Land Use Policy Recommendations——Based on Empirical Analysis of The Effect of Agricultural Land Expansion on Greenhouse Gas Emission

Zhaojie Wang

O'Neill School of Public and Environmental Affairs, Indiana University Bloomington, IN, USA

Abstract. Agricultural activities emit about 10-12% greenhouse gases each year. Moreover, the total amount is on the rise, which can significantly contribute to global warming. To get governments to pay more attention to greenhouse gas emissions from agricultural activities, so that the necessary regulations can be implemented. This paper studies the relationship between agricultural land expansion and greenhouse gas emissions. My initial hypothesis was that ceteris paribus, the agricultural land expansion would cause a rise in greenhouse gas emissions. I obtained 25 years of data for 12 countries from the World Development Indicators Database of the World Bank. And created a panel data frame. The variables exist in it are: Total greenhouse gas emissions (kt of CO2 equivalent), Energy use (kg of oil equivalent per capita), Alternative and nuclear energy (% of total energy use), Agricultural land (% of land area), GDP per capita (current US$). In order to cope with the possible heteroscedasticity problem in the regression, several of the variables were processed logarithms. This paper used a Two-way Fixed effect model which controls fixed effects from both cross section and time. The explained variable is log (GHG emissions), and the explanatory variables are log (energy use per capita), log (GDP per capita), clean energy using rate, and percentage of agricultural land in total land. However due to the presence of multicollinearity, log (GDP per capita) was removed. This paper also tested the autocorrelation and used the Cochrane-Orcutt Iterative Process to estimate the autocorrelation coefficient to transform all variables so that solve the autocorrelation issue. The final regression model is as follows: **ln(Greenhouse Gases emission)=β_i + 1.06*ln(Energy use per capita) – 0.009*(Alternative and nuclear energy % of total energy use) + 0.0074*(Agricultural land % of land area). β_i =**

```
> fixef(plm_2FX)
    AUS     BRA     CHN     DEU     FRA     IND     JPN     KOR     THA      UK     USA     VNM
3.6074  5.9772  7.6216  4.6262  4.2196  7.4450  5.2377  4.2565  4.5428  4.2132  5.8597  5.2504
```

. The overall F-test of this model is significant, and the individual t-tests of the coefficients of each variable are all significant. R^2=92.18, adjR^2=91.04. This shows that the explanatory power of the model is strong. This outcome is consistent with my initial hypothesis that agricultural land expansion does lead to an increase in greenhouse gas emissions. The other two coefficients are also consistent with our common sense. This result can help governments to decide on the planning of land use. If the current amount of energy used and the types of energy used do not change, a country develops agriculture, it will inevitably lead to more greenhouse gas emissions. More Importantly, it is not like industrial manufacture, normally there are few regulations and policies on agricultural GHG emissions. But the quantitative results of the model tell us that as agriculture expands, it is necessary for the government to implement regulations and policies on it.

Keywords: Land Use Policy; Land Economics; Agricultural Land Expansion; Greenhouse Gas Emission; Fixed Effects Regression Model.

1. Introduction

In recent years, with the improvement of human awareness of global warming, governments continue to introduce emission reduction policies, and citizens pay more and more attention to environmental protection in their daily life. In terms of technology, clean energy and electric vehicles are widespread in cities. Refrigerators and air conditioners that do not use Freon as refrigerant have also appeared on the market. In terms of industrial production, both developed and developing countries have begun to limit the emissions of enterprises and take carbon neutrality as the guiding direction. At the 21st United Nations Climate Change Conference, 178 parties signed the Paris Agreement to jointly take actions to respond to climate change, including the protection of forests

and the limitation of carbon emissions. Although humanity has made great progress in reducing emissions and has a clear plan for the future, there is still a huge hidden danger. There is almost no regulation of greenhouse gas emissions related to agriculture in all countries. In fact, the greenhouse gases produced by agriculture and animal husbandry are very large. Agriculture is the basis for sustaining people's livelihood, with the growth of population, the production of the agricultural sector can only increase. Therefore, humans must be vigilant about the impact of agriculture on global warming. Understand how the expansion of agricultural land will affect greenhouse gas emissions.

