REVISTA ESPACIOS

The pluviosity effect on occupational diseases in Colombia during 2010–2014, a quasi–experimental approach

El efecto de la pluviosidad en las enfermedades de origen laboral en Colombia para el período 2010 – 2014, un enfoque quasi-experimental

CUÉLLAR MORENO, Carlos Andrés [1]
ÁLVARADO TORRES, Jhon Edisson [2]
MONROY MÁRTINEZ, Sorayda [3]

Abstract

In this work, we provide a formal approach to the econometric relationship between the volume of rainfall and occupational diseases in Colombia for the period 2010 – 2014. Our main result is that, with a random effects panel data model at the municipal level with month – by – month and year – by – year effects, we obtained evidence of an increase in occupational diseases due to an increasing volume of rainfall. Finally, we bring outlines for future research in this topic.
key words: health economics, Environmental economics, Applied econometrics

Resumen

En el presente documento, se brinda una aproximación formal al análisis de la relación empírica entre el nivel de lluvias y las enfermedades de origen laboral (EOL) en Colombia para el período 2010 – 2014. El principal hallazgo es que usando un modelo de datos panel con efectos aleatorios pero explicitando efectos de mes y de año, existe evidencia de un incremento de EOL por mayor pluviosidad. Finalmente, sugerimos dos líneas de investigación para el futuro de esta área de investigación.
Palabras clave: economía de la salud, economía ambiental, econometría aplicada

1. Introduction

The econometric relationship between the volume of rainfall and occupational diseases in Colombia is a topic of recent interest worldwide. According to Henseler, M. & Schumacher, I. (2018), to understand the impact of the pluviosity on the workers´ health is relevant, because it evidences indirectly the lags in labor productivity and as a consequence it will affect the country's economic growth process; for example, Barrios, S., Bertinelli, L., & Strobl, E. (2010) and Lanzafame, M. (2012), provide empirical evidence of the lag that rainfall generates in terms of economic growth for African nations by having a negative impact on the productivity of the labor factor and the total productivity of the used inputs.

[1] Profesor asistente – Tiempo completo, Facultad de Medio ambiente y recursos naturales. Universidad Distrital – Francisco José de Caldas. cacuellarm@udistrital.edu.co
[2] Profesor titular – Tiempo completo, Facultad de Medio ambiente y recursos naturales. Universidad Distrital – Francisco José de Caldas. jealvaradot@udistrital.edu.co
[3] Profesor de tiempo completo ocasional, Facultad de Medio ambiente y recursos naturales. Universidad Distrital – Francisco José de Caldas. sorymm29@yahoo.es

In our case, we will concentrate on the effect that the volume of rainfall has on the recognized diseases of labouring origin by the ARL (Risk labour management office) in Colombia, since it is (as already mentioned above) a previous step to the analysis of the inputs that affect economic growth; such as: labor, capital and total factor productivity. The above has been considered a topic of interest because at the level of Colombia a similar analysis has not yet been found, and therefore it is a good opportunity to provide exploratory empirical evidence on a subject as relevant as it is the level of rainfall and its effects on Colombian workers´ health.

2. Empirical Strategy

In order to make the econometric analysis, we have a municipal and unbalanced panel for 5 years, the research design that has been chosen is a panel data model with random effects (the best model obtained) and explored the behavior of the model with fixed effects that allows cleaning the group and time fixed effects that do not vary throughout time. These fixed effects are, among others, the availability of access to health services, the geographical location of each municipality and their height. Aditionally, geographic information is available at the regional level, to control for heterogeneous effects as there are 5 different political - administrative regions in the country.

On the other hand, these are the main reasons for exploring a panel data research design:

- Having a panel data is considered an enormous advantage for applied research, because we have available more data and each individual can be followed throughout the available periods. On the other hand, this econometric technique allows to capture unobservable heterogeneity between individual units and over time at the same moment, which is not possible using the Ordinary Least Squares (OLS) method.

