IEREK PRESS
Academic Research Community
CrossMark
http://www.press.ierek.com
ESSD
ISSN (Print: 2357-0849, online: 2357-0857)
International Journal on:
Environmental Science and Sustainable Development

DOI: 10.21625/essd.v4i2.556

In Search of a Tool to Support Planning Inside Large Cities: the SustaIn-LED Model

Giuseppe Umberto Cantafio[1], Petros Ieromonachou[2]

[1]*Business, University of Sunderland in London, 197 Marsh Wall Docklands, London, E14 9SG, London, United Kingdom*

[2]*Business School, The University of Greenwich*

Abstract

The aim of the present study is to investigate the linkages between local economic development, innovation, and environmental sustainability inside urban areas.

Can innovation affect the improvement of the quality of life inside urban areas? This research question comes from the consideration that usually innovation and growth in general are considered sources of conflict in affecting the livability of large cities.

The objective of the paper is to design a model — the "SustaIn-Led" - to connect levels of environmental sustainability, quality of life, and economic development inside metropolitan areas, taking into account also innovation processes, activated by the innovation policies and by the knowledge economy.

The study takes in consideration the 53 largest United States metropolitan areas with a population over 1 million, with a time series from the years 2000 through 2015.

This has been done because of a two-fold reason: (1) the US among high-income countries is the one with the highest number of universities, patents, and citations; (2) several studies have shown that innovation occurs in large cities.

The first part of the present study has carried out the identification of the variables to represent and significantly explain the phenomena – local economic development, innovation, and environmental sustainability – linked to the design of the SustaIn-LED model.

Environmental sustainability in urban areas in this paper is represented by means of the Air Quality Index (AQI), while the number of workers synthetically quantifies local economic development. Correlation and multiple regression analyses are conducted in order to examine the relationship between the three main indicators.

The multiple regressions for the year 2015 produced a low p-value, indicating that the predictors are significant in the regression analysis. Similar results of p-value are shown in all the years from 2000 to 2013. For 2015, the results showed that part of the variance in the measure of total workers of the metropolitan areas could be predicted by measures of innovation and air quality. Higher R^2 values have been registered for the years from 2000 through 2013.

The development of the SustaIn-LED model could be utilized in urban regeneration processes to help in the design of new urban planning policies inside large cities by means of a better comprehension of environmental and economic implications caused by the implementation of innovation policies.

Keywords
SustaIn-LED Model; Sustainability; Sustainable Development; City Planning

1. Introduction

Over half of the global population lives in cities. The trend toward urbanization all over the world is a key challenge that suggests the need for developing models regarding urban organization and sustainable development. While economic development is essential, other factors control the urban areas' overall sustainability, competitiveness, and growth, including its innovation ecosystem, the quality of the environment, and the quality of life. These factors not only represent an advantage for urban areas in order to achieve economic growth, but also generate a stable economic and social environment.

The aim of the present study is to find out whether or not a relationship exists within Local Economic Development and innovation, driven by Sustainability, within urban areas. This will be carried out by means of the objective represented by the SustaIn-LED model, developed in order to capture and connect standards of sustainability, innovation, and economic development.

The success of a city does occur because of several factors. Many of them are localization factors, depending on different reasons. The term "success" can be translated with the use of different criteria that will be utilized in the present study, for measuring the sustainability, innovation, and economic development within urban areas. Today, people are attracted to urban areas that offer them not only decent jobs, but also a high standard of living.

Different studies have highlighted the importance of sustainability and sustainable planning within the cities, not only in the US, but also in most developed countries. As an example, a study made in Portland, OR (USA) analysed the role of the local government during the creation of green cluster and city development strategies for the transition to a Green Urban Economy (Zimmerman, 2013). Other studies showed models developed for the application of sustainability to cities (Newman, 1999) or regarding growth and innovation inside cities (Bettencourt et al., 2007) or others more recently investigated the role of knowledge in urban areas for achieving sustainable and more equitable growth (Pastor & Benner, 2015).

This study takes into account the United States metropolitan areas because among high-income countries, the USA is the one with the highest level of innovation, with the highest number of universities, patents and citations. This selection has been done because several studies have shown that innovation occurs in large cities (Sedgley & Elmslie, 2011; Lee & Rodriguez-Pose, 2014). Other studies provided insightful examinations of how innovation occurs in cities. Some other studies endorsed the notion that cities are places of innovation and creativeness (Florida, 2010; Bettencourt & West, 2010; Shearmur, 2012; Glaeser, 2014). This is mainly due to three kinds of inputs: human capital, financial resources, and knowledge sharing. For instance, research labs pay experts and engineers to generate valuable new ideas. A recent study points out that cities generate also more unconventional innovation (Gaetani & Berkes, 2015).

This research in the future could help in the decision-making process for the implementation of innovation policies in urban areas, trying to draw possible elements that can be transferred and less positive elements that should be analyzed more in depth before being transferred to other regions. Moreover, it could be used as an effective tool for evaluating how urban areas can achieve higher standards of sustainability, innovation and economic development. A better planning of cities could result in different outcomes, such as job creation. In turn, this could result, under a Keynesian point of view, in an increase of consumption and tax revenues for the area, which could raise the quality of the services, the GDP of the region and other parameters for achieving a better quality of life.

2. Methodology

Extensive and repeated visits over the last years have formed a knowledge base about many aspects of urban areas in the USA and Europe. Research during my doctorate as a PhD generated quantitative and qualitative data for this study. As an exploratory study the results of the research were not expected at the beginning of the investigation.

2.1. The scope of study

There was a time where cities were delimited by their political boundaries. In recent decades urban settlements have been rapidly growing into larger urban areas or metropolitan areas, with the city center as the nucleus of these urban entities. For instance, New York City in 2013 had an estimated population of circa 8M of people, whereas 20M lived in the New York- Newark-Jersey City metropolitan area. The broad definition of metropolitan area is that of a core area comprising a large population nucleus, together with adjacent communities that have a high degree of economic and social integration with that core. More specifically, in the United States a metropolitan statistical area consists of one or more counties that contain at least one urbanized area that has a population of at least 50,000 or more inhabitants (2010 census summary file). The present study has considered USA metropolitan areas with a population larger than 1 million. The reasons for this choice are explained in the introduction section; this is supported also by the reason that in terms of innovation larger cities have more advantages than smaller cities, not only because of the larger number of people, but also because of more networks, as new ideas come from the combination of existing ideas. Furthermore, in larger cities there is more information spillover; the connections are more varied in larger cities, as there is more possibility to meet diverse people that eventually make something innovative happening more likely, in comparison with smaller cities (Johnson, 2011).
The US metropolitan areas analyzed in the present study are listed in table 1. The metropolitan areas are listed in alphabetical order.