2. Background

According to the *IPCC AR4 Climate Change 2007: Mitigation of Climate Change* "Agriculture accounted for an estimated emission 10-12% of total global anthropogenic emissions of greenhouse gases (GHGs)). CH4 contributes 3.3 GtCO2-eq/yr and N2O 2.8 GtCO2-eq/yr. Of global anthropogenic emissions in 2005, agriculture accounts for about 60% of N2O and about 50% of CH4 (medium agreement, medium evidence). " (IPCC AR4 Ch.8). This means that agriculture not only produces a lot of carbon dioxide, but it is also a major source of N2O and methane. Methane is produced mainly through the excretion and fart of livestock. On farms in some countries, people use manure as fertilizer on their fields. They build Septic tanks to turn animal waste and decaying plants into fertilizer. More methane is produced in that process. Methane is a far more dangerous greenhouse gas than carbon dioxide. It is about 20 times more potent than carbon dioxide in contributing to the greenhouse effect. So even relatively small changes in methane emissions can have an effect on global warming. (John 2011) As for N2O, it is present in very low levels in the atmosphere and is defined as a trace gas by the IPCC. But its warming effect potential is 298 times that of carbon dioxide (IPCC, 2007). And crucially, it can remain in the atmosphere for a long time, and as it accumulates, it will have an increasingly significant impact on global warming in the future. In agriculture, N2O is produced by nitrifying and denitrifying microorganisms living in the soil. Because nitrogen is one of the essential nutrients for crop growth, people apply nitrogen fertilizer to the farmland, which promotes microbial activity and converts nitrogen into N2O through nitrification and denitrification. So, both N2O and methane emissions are closely related to agricultural activities, and according to IPCC, global agricultural CH4 and N2O emissions have increased by more than 20% in the past decade. Consider that agricultural activity contributes to global warming, which is getting worse. I think it is important for governments to be clear about the potential threat posed by the expansion of agricultural land. Thus, I want to investigate the relationship between agricultural land growth and greenhouse gas emissions.

Based on current knowledge, my original hypothesis is that the expansion of agricultural land will lead to an increase in greenhouse gas emissions which exacerbate global warming. To prove this hypothesis, I needed to use data to run a regression model. In the model, I take the percentage of agricultural land in the total land as the independent variable. The change of this variable can represent the expansion and contraction of agricultural land. And the Green House emissions should be naturally the dependent variable. I also want to add energy consumption and the proportion of clean energy to total energy into the model. Because these two factors have a decisive impact on GHG emissions. The specification errors caused by omitting them can be serious. In addition, I want to add the variable of GDP per capita, since it can reflect to some extent the consumption power of the people of a country and the overall development and prosperity of the country. They are all likely to have an impact on GHG emissions.

3. Data and Methodology

3.1. Data

I obtained the initial data for the experiment from the World Bank - World Development Indicators Database[1]and I created a panel data frame which include 12 countries, they are United States, China, South Korea, Japan, Australia, Thailand, Vietnam, France, Germany, Britain, Brazil, India. For each country, it has observations from 1990 to 2014, 25 years and the data are arranged as follows:

Table 1. Variable presentation

	lnGHG	lnErgpc	Cleanerg_rate	lnGDPpc	Agland_rate	Year
1	14.99117	6.642481	0.9187117	5.761689	53.75170	1990
2	15.03566	6.602387	0.9893552	5.808569	54.35426	1991
3	15.07301	6.623572	1.0140210	5.903891	54.66483	1992
4	15.13630	6.669661	1.1403410	5.933279	55.07949	1993
5	15.17801	6.704614	1.6385029	6.160136	55.56831	1994
6	15.27627	6.764848	1.7017239	6.412896	55.56812	1995

lnGHG=log(GHG)
GHG = Total greenhouse gas emissions (kt of CO2 equivalent)
lnErgpc = log(Ergpc)
Ergpc = Energy use (kg of oil equivalent per capita)
lnGDPpc = log(GDPpc)
GDPpc = GDP per capita (current US$)
Cleanerg_rate = Alternative and nuclear energy (% of total energy use)
Agland_rate = Agricultural land (% of land area)

Table 2. Statistical description

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
lnGHG	1	299	13.70	1.10	13.47	13.66	0.80	11.21	16.30	5.09	0.43	-0.13	0.06
lnErgpc	2	299	7.69	0.95	8.20	7.76	0.74	5.58	8.99	3.41	-0.63	-0.89	0.06
Cleanerg_rate	3	299	10.95	11.51	10.30	8.61	9.94	0.39	49.59	49.20	1.86	3.17	0.67
lnGDPpc	4	299	9.04	1.69	9.89	9.22	1.20	4.57	11.13	6.56	-0.76	-0.76	0.10
Agland_rate	5	299	43.57	17.21	47.81	43.86	18.71	12.40	75.24	62.85	-0.25	-1.00	1.00

Table 3. Variables in the Regression Model

Variable[2]	Description
lnGHG	The log value of Total greenhouse gas emissions (kt of CO2 equivalent), Greenhouse gases include CO_2, N_2O, Freon, CH_4, etc.
lnErgpc	The log value of Energy use (kg of oil equivalent per capita). “Energy use refers to use of primary energy before transformation to other end-use fuels, which is equal to indigenous production plus imports and stock changes, minus exports and fuels supplied to ships and aircraft engaged in international transport.” (The World Bank)
Cleanerg_rate	Alternative and nuclear energy (% of total energy use). “Clean energy is noncarbohydrate energy that does not produce carbon dioxide when generated. It

[1] https://databank.worldbank.org/source/world-development-indicators/Type/TABLE/preview/on#

[2] We take a logarithmic treatment for some of the variables. To alleviate the possible **heteroscedasticity** problem.