- In the nested data regression models, different hypotheses are made about the behavior of the residuals, the most elementary and the most consistent is that of Fixed and Random Effects, which assumes that the error term (ε_{it}) can be decomposed into two parts: a fixed one, constant part for each individual (v_i) and another one that is random and meets the OLS requirements (u_{it}) remaining at the end ($\varepsilon_{it} = v_i + u_{it}$), which is equivalent to obtain a general trend by regression giving each individual a different point of origin.

- If it is compared with panel data models with random effects, in which the same specification as that of fixed effects is provided with the difference that v_i, instead of being a fixed value for each individual and constant over time for each individual, it is a random variable with mean (v_i) and variance ($\sigma (v\_i) \neq 0$). This model is more efficient (the variance of the estimates is less) but it is less consistent than the fixed effects model, that is, the random effects model is more accurate in calculating the value of the parameter, but the bias is most likely to be greater than that of fixed effects specification. In empirical work, one must choose between both of them running hausman tests.

Regarding the assumptions of the model that must hold, see (Wooldridge, 2010), we have:

- The group and time fixed effects are needed to be plausible.

- Strict exogeneity should be assumed.

- The imbalance must not be correlated with the observable characteristics.

- The imbalance must not be correlated with the idiosyncratic error.

The last two have been econometrically validated with a random imbalance test and this was significant and assuming that the first two are fully met, we have the following model to estimate:

$$y_{it} = \alpha + \beta X_{it} + v_i + u_{it}$$

So we have:

y_{it} = Occupational diseases recognized by the ARLs in Colombia at the municipal level on a monthly basis.

$\ln X_{it}$ = Growth rate of the pluovisity volume at the municipal level on a monthly basis.

i.month = A Dummy variable that captures the effect for each of the months in relation to the first month of the year, namely January of our relationship of interest.

i.year = A Dummy variable that captures the effect for each of the years in relation to the first year of the panel, namely 2010 of our relationship of interest.

Finally, as can be seen in Figure 1, it is expected to obtain a positive relationship between the growth rate of rainfall and the number of occupational diseases at the municipal level, since as rainfall increases, it is expected that they will increase. Now, this gives evidence that the respiratory diseases and according to Chang, T., Zivin Graff, J., Gross, T., & Neidell, M. (2016, 2019), at the firm level, it is not so important whether the work activity is carried out indoors or outdoors, the negative effects of climate will affect the workers´ health and will generate collateral effects on labor productivity.

Figure 1
Scatter graph Occupational illness
vs. Precipitation growth rate
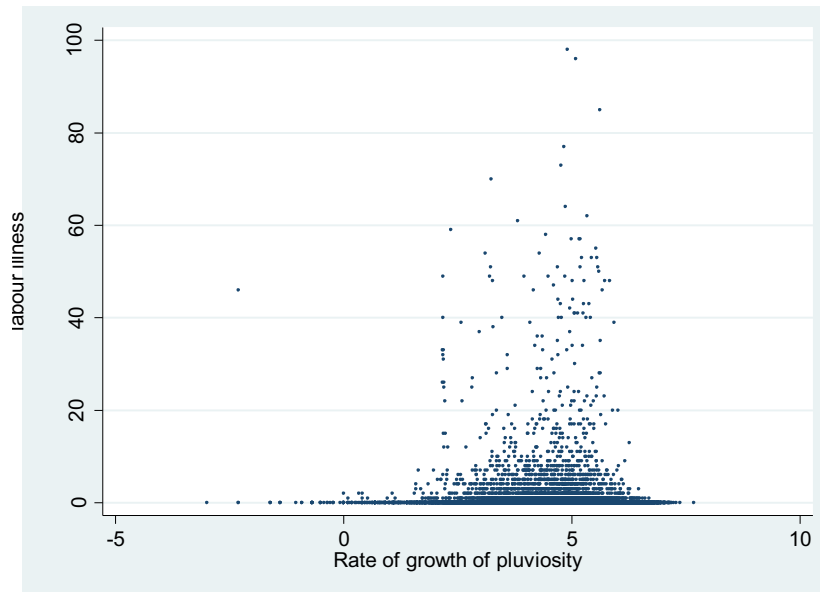
Source: Authors' calculations, using Stata 16.1 Software

3. Results

3.1. Data analysis

In this work, we used three databases with broad national recognition in Colombia, they are: i) The municipal panel of CEDE – UNIANDES, ii) The database on occupational risks of FASECOLDA and iii) The database of IDEAM on rainfall and temperature for 27 departments and 639 municipalities that remained after joining the bases; it is added that the information on temperature is not susceptible to be used, since the number of missings values

is very high. In addition, the longitudinal cut used is the years 2010 – 2014 and the observation unit in all cases is the municipal level, see table 1 for descriptive statistics.