2.2. The database and indicators

For the present study it has been collected a multifaceted database, looking across various aspects of sustainability, technological innovation, and economic development, to get the full picture. The Sustainability-Innovation-Local Economic Development model (SustaIn-Led model) assesses three parameters: air quality, jobs, and technological innovation. Regression analyses within each parameter are carried out in the study. In the following table are clarified the indicators relative to the three main topics included in the development of the SustaIn-LED model.

Table 2. Indicators chosen for the SustaIn-LED model

Topic	**Indicator**	**Units**	**Source**	**Geographic scale**	**Description of Indicator Inputs**
Environmental Sustainability	Air Quality	Median Air Quality Index	Environmental Protection Agency	Metropolitan Areas with population greater than 1 Million	The AQI is based on the five ”criteria” pollutants regulated under the Clean Air Act: ground-level ozone, particulate matter, carbon monoxide, sulfur dioxide, and nitrogen dioxide.
Innovation	Utility patents	Utility Patents per 100 Workers	US Patent and Trademark Office Patent Technology Monitoring Team (PTMT)		Utility Patents Originated During 2015 (patent origin is determined by the residence of the first-named inventor)

Continued on next page

Table 2 continued

Local Economic Development	Employment	Employment rate over Population	US Bureau of Labor Statistics and American Community 2015 Survey 1-Year Estimates		Employment / Population rate (16 years and over)

The reasons for the choice of the three indicators shown in table 2, representing the three main topics of the present study, are now explained:

- To represent local economic development, several indicators have been analyzed: total employment, number of jobs of local clusters, wage, specialization (as national employment share), and active population (16 years old). As an outcome of this phase, total employment has been selected as the most appropriate indicator for the research. It is a better fit than employment of local clusters because it comprises all the types of jobs within economy, including the ones related to exporting industry. The total employment helps to explain more and it makes more sense for this investigation. It also represents broader facts affecting the community.
- Regarding the choice of the parameter representing innovation, there are two separate schools of thought for analyzing innovation at a broad scale. Many academics in their studies have utilized patent data (Feldman & Lendl, 2010; Sedgley & Elmslie, 2011). Another body of researchers believes that the best measure for innovation is represented by surveys. However, a study conducted by (Acs, Anselin & Varga, 2002) determined that patents deliver a reasonably precise evaluation of the geography of product innovation throughout US metropolitan areas (MSAs), but patents do not control their legitimacy in smaller town or non-urban areas, nor for other types of economic innovation, like for instance non-scientific fields or more humanistic areas. The latter does not represent an issue, since the present study has been carried out with data of large metropolitan areas, with a population of over 1 million.
- For environmental sustainability, the median value of the AQI of the metropolitan area has been utilized. The reason that lay behind is that high amounts of air pollution are connected with frequent adverse health effects for example increased risk for heart disease and lung illnesses. The air quality is a measure of the tendency for the adverse health consequences related with air pollution. In this study air quality is quantified utilizing data from the Environmental Protection Agency (EPA) through Air Quality Index (AQI). The data provided tells how polluted the air is. An AQI measure of 100 usually corresponds to the national air quality standard for the pollutant, which is the level that EPA has fixed to defend public health.

The model used for multiple linear regression, given the observations, is:

$$Y_1 = \alpha \mid \beta_1 X_{1,1} \mid \beta_2 X_{2,2} \tag{1}$$

It is useful to recall which variables are taken into account for the study at this level:

- Dependent variable: Number of Workers in the American metropolitan areas; Y_1 in the equation (1)
- Independent variable no.1: Number of patents in the American metropolitan areas; $X_{2,1}$ in the equation (1)
- Independent variable no.2: Air Quality Index in the American metropolitan areas; $X_{2,2}$ in the equation (1)

Table 1. US metropolitan areas studied in the SustaIn-LED listed by population

1	New York-Newark-Jersey City, NY-NJ-PA	2	Los Angeles-Long Beach-Anaheim, CA
3	Chicago-Naperville-Elgin, IL-IN-WI	4	Dallas-Fort Worth-Arlington, TX
5	Houston-The Woodlands-Sugar Land, TX	6	Washington-Arlington-Alexandria, DC-VA-MD-WV
7	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	8	Miami-Fort Lauderdale-West Palm Beach, FL
9	Atlanta-Sandy Springs-Roswell, GA	10	Boston-Cambridge-Newton, MA-NH
11	San Francisco–Oakland–Hayward, CA	12	Phoenix-Mesa-Scottsdale, AZ
13	Riverside-San Bernardino-Ontario, CA	14	Detroit-Warren-Dearborn, MI
15	Seattle-Tacoma-Bellevue, WA	16	Minneapolis-St. Paul-Bloomington, MN-WI
17	San Diego-Carlsbad, CA	18	Tampa-St. Petersburg-Clearwater, FL
19	Denver-Aurora-Lakewood, CO	20	St. Louis, MO-IL
21	Baltimore-Columbia-Towson, MD	22	Charlotte-Concord-Gastonia, NC-SC
23	Portland-Vancouver-Hillsboro, OR-WA	24	Orlando-Kissimmee-Sanford, FL
25	San Antonio-New Braunfels, TX	26	Pittsburgh, PA
27	Sacramento–Roseville–Arden-Arcade, CA	28	Cincinnati, OH-KY-IN
29	Las Vegas-Henderson-Paradise, NV	30	Kansas City, MO-KS
31	Cleveland-Elyria, OH	32	Columbus, OH
33	Austin-Round Rock, TX	34	Indianapolis-Carmel-Anderson, IN
35	San Jose-Sunnyvale-Santa Clara, CA	36	Nashville-Davidson–Murfreesboro–Franklin, TN
37	Virginia Beach-Norfolk-Newport News, VA-NC	38	Providence-Warwick, RI-MA
39	Milwaukee-Waukesha-West Allis, WI	40	Jacksonville, FL
41	Oklahoma City, OK	42	Memphis, TN-MS-AR
43	Louisville/Jefferson County, KY-IN	44	Raleigh, NC
45	Richmond, VA	46	New Orleans-Metairie, LA
47	Hartford-West Hartford-East Hartford, CT	48	Salt Lake City, UT
49	Birmingham-Hoover, AL	50	Buffalo-Cheektowaga-Niagara Falls, NY
51	Rochester, NY	52	Grand Rapids-Wyoming, MI
53	Tucson, AZ		

