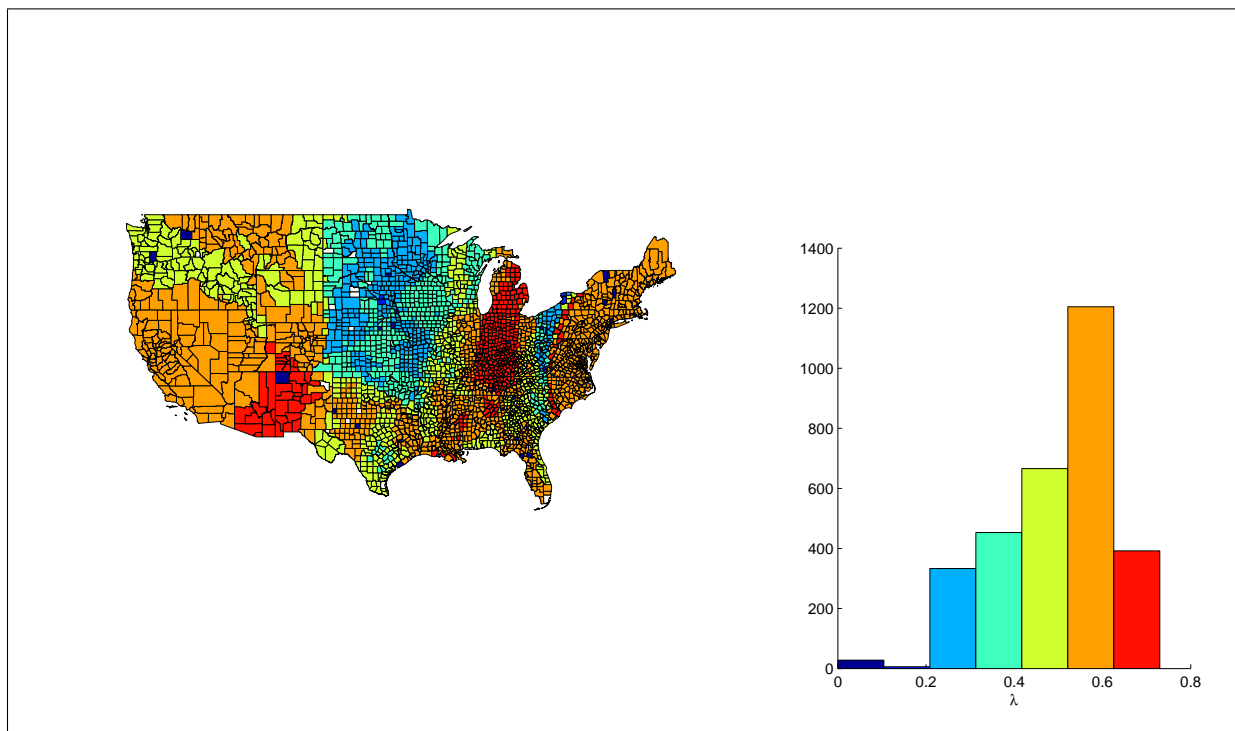
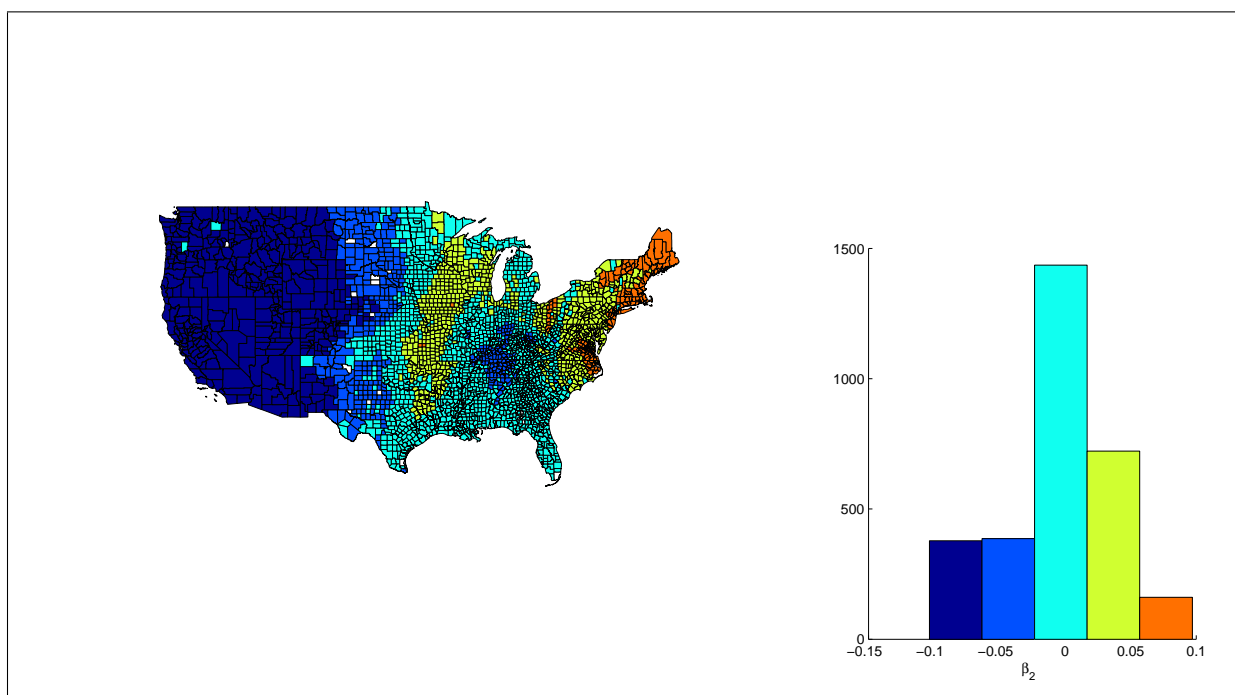