Table 1
Descriptive statistics

Variable	Observations	Mean	Standard deviation	Mín.	Máx.
Labour illness	32.059	0,4059702	3,05821	0	98
Pluovisity	29.006	172,9215	158,2267	0	2.150
Pluviosity (square)	29.006	54.936,68	118.791	0	4.622.500
Ln (Pluviosity)	28.634	4,738214	1,050091	-2,995732	7,673223

Source: Authors´ Calculations. Using Software Stata 16.1

The parameters of interest for the regression model are ARL – certified occupational disease and the volume of rainfall at the municipal level. On the other hand, the dummy variables have been taken to classify by geographic region and control for heterogeneous regional effects.

Regarding the descriptive statistics, since the variables occupational illness and precipitation come from different databases, there is not the same level of observations (hence the unbalance of the panel), but it has already been mentioned that when performing the random imbalance test, it has been possible to verify that the imbalance is plausibly random. On the other hand, when analyzing the standard deviation, it is considered ex – ante that it will be more appropriate to work with the logarithm of precipitation, and thus eliminate the effect of the unit level of the variable on the coefficient.

3.2. Results and robustness

First, according to table 2, a regression by Ordinary Least Squares (Model 1) were run, in order to have a starting point about the econometric relationship between occupational diseases and the volume of rainfall at the municipal level in Colombia, and we obtained a consistent estimator at the 1% level of significance, although negative; that is, if the volume of rainfall increases by one unit, occupational diseases will decrease by – 0.000769 units, which is an opposite effect to that initially expected; Also, the exercise was done taking the logarithmic transformation of rainfall (Model 2), and the estimator is consistent but at the same time it indicates a negative relationship between the rate of growth of rainfall and occupational diseases; that is, if the growth rate of rainfall increases by 1% occupational diseases will decrease by 0.001013 units.

To verify if the existing relationship for our panel is negative, four additional regressions were specified and run and thus to be able to test the robustness of the results initially obtained, as follows:

I.   Taking two lags for the precipitation variable (Model 3) and although the coefficients for the lags are not significant, this is significant for the variable without the 10% transformation and the negative sign persists.

II. Using the squared of the precipitation variable, to explore non-linear effects (Model 4).

III. Another model in which the natural logarithm of the rainfall volume has been used to correct for the unit of the variable and including the effects month by month and year by year related to the first month of the year and the first year of the panel (Model 5).

IV. Two models correcting for heterogeneous regional effects for different specifications of the independent variable of interest (Models 6 and 7), which is not significant in any of the parameters of interest.