3. Results and discussion

This section shows the results of the linear regression that has been carried out to investigate the linkages between economic development, economic clusters and environmental sustainability inside urban areas. The results have been drawn with the support of the software Microsoft Excel and Stata. Data from 2013 have been used for running the first linear regression. The rationale behind it is that Local Patents and Air Quality Index (independent variables) do have an effect on the total number of workers in the metropolitan area (dependent variable).

Table 3. Summary statistics

Variable	Observations	Obs. with missing data	Obs. without missing data	Min.	Max.	Mean	Std. deviation
Y	53	0	53	305393	7870219	1322172.92	1309411.907
AQI	53	0	53	36	97	50.547	10.5
Tot. Patents	53	0	53	92	12899	1886.491	2504.574

Correlation matrix:	AQI	Tot. Patents	Y
AQI	1	0.146	0.291
Tot. Patents	0.146	1	0.569
Y	0.291	0.569	1

Regression of variable Y:	
Goodness of fit statistics (Y):	
Observations	53.000
Sum of weights	53.000
DF	50.000
R^2	0.368
Adjusted R^2	0.342
MSE	1127512333648.42
RMSE	1061843.837
MAPE	49.542
DW	1.837
Cp	3.000
AIC	1473.717
SBC	1479.627
PC	0.708

Analysis of variance (Y):					
Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	2	32781479498312.6	16390739749156.3	14.537	< 0.0001
Error	50	56375616682421.1	1127512333648.42		
Corrected Total	52	89157096180733.7			

As a double-check, the data have been analyzed also using another software, Stata 13.0. The robustness check after an OLS regression is performed to see if none of the Ordinary Least Squares (OLS) assumptions are violated. The OLS is a method for estimating the unknown parameters in a linear regression model, with the goal of minimizing the sum of the squares of the differences between the observed responses in the given dataset and those predicted by a linear function of a set of explanatory variables.

For the year 2013, the process has shown results as an R-squared for the multivariate correlation a value of 0.368, meaning that almost 40% of the dependent variable can be explained by the independent variables (number of patents and environmental pollution). Other results for the R-squared value show a value of 0.569 between the variables Y and Total Patents, meaning that almost the 57% of the number of jobs (taken as a measure of economic development) in the largest US metro areas can be explained by the number of patents (assumed as a measure of technological innovation).

The F-ratio tests output, analyzing whether the overall regression model is a good fit for the data, shows that the independent variables statistically and significantly predict the dependent variable, F (2, 50) = 14.537 and p $0.0001 < 0.0005$. This means that the regression model is a good fit of the data. A small p (≤ 0.05), rejects the null hypothesis. This is strong evidence that the null hypothesis is invalid.

The Analysis Of Variance single factor (ANOVA) test, for the year 2015 has been performed between the three variables Y, Total Patents and AQI, namely employment, total patents and Air Quality Index in the metro areas. As shown in Table 4, the analysis gives an F-value greater than F-critical, being $55.61 > 3.05$. Similar results have been obtained for the other years of the time series: in 2013, 2011, 2008, 2005, 2002, and 2000 the F-value obtained was always greater than F-critical in the ANOVA test, therefore the null hypothesis can be rejected.

Table 4. ANOVA test with three variables for the year 2015

ANOVA: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Total Workers	53	86068566	1623935	2.50995E+12		
AQI	53	2708	51.09434	105.202		
Tot. Patents	53	105903	1998.17	7527260.528		
ANOVA						
Source of Variation		df	MS	F	P-value	F crit
Between Groups	9.31E+13	2	4.65E+13	55.615	5.84022E-19	3.054
Within Groups	1.31E+14	156	8.37E+11			
Total	2.24E+14	158				

Table 5. ANOVA test with two variables Y and AQI for the year 2013

ANOVA: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Y	53	86068566	1623935.208	2.50995E+12		
AQI	53	2708	51.094	105.202		
ANOVA						
Source of Variation		df	MS	F	P-value	F crit
Between Groups	6.98805E+13	1	6.98805E+13	55.682	2.69221E-11	3.932
Within Groups	1.30517E+14	104	1.25497E+12			
Total	2.00398E+14	105				

The ANOVA test for the year 2015, between the variables Y and AQI, namely employment and Air pollution, gives an F-value greater than F-critical, being 55.68>3.93 (as shown in the table above).

An F-statistic value greater than the critical value is equivalent to a p-value less than alpha and both mean that the null hypothesis can be rejected.
Similarly, the ANOVA test for 2015 performed for the variables Y and Tot. Patents gives F-value>F-critical, with 55.54>3.93. For the year 2013, the ANOVA tests performed between Y and AQI present F-value>F-critical, being 54.03>5.93 and between Y and Total Patents an F-value>F-critical, with 53.88>3.93. Therefore, also for the year 2013, the null hypothesis can be rejected. Given these analyses, the results of the regression from the data can be considered robust.

3.1. Interpretation of the results

The F-ratio tests whether the overall regression model is a good fit for the data. The output shows that the independent variables statistically and significantly predict the dependent variable F (2, 50) = 24.56 and p 0.0000< 0.0005 (i.e., the regression model is a good fit of the data). The results show an R-squared value of 0.495, meaning that the independent variables (AQI and I, namely Air Quality and Innovation) explain 49.5% of the variability of our dependent variable Y (Total Employment of the Metropolitan area).