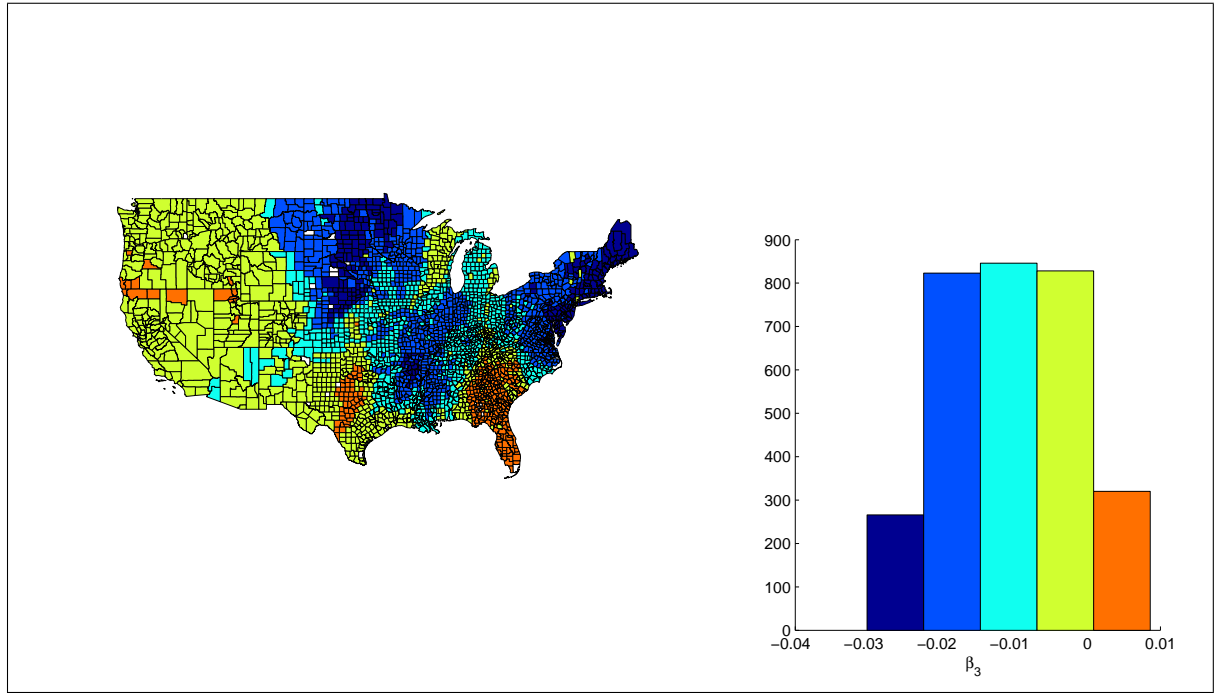
Figure 5: Local spatial Durbin model parameter estimates based on $m=456$