Table 2

Econometric results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Labor illness								
Pluviosity	-0,000769*** (.0000697)		-0,0001231* (0,0000759)	-0,0003893*** (0,0001199)			-0,0002966 (0,0003147)	
L1.pluviosity			-0,0001082 (0,0000759)					
L2.pluviosity			-0,0000431 (0,0000688)					
Pluviosity2				2,85e-07* (1.48e-07)			1.81e-07 (1.63e-07)	
Ln (Pluviosity)		-0,1013334*** (0,0186971)			**-0,0229938** (0,0090384)	-0,042701 (0,0629651)		**-0,0236472*** (0,0090264)**
Month								
2				0,0515659 (0,031424)	**0,0602152*** (0,0329168)**	0,0584311* (0,0330741)	0,0486281 (0,0314095)	**0,0604554*** (0,0329131)**
3				0,0934532*** (0,0320486)	**0,1016113*** (0,0338132)**	0,101234*** (0,0340922)	0,0872336*** (0,0320927)	**0,1021161*** (0,0338066)**
4				0,0997253*** (0,0339737)	**0,1003762*** (0,0357376)**	0,1000722*** (0,0360993)	0,0872731** (0,0341328)	**0,1011915*** (0,0357267)**
5				0,11219*** (0,0338494)	**0,1162248*** (0,0356641)**	0,1174726*** (0,035995)	0,1019194*** (0,033964)	**0,1170493*** (0,0356543)**
6				0,0574977* (0,032832)	**0,0688354*** (0,0345562)**	0,0719987** (0,0350008)	0,0543093* (0,0330455)	**0,0694024*** (0,0345496)**
7				0,056769* (0,0333273)	**0,0705537*** (0,0351215)**	0,0734642** (0,0354987)	0,0568308* (0,0335839)	**0,0711101*** (0,0351147)**
8				0,1064164*** (0,0331681)	**0,1224657*** (0,0349355)**	0,1268908*** (0,0354088)	0,1083101*** (0,0334472)	**0,1229605*** (0,0349294)**
9				0,0791552** (0,0332949)	**0,0970599*** (0,0350705)**	0,1013524*** (0,0355085)	0,0827003** (0,0335892)	**0,0975444*** (0,0350644)**
10				0,0862751** (0,034556)	**0,1003862*** (0,0366247)**	0,1020879** (0,0368015)	0,0849913** (0,0347598)	**0,1012354*** (0,0366149)**
11				0,0730086** (0,0350923)	**0,0889977*** (0,0370335)**	0,0897043** (0,0371884)	0,0747877** (0,0353007)	**0,0898536*** (0,0370236)**
12				0,005777 (0,0338834)	**0,0302531** (0,0357152)	0,0308074 (0,035773)	0,0156531 (0,0340532)	**0,0307374** (0,0357083)
Year								
2011					**0,021393** (0,0190459)	0,0205594 (0,0190703)	0,0189767 (0,0188153)	**0,0213649** (0,0190427)
2012					**0,1286856*** (0,0194558)**	0,1280833*** (0,0194753)	0,1261648*** (0,0192268)	**0,128466*** (0,0194524)**
2013					**0,1269778*** (0,0205248)**	0,1264962*** (0,0205344)	0,1253264*** (0,0202645)	**0,1269746*** (0,0205198)**
2014					**0,1750769** (0,039532)	0,1767622*** (0,0395971)	0,1702605*** (0,0391375)	**0,1763906*** (0,0395232)**
Interaction Inpluviosity*region y pluviosity*region								
Interaction region A						0,0221233 (0,0636024)	0,0000121 (0,0003093)	
Interaction region C						0,0002169 (0,064866)	0,0000747 (0,0003246)	
Interaction region P						0,0330797 (0,0663468)	0,000078 (0,0003241)	
Interaction region O						0,0321118 (0,0663679)	0,0001152 (0,0003269)	
Interaction region Am						0 (omitted)	0 (omitted)	
Constant	0,5226002*** (0,0261966)	0,8736587*** (0,0935599)	0,4472052*** (0,0168401)	0,3734116*** (0,0231283)	**0,3533467*** (0,0425201)**	0,3486896*** (0,0441399)	0,2890979*** (0,0266969)	**0,3187567*** (0,1100412)**
Observations	28.454	28.088	26.728 (Grupos 639)	28.454 (Grupos 639)	**28.088 (Grupos 639)**	28.088 (Grupos 639)	28.454 (Grupos 639)	**28.088 (Grupos 639)**
R- square	0,0016	0,0013	N.A.	N.A.	**N.A.**	N.A.	N.A.	