The same model has been applied for the years: 2000, 2002, 2005, 2008, 2011, 2013, and 2015. For each year all the indicators listed in Table 1 have been collected. With the data collected before 2007/2008, we have the opportunity to have information before and after the economic crisis of 2007/2008. In the following chart the R-squared values of the years 2000, 2002, 2005, 2008, 2011, 2013 and 2015 are shown.

From the following table can be seen that the maximum R-squared value is in 2000, while the minimum is reached in 2015, with a progressive decrease of this value. This means that the correlation among the selected variables has diminished over the years. The variance in the measure of total workers in metro areas predicted by measures of innovation and air quality has shrunk from more than 60% in 2000 to about 36% in 2015, as shown in the following table 6.

Table 6. R-squared results SustaIn-LED model for the years 2000:2015

Year	R-squared
2000	0.611
2002	0.473
2005	0.442
2008	0.420
2011	0.400
2013	0.368
2015	0.359

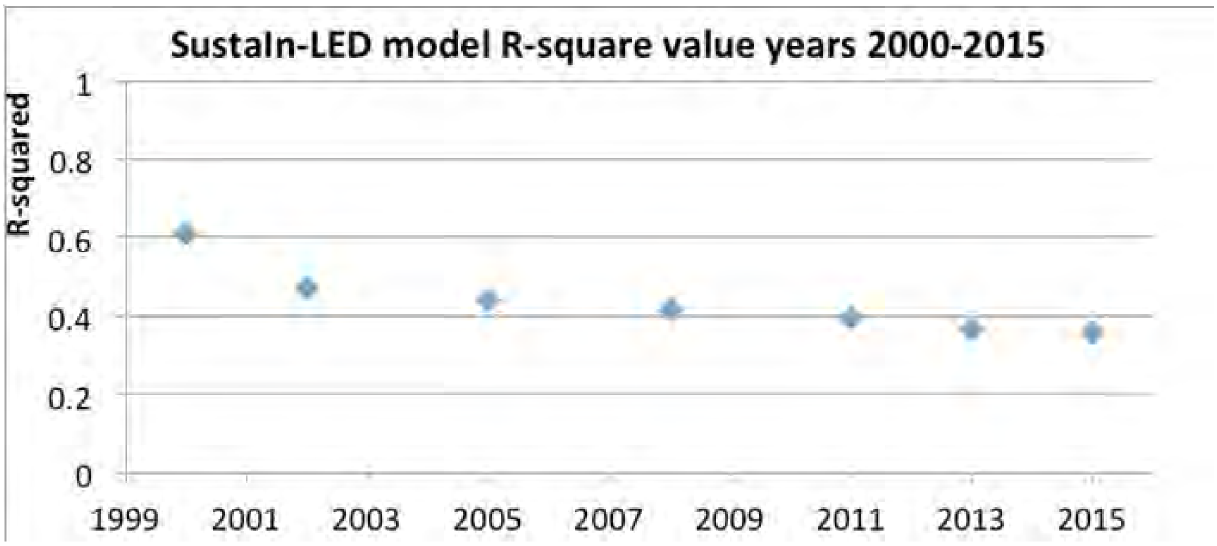

Figure 1. R-squared coefficient from the linear regressions of the SustaIn-LED model in the years 2000-2015

It is probable that the negative trend phenomenon can be explained with the various economic crises that happened during the years taken into account in the study.

In the 2000s an energy crisis swept the western world, known as the 2000s energy crisis - Since 2003, a rise in prices was produced by continuous global rises in petroleum demand combined with production stagnation, the falling value of the US dollar, and a multitude of other minor causes. Another crisis was in 2000–2001, the California electricity crisis, caused by market manipulation by Enron and failed deregulation; this resulted in multiple large-scale power outages. Last but not least, another reason that can explain the results is the financial crisis of 2007–09, also identified as the global financial crisis and the 2008-09 financial crisis, considered by numerous economists to be the worst financial crisis that has happened since the Great Depression of the 1930s.

All these crises had an impact on various sectors of the economy, starting with the financial market, with a drop of various stock market indexes, e.g. in June 2008 the Dow Jones Industrial Average (DJIA) had fallen 20% from its October 11, 2007 high. Also, the Standard and Poor's stock market index (S&P) 500 lost approximately 50% of its value in the period 2007-2009. Other effects of these crises have been seen on the global economy as a whole and on the US domestic economy, with a drop of the Gross Domestic Product with an annual rate of the 6% until 2009, and impacts on the distribution of the wealth. 77% of the richest families had a decrease in total wealth.

It is difficult to say which is the cause and which is the consequence in these cases, but we can say that there is a relation among the economic and energy crises of the 2000s. A reason for the job losses from 2008 on lays behind the fact that during the economic crisis period, as the Nobel Prize in Economic Sciences Paul Krugman wrote, "countries faced with economic crisis were urged by Washington to raise interest rates, cut public spending and increase taxes". As a result to the increase in interest rates and taxes, firms stopped investing and hiring workers, "as in past periods of depression, the decline in the rate of profit reduced business investment, which in turn resulted in slower growth and higher rates of unemployment" (Moseley, 2009).

As a result of the cut in public spending, public spending in initiatives for social well-being, such as pollution remediation and sustainability measures decreased after the period of economic crisis. The drop of the R-squared value, representing the variance in the measure of total workers in metro areas predicted by measures of innovation and air quality, as shown in the previous figure, can be explained by the economics and social facts aforementioned.

In 2015, the unadjusted multiple R = 0.359, but that the adjusted multiple R is 0.334. This rather small change is due to the fact that a relatively large number of observations are being predicted with a relatively small number of variables.

3.2. Linkages between the dependant variable and one independent variable

After having analyzed the multivariate function, it is interesting to separately explore the linkages between the dependent variable and one of the two independent variables. The relation between the number of workers and the number of patents has been analyzed, as well as the relation between the number of workers and the air quality. The results of these analyses are interesting for determining the changes over the year of the correlation between innovation and local economic development, and the correlation between environmental sustainability and economic development.