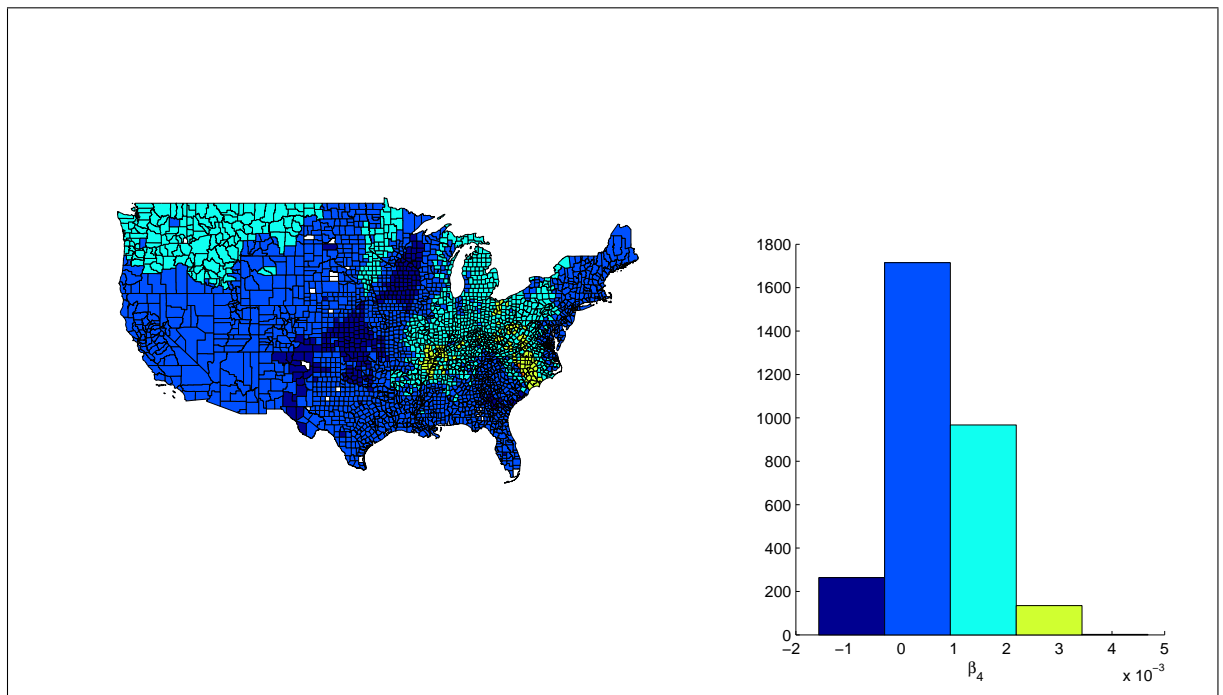
(a) County's Spatial dependence



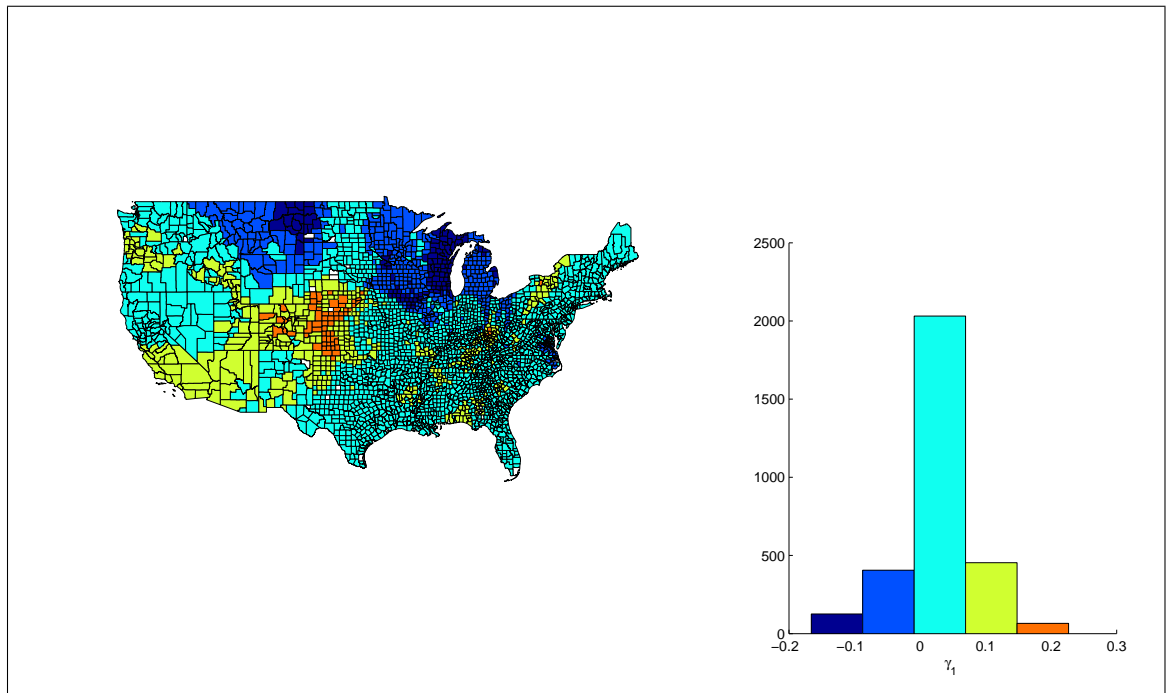
(b) County's ethnic heterogeneity



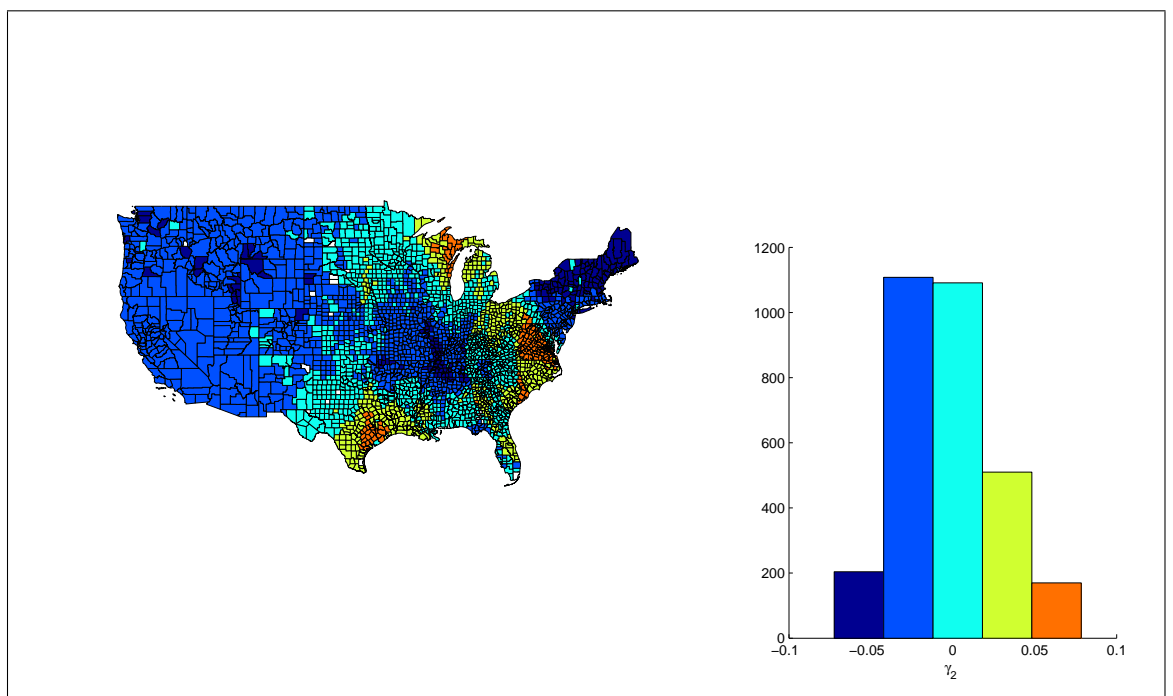
(c) County's income inequality



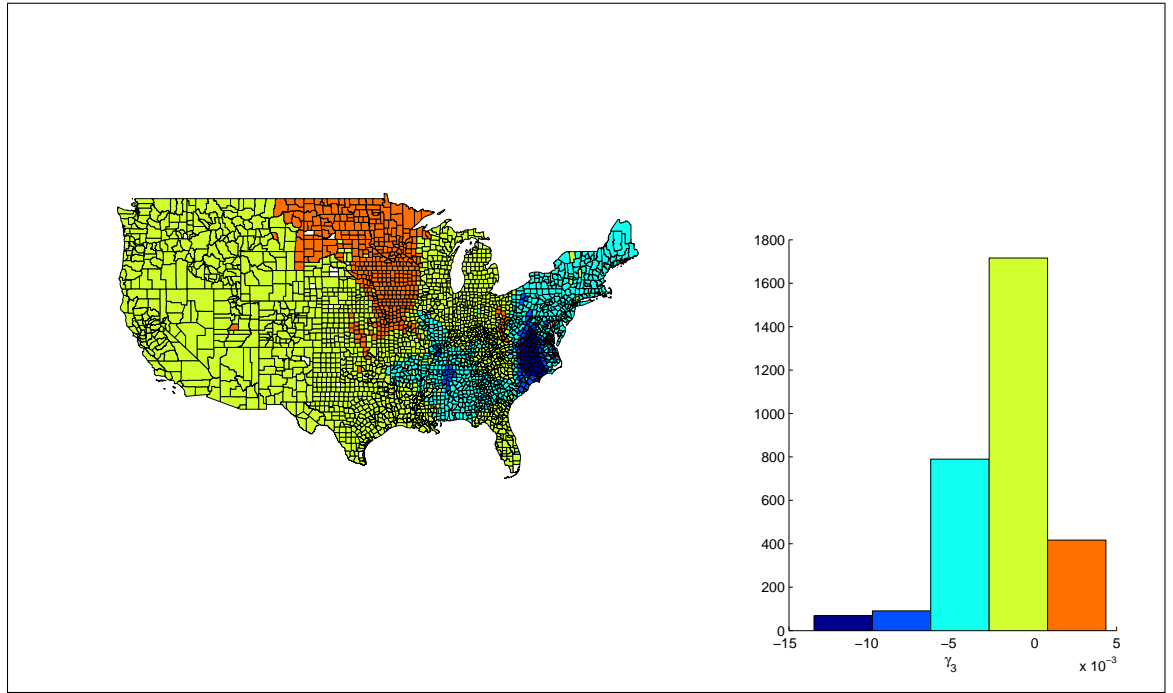
(d) County's linguistic heterogeneity



(e) County's neighbors ethnic heterogeneity



(f) County's neighbors income inequality



(g) County's neighbors linguistic heterogeneity

References

- [1] Alesina A., Baqir R., Easterly W. (1999) Public Goods and Ethnic Divisions *The Quarterly Journal of Economics*, 114 (4): 1243-1284.
- [2] Alesina A., Baquir R., Hoxby C. (2002) Political Jurisdictions in Heterogenous Communities, Feb., (unpublished manuscript).
- [3] Alesina A., La Ferrara E. (2000) Participation in Heterogeneous Communities *The Quarterly Journal of Economics*, 115 (3): 847-904.
- [4] Alesina A., Devleeschauwer A., Easterly W., Kurlat S., Wacziarg R. (2003) Fractionalization *Journal of Economic Growth*, 8: 155-194.
- [5] Anselin L. (1988a) *Spatial Econometrics: Methods and Models*
- [6] Anselin L., Bera A. (1998) Spatial dependence in linear regression models with an application to spatial econometrics, in A. Ullah et D.E.A. Giles (eds), *Handbook of Applied Economics Statistics*, Springer-Verlag, Berlin, 21-74.
- [7] Bergstrom T.C., Goodman R.P. (1973) Private Demands for Public Goods *The American Economic Review*, 63 (3): 280-296.