Standard errors in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Source: Authors´ Calculations, using software Stata 16.1

With these in mind, it is highlighted that the most appropriate specification is to estimate Model 5, the month-to-month effects, using a dummy variable that finds the effects in relation to the first month of the year and the result shows that if there is a positive relationship between rainfall and the number of occupational diseases for those months of the rainy season in the country, while for the months of the dry season the result is not significant; that is, in the rainy season there is an increase in occupational diseases in Colombia, but there is no significant statistical evidence to ensure that the same occurs in the dry season, see Figure 2. Furthermore, the

same exercise was performed to calculate the effects by year in relation to the initial year 2010, and it is obtained that these have a positive sign for the 4 years; although the coefficients are significant only for the years 2012, 2013 and 2014; which provides empirical **evidence in favor of the initial research question.**

The same initial results are obtained in terms of the sign of the coefficient but significant at 1% (Models 1 and 2), at 5% (Model 3) and the coefficients of the last two models are not statistically significant respectively: Finally, a regression model with random effects (model 8) was specified and a Hausman test was performed to determine which of the two methodologies would be the most appropriate, and no significant differences were found in the coefficients of both models, thus confirming that the most appropriate model to solve our research question is a panel data with random effects and not with fixed effects, because the variance of the estimator is lower in the model with random effects. So we have consistent and efficient estimators with this random effects model.

4. Conclusions

As a starting point, it has ben possible to verify that the econometric relationship between the level of rainfall and occupational diseases is robust according to different specifications of the variables of interest. Now, although in global terms it has been found that the relationship is negative, that is, an increase in rainfall or in the rate of growth of rainfall generates a decrease in occupational diseases. It has been found that when estimating the monthly and annual effects this relationship is reversed and consistent results are obtained compared to the fact that the higher the level of rainfall or its growth rate, then the more occupational diseases will be generated.

This should be taken with caution given the quasi – experimental design of this econometric exercise, because approaching the most recent developments in the analysis of environmental factors and the performance of workers, i.e. Chang, T., Zivin Graff, J., Gross, T., & Neidell, M. (2016, 2019), it can be argued that this relationship in principle is consistent with the results they obtain, when the level of rainfall increases in a Municipality, respiratory diseases are going to increase and therefore it is likely that occupational diseases will increase because this is how labour activities are carried out indoors and not outdoors, the change in temperature (downward) that generates the increased rainfall is likely to result in the abscence of optimal humidity and general health conditions in workplaces, workers end up getting sick from being exposed to these climatic conditions for such long periods of time.

Given the above, it is evident that there is still a long way to go for validating this type of econometric relationship in a country like Colombia, and a possible future research course is to design an economic growth model that allows testing whether climate factors actually decrease. the productivity of the inputs in general and the total productivity of the inputs in particular. Another possible course of research is to design randomized controlled experiments at the firm level to validate the effects of climate change on the health of the working population, following the experimental approach that is evidenced in much of the literature on health and environmental economics for this issue in particular and, perhaps this way, it can be corrected for possible endogeneity and obtain a global relationship such as that initially expected by the literature.

Finally, in terms of economic policy recommendations, it is considered essential to continue promoting governmental programmes against the harmful effects of climate change, since obviously in the short term negative effects are already being evidenced on the workers´ health that will provoke reductions, and that according to the recent literature has negative impacts on the process of economic growth in Colombia, it is neccesary more research in the future in this topic as well.

References

Barrios, S., Bertinelli, L., & Strobl, E. (2010). Trends in rainfall and economic growth in Africa: a neglected cause of the African growth tragedy, Review of economics and statistics, 92 (2): 350 – 366.

Chang, T., Zivin Graff, J., Gross, T., & Neidell, M. (2016). Particulate pollution and the productivity of pear packers. American Economic Journal: Economic Policy, 8 (3): 141 – 169.

Chang, T., Zivin Graff, J., Gross, T., & Neidell, M. (2019). The effect of pollution on worker productivity: evidence from call center workers in China. American Economic Journal: Applied Economics, 11(1): 151 – 172.

Henseler, M., & Schumacher, I. (2018). The impact of weather on economic growth and its production factors, Climatic change, 154: 417 – 433.

Lanzafame, M. (2012). Temperature, rainfall and economic growth in Africa. Empirical Economics, 46: 1 – 18. Journal of the Institute for Advanced Studies, Vienna, Austria.

Wooldridge, J. (2010). Econometric analysis of cross section and panel data. The MIT press. Cambridge, Massachusetts. Second edition