The series of data going from 2000 to 2015 shows a negative trend of the R-squared coefficient, namely a diminishing linkage among the dependent variable (number of workers in the metro areas) and the independent variable (number of patents registered in the metro areas). The R-squared value decreased from a maximum value of 0.542 in 2000 to a minimum of 0.254 in 2015. The negative trend of the coefficient of determination could mean that the model was more robust in the first years taken into account for the model, while in the meantime, some other variables could have affected the dependent variables (number of workers). Among these variables there could be the 9/11 factor, happened in 2001, and the financial crisis of 2007-2008. These events have shaken the market and could have affected the relationship hypothesized in the model, among innovation and economic development in the American metro areas.

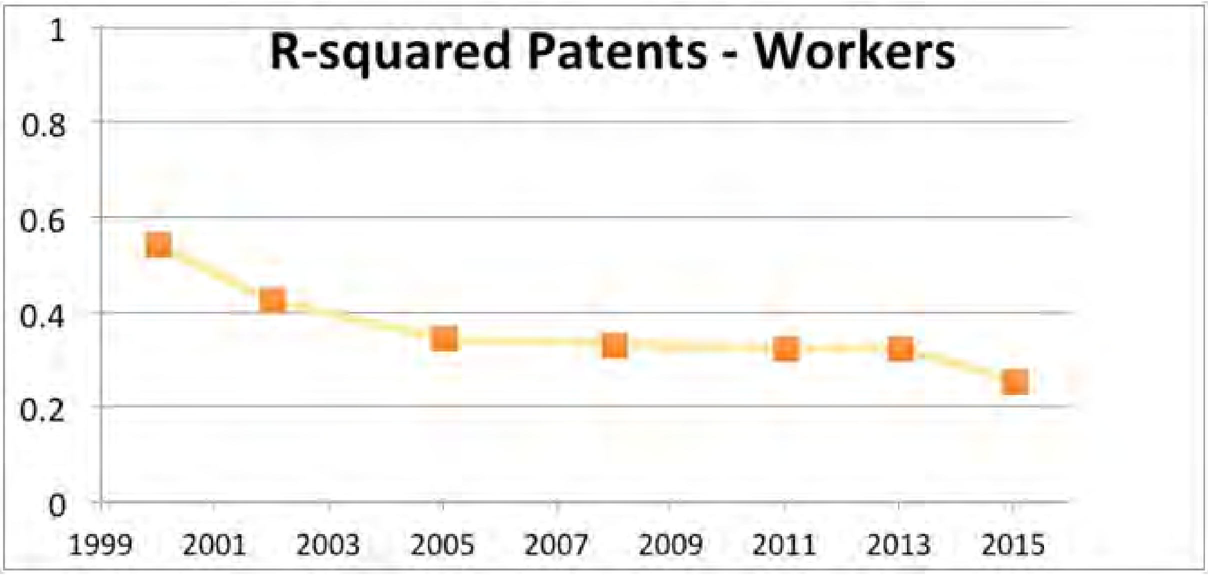

Figure 2. Coefficient of determination (R-squared) Patents-Workers for US metro areas from 2000-2015

Higher results for the R-squared value are provided if two outliers are excluded: the metropolitan area of San Jose, because of the presence of Silicon Valley, which represents a unique case in all the world, given the world-class high-tech companies located in its metro area, and the metropolitan area of Detroit, exposed to an economic crisis in the automotive industry, the main source of jobs for the workers of the area.

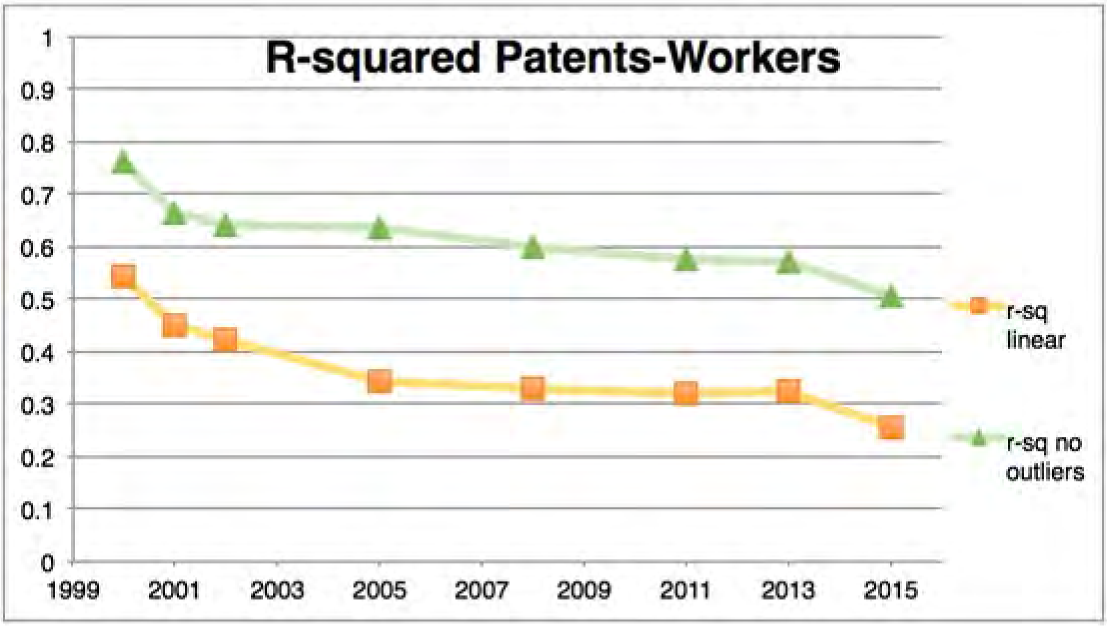

Figure 3. Coefficient of determination trend from 2000 to 2015 between patents and workers for US metro areas with and without outliers.

Analyzing the relation between the number of workers and the air quality, it has been discovered that there is a weak correlation among these indicators. The R-squared of the regression among these indicators is about 0.09 for the year 2000, while for 2015, the R-square is around 0.15, indicating a rather low correlation.