- [8] Brock W and Durlauf S. Growth Empirics and Reality, *World Bank Economic Review*, 15, 3: 229–272, 2001.
- [9] DiMaggio P.J., Anheier H.K. (1990) The Sociology of Nonprofit Organizations and Sectors *Annual Review of Sociology*, 16 :137-159.
- [10] Durlauf S. N. (2002) On The Empirics of Social Capital, *The Economic Journal*, 112: 459–479.
- [11] Durlauf S.N., Fafchamps M. (2003) Empirical Studies of Social Capital: A Critical Survey, www.ssc.wisc.edu/econ/archive/wp2003-12.pdf.2003.
- [12] Ertur C., Le Gallo J., LeSage J.P. (2003) Local versus Global Convergence in Europe: A Bayesian Spatial Econometric Approach. REAL Discussion paper, REAL 03-T-28.
- [13] Ertur C., Koch W. (2005) Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence.
- [14] James E. (1993) Why Do Different Countries Choose a Different Public-Private Mix of Educational Services? *The Journal of Human Resources*, 28 (3): 571-592.
- [15] Lassen D.D. (2006) Ethnic Divisions, Trust, and the Size of the Informal Sector *Journal of Economic Behaviour & Organization* Vol. xxx(2006)xxx-xxx (article in press).
- [16] Lesage J.P (1997) Bayesian Estimation of Spatial Autoregressive Models *International Regional Science Review*, 20: 113-129.
- [17] McMillen D.P. (1996) One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics* 40: 100-124.
- [18] Miguel E., Gugerty M.K. (2004) Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. Unpublished manuscript available at [http : //emlab.berkeley.edu/users/emiguel/miguel_tribes.pdf](http://emlab.berkeley.edu/users/emiguel/miguel_tribes.pdf)
- [19] Okten C., Osili U.O. (2005) Ethnic Diversity and Charitable Giving. Unpublished manuscript available at [http : //team.univ – paris1.fr/espe2005/papers/osili_paper.pdf](http://team.univ-paris1.fr/espe2005/papers/osili_paper.pdf)
- [20] Pace R.Kelley, LeSage J.P (2000a) Closed Form Maximum Likelihood Estimates for Spatial Problems. Unpublished manuscript available at [http : //www.spatial – statistics.com](http://www.spatial-statistics.com)
- [21] Pace R.Kelley, LeSage J.P (2000b) Using Matrix Exponentials to Explore Spatial Structure in Regression relationship.

- [22] Pace R.K., Barry R. (1997) Quick Computation of Regression s with a Spatially Autoregressive Dependent Variable. *Geographical Analysis*, 39(3):232-247.
- [23] Pace R.K, LeSage J.P (2004) Spatial autoregressive local estimation. *Recent Advances in Spatial Econometrics* Getis Arthur, Mur Jesus and Henri Zoller (eds), Palgrave Publishers, 3:31-51.
- [24] Rose-Ackerman S. (1996) Altruism, Nonprofits, and Economic Theory *Journal of Economic Leterature*, XXXIV: 701-728.
- [25] Salamon L.M., Sokolowski S.W., Anheier H.K. (2000) Social Origins of Civil Society: an Overview *Working Papers of The Johns Hopkins Comparative Nonprofit Sector Project*, Johns Hopkins University.
- [26] Weisbrod B. (1977) The voluntary nonprofit sector: an economic analysis. Lexington Books D.C. Health & Co Lexington, Massachusetts Toronto.
- [27] Weisbrod B. (1986) Toward a Theory of the Voluntary Nonprofit Sector in a Three-sector Economy in *The Economics of Nonprofit Institutions*, ed. S.Rose-Ackerman, Columbia University.