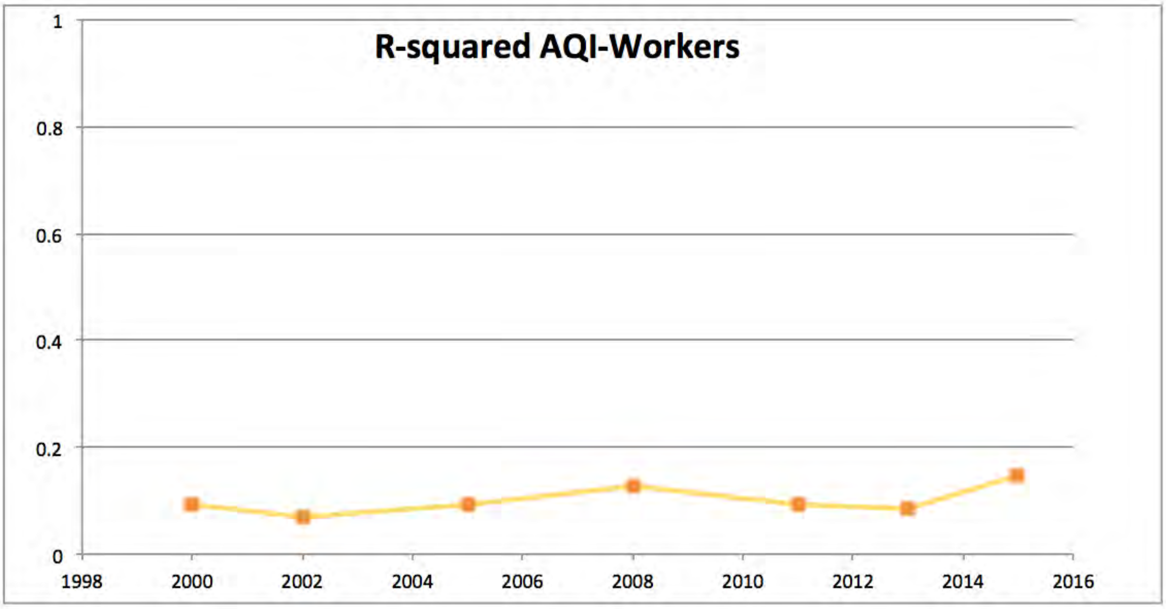

Figure 4. R-squared between AQI-Workers for the US metro areas in the period 2000-2015

Moreover, a positive correlation between air pollution and the number of workers, in this case, represents that the more polluted is the metro area, the more workers are present in the metro area, therefore the hypothesis advocated at the beginning of the study of a relationship between environmental sustainability and economic development cannot be depicted by the indicator of Air Quality Index.

3.3. Alternative fit to the data

The usual correlation between two variables is represented by a straight line. Namely, if we increase the predictor by 1 unit, the response always rises by X units. However, not all data have a linear relationship, and the model must be fitting to the curves present in the data.
Among all the non-linear curves, the one that fits the data collected best is the power function. For instance, the following charts display the number of patents produced and the number of workers per US metro areas in 2005. It can be seen that a power regression is a better fit than the linear regression for the distribution, with an R-squared value of 0.517 instead of 0.344.

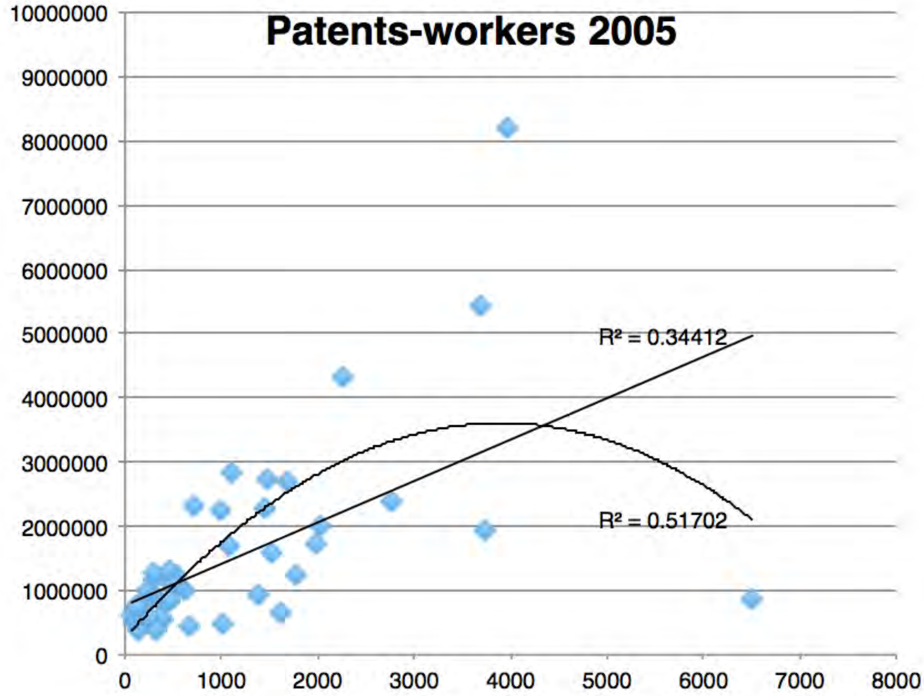

Figure 5. Linear and power regression patents - workers 2005

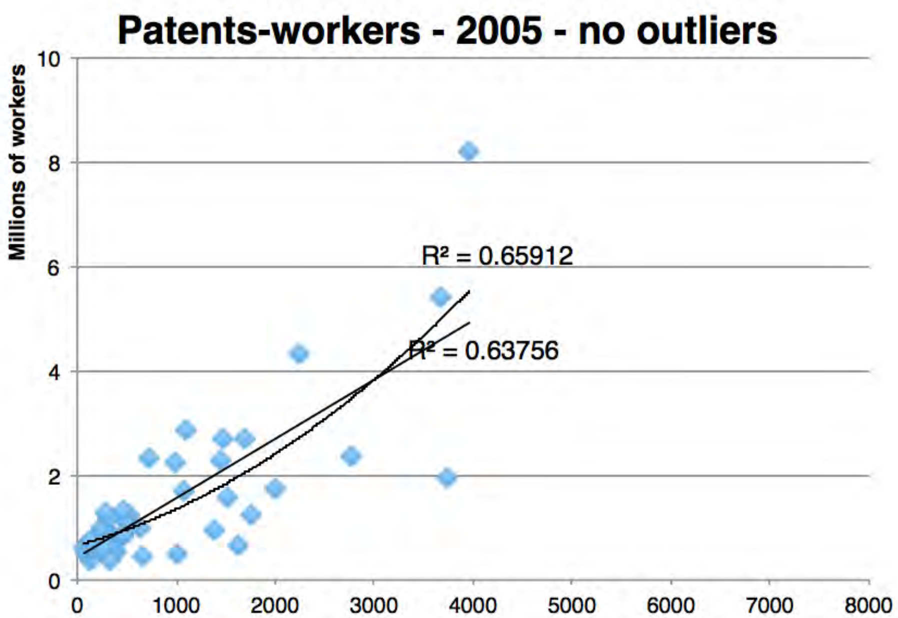

Figure 6. Linear and polynomial regression without two outliers.

If we keep on exploring at this relationship, eliminating two outliers from the data set, that are the cities of Detroit and San Jose, we can see from figure 6 that the correlation is even stronger (R2 from 0.344 to 0.637).

A better indicator for depicting the quality of environmental sustainability within urban areas is provided by Wallet Hub, which calculated the "Greenest Cities in America" (Bernardo, 2016). Another useful indicator for this purpose could be the one provided by Siemens and the Economist Intelligence Unit (Unit EI, 2009).

In order to define the greenest cities in America, Wallet Hub's analysts studied the 100 most populated cities in the United States throughout four main dimensions: Environment, Transportation, Energy Sources and Lifestyle & Policy. Some examples of indicators within the GCI: Median Air Quality Index (AQI), Greenhouse-Gas Emissions per Capita, Green Spaces (percentage of parkland), and Water Quality.

The cons of the study are represented by the fact that the data have been collected at the city level, hence if we want to compare the Greenest cities in America with the indicators collected in the SustaIn-LED database, we have to considerate the proxy that the rest of the indicators are gathered at the metropolitan area level (Bernardo, 2016). The Greenest cities in America for 2016 are somewhat correlated with the number of workers in the metro area, with an R-squared value of the linear regression of 0.0727. If we hypothesize a power regression instead of a linear regression, we can appreciate an R-squared value increased to 0.101. It is interesting also to look at the relationship between the number of patents inside the metro area and the Greenest Cities in America for 2016, by looking at the following chart (Fig. 7).

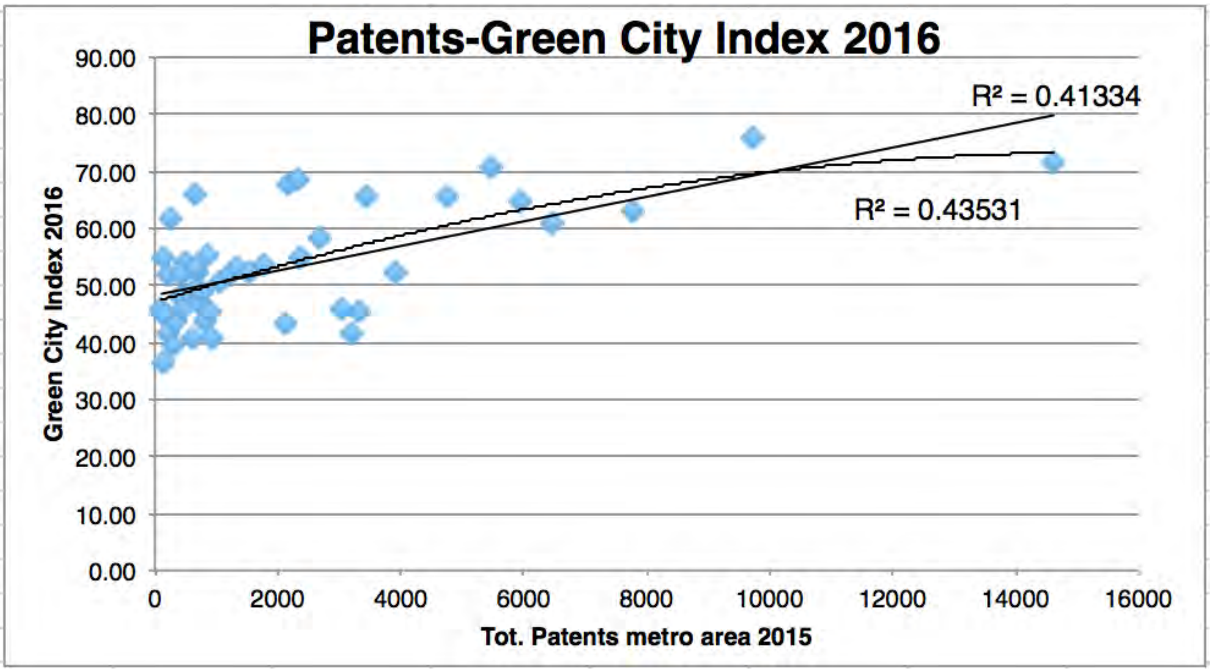

Figure 7. Regression between Patents and Green City Index for US metro areas

In this case, the rationale is slightly different from the one originally hypothesized in the model. In this circumstance, the present study explores if innovation has an impact on environmental sustainability within metropolitan areas. From the data analyzed, it seems that there is a relation among innovation and sustainability. This has been done by utilizing Utility Patents and the Green City Index, and more than 41% of the variable of the Greenest cities in America for 2016 can be explained by the variance of the number of patents registered in the metro area in 2015, as if there was a cause-effect relationship. If we use a power relation instead of a linear one as a fit to the data, the coefficient of determination goes from 41% to 43.5%, therefore representing a better fit to the data. If we keep on exploring at this relation, eliminating two outliers from the data set, that are the cities of Detroit and New Orleans, we can see that the relationship is even stronger.

The city of Detroit, in 2013, filed the largest municipal bankruptcy case in US history, which it successfully exited on December 10, 2014. However, poverty, crime, and urban blight continue to be ongoing difficulties in the urban area. The city of New Orleans, hit by Hurricane Katrina in 2005, had devastating effects also on the industrial sector and on the production of patents and trademarks; the average of patents produced in the period 2000-2004 was 145 per year, while in the period 2005-2015 the average fell to 104 patents per year.

With the pondered exclusion of the two outliers, the relation between innovation and sustainability is even stronger. The coefficient of determination rises from 0.413 to 0.439, meaning that 43.9% of the variance in the dependent variable (tot. patents in the metro area in 2015) is explained by the independent variable. If a polynomial relation as a fit to the data is hypothesized, the R-squared value rises to 0.474 (without the exclusion of the outliers the value was 0.435).

4. Conclusions

The aim of the present study has been to find out whether or not a relationship exists within local economic development and clusters, driven by sustainability, in urban areas.
Nowadays, cities in the world are dealing with sustainability issues. It has become a central topic of discussion in every single decision that has to be made concerning everybody's life and affecting the livability of urban areas. With regards to local economic development, in this study it has not been intended as mere economic growth. It rather encompasses environmental planning, business development, infrastructure provision, real estate development, and finance.
These themes are analyzed in the present research, as the attempt of the present study to merge the three main topics of local economic development, innovation, and environmental sustainability, which at first can seem as distant concepts, but are actually intertwined topics, and they all concur to the development and support contemporary urban areas, helping them to be competitive, inclusive, and economically strong.
The present study has carried out an econometric model to find out the significance of the correlation between the three main topics aforementioned. This has been accomplished with a time series from the years 2000 through 2015 in American metropolitan areas with a population larger than 1 million. The results of the year 2015 showed that 35.9% of the variance in the measure of total workers in metro areas can be predicted by measures of innovation and air quality. Higher R^2 values have been registered for the years 2000 through 2013. The highest R^2 value has been scored in 2000, with a R^2 value of 0.611, meaning that the 61.1% of the variance in the measure of total workers can be predicted by innovation and air quality. A more consistent correlation that has been found is the one between two variables: the number of patents and workers in metropolitan areas of the US, excluding two outliers. The highest coefficient of determination from the database, excluding two outliers, was 0.76 in the year 2000.
Concerning the limitations of the model, an improved database and data-retrieval will help to collect more significant data.
In the future, we can consider two possible scenarios.

- Will it be more efficient to collect more data, improving the database and utilizing the current indicators?

- Will it be better to choose different indicators in the study?

A first attempt has shown that substituting the indicator of Air Quality Index (AQI), representing the topic of environmental sustainability, has produced better results in terms of R^2 (coefficient of determination). The problem presented in the results of the negative correlation between sustainability and economic development has been partly solved by using a new indicator representing environmental sustainability, namely the Greenest cities in America or another like the Green City Index (GCI). More work has to be done to improve the results of the multivariate regression, by inserting the Greenest cities in America inside the SustaIn-LED model and verifying if there is a higher correlation, inserting the data about patents of 2016, that still have to be released by the United States Patent and Trademark Office in their annual report.
Some potential future work could be the application of the SustaIn-LED model to European urban areas. It could represent an important tool for policymakers, offering a comprehensive picture with the possible implementation

of a visual tool like a GIS platform for visualizing the parameters taken into account in the present research and that could help policymakers decide which policies need to be implemented to improve the quality of life inside urban areas.

5. References

1. 2010 Census Summary File 1. (2012). Retrieved from https://www.census.gov/prod/cen2010/doc/sf1.pdf
2. Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy,31*(7), 1069-1085. doi:10.1016/s0048-7333(01)00184-6
3. Pastor, M. and Benner, C. (2015). *Equity, Growth, and Community: What the Nation Can Learn from America's Metro Areas.* California: University of California Press. DOI: https://doi.org/10.1525/luminos.6
4. Bernardo R. (2018). 2017's Greenest Cities in America (WalletHub). Retrieved from https://wallethub.com /edu/most-least-green-cities/16246/#methodology
5. Bettencourt, L., Lobo, J., Helbing, D., Kuhnert, C., & West, G. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings Of The National Academy Of Sciences, 104*(17), 7301-7306. doi: 10.1073/pnas.0610172104
6. Bettencourt, L., & West, G. (2010). A unified theory of urban living. *Nature, 467*(7318), 912-913. doi: 10.1038/467912a
7. Feldman, M., & Lendel, I. (2010). Under the Lens: The Geography of Optical Science as an Emerging Industry. *Economic Geography, 86*(2), 147-171. doi: 10.1111/j.1944-8287.2010.01068.x.
8. Florida, R. (2009). *Who's your city?*. New York: Basic books.
9. Gaetani, R., & Berkes, E. (2015). The Geography of Unconventional Innovation. 2015 Meeting Papers, Society For Economic Dynamics, (896).
10. Glaeser, E. (2014). *Triumph of the city*. New York: Penguin Books.
11. Johnson, S. (2011). Where good ideas come from: the natural history of innovation. *Choice Reviews Online, 48*(10), 48-5785-48-5785. doi: 10.5860/choice.48-5785
12. Lee, N., & Rodríguez-Pose, A. (2014). Creativity, Cities, and Innovation. *Environment And Planning A: Economy And Space, 46(*5), 1139-1159. doi: 10.1068/a46202 .
13. Newman, P. (1999). Sustainability and cities: extending the metabolism model. *Landscape And Urban Planning, 44*(4), 219-226. doi: 10.1016/s0169-2046(99)00009-2 .
14. Sedgley, N., & Elmslie, B. (2011). Do We Still Need Cities? Evidence on Rates of Innovation from Count Data Models of Metropolitan Statistical Area Patents. *American Journal Of Economics And Sociology, 70*(1), 86-108. doi: 10.1111/j.1536-7150.2010.00764.x .
15. Shearmur, R. (2012). Are cities the font of innovation? A critical review of the literature on cities and innovation. *Cities, 29*, S9-S18. doi: 10.1016/j.cities.2012.06.008.
16. Unit EI, Siemens AG. (2009). European Green City Index (p. 51). Munich: Siemens AG. U.S. Patent and Trademark Office. (2000). .
17. United States Patent Grants by State, County, and Metropolitan Area (Utility Patents, 1990 - 1999). Office for Patent and Trademark Information Washington, Dc 20231.
18. Zimmermann, M. (2013). *The Economy of Green Cities*. Springer.