	includes hydropower and nuclear, geothermal, and solar power, among others." (The World Bank)
Agland_rate	Agricultural land (% of land area). Agricultural land refers to the share of land area that is arable, under permanent crops, and under permanent pastures. Arable land includes land defined by the FAO as land under temporary crops (double-cropped areas are counted once), temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land temporarily fallow.
lnGDPpc[3]	The log value of GDP per capita (current US$)

Potential specification errors:

There are other factors that can have an impact on Green House gases Emission. And their influence is not captured by my explanatory variables. Such as the annual change in forest area in a country. Deforestation causes the carbon fixed in the trees to go into the atmosphere. In contrast, the birth of new trees absorbs carbon dioxide from the atmosphere. I will discuss the details of forests in a subsequent section. The environmental protection policies of different countries can also have a significant impact on GHG emissions. In addition, how seriously people in different countries attach to the problem of global warming also affects their greenhouse gas emissions. Especially for countries below the sea level, where people are likely to do huge effort to reduce greenhouse gas emissions in production and living.

3.2. Methodology

As we know, there are various types of greenhouse gases. CO_2, N_2O, Freon, CH_4 are the main greenhouse gases in the earth's atmosphere. The activities that produce them are complex, such as energy consumption, deforestation, agriculture, transportation, and the use of household appliances. Although the four variables in my model represent a large number of main emission activities, the omitted variables issue still needs to be worried. The omitted variables will add the composite error to our multiple linear regression model. Therefore, our basic model assumption at the beginning is as follows:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \cdots + \beta_k X_{kit} + (a_i + u_{it})$$

3.2.1 Fixed Effect or Random Effect

For the error a_i, I am not clear whether it is related to the independent variables. If it is correlated, then we cannot model using the OLS method because the estimates will be biased. If it has nothing to do with the independent variables, the estimation of OLS method can be not biased, but the composite error will be serially correlated.

From the perspective of the data frame, it contains 12 countries, it is reasonable to believe that there are cross-sectional effects various from country to country. Most intuitively, countries differ in size and geographical conditions, and the ways of economic development. As a result, for a single country, one of the variables over 25 years can be smooth, but when we combine data together, one variable's observation value can jump from country to country. However, as I analyzed in the previous paragraph, it is not obvious whether the cross-sectional effect that switch across countries is brought by a variable inside or outside my model. I can only make some assumptions first.

Endogenous assumptions:

We can suppose each country has special features. For example, the economies of China and India are based on manufacturing that are big greenhouse gas polluters, and these countries use less clean technologies in production. On the other hand, developed countries such as the United States and the United Kingdom have completed industrial transformation and upgrading. Technology and information industries account for a large proportion, and pollution treatment equipment is widely

[3] This variable was later dropped from the final model due to multicollinearity problems. For details, please see the Technical Appendix, Testing and Correcting for NMC section.

used in production. So for them, the same amount of GDP or Energy use may be associated with smaller greenhouse gas emissions than for those developing countries. As for countries like Thailand, which are dominated by tourism and services, their emissions even be smaller. Because of **energy use per capita** and **GDP per capita** in my model, the effect brought by country's special features will exist in the model in an endogenous way. In other words, composite error is correlated with our explanatory variables. If this is the case, we can solve it using the Fixed Effect Model.

Exogenous assumptions:

There are also effects that may not be correlated with the independent variables but vary across countries and directly affect our dependent variable (GHG emissions). Like deforestation. Part of the carbon emissions from deforestation can be captured by the energy use variable, but the amount of carbon that has been stored in forests and released directly into the atmosphere because of deforestation cannot be captured. This part of carbon emissions will have a direct impact on the dependent variable, but not relate to the explanatory variables, and it varies with different countries. At this point, there is an exogeneity effect.[4] It can be solved by the Random Effect Model.

Because I used long Panel data, which are more suitable for taking fixed effect model. Secondly, as I analyzed in endogenous assumptions, we cannot rule out the possibility of composite error and explanatory variable correlation, so Fixed effect model is more appropriate.

3.2.2 One-Way or Two-Way

Consider that I used data from 1990 to 2014. During this 25-year period, the issue of global warming has received increasing attention. Governments have gradually introduced various policies, and the technology to reduce emissions has developed rapidly. I believe the time effect exists, so it is more appropriate to use the **Two-Way Fixed Effect Model**.

4. Regression Analysis Results

```
Coefficients:
               Estimate Std. Error t-value  Pr(>|t|)
lnErgpc       1.0645784  0.0225906 47.1248 < 2.2e-16 ***
Cleanerg_rate -0.0094362  0.0018225 -5.1777 4.503e-07 ***
Agland_rate    0.0074388  0.0016322  4.5575 7.972e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    9.6818
Residual Sum of Squares: 0.7564
R-Squared:      0.92187
Adj. R-Squared: 0.91045
F-statistic: 1022.64 on 3 and 260 DF, p-value: < 2.22e-16
```

Figure 1. Summary of regression analysis

(F) Significance Test :

H0: All the coefficients equal to zero.

H1: At least one of the coefficients is not zero.

Given an F value of [1022.64], and a p-value of [<2.22e-16], we can reject H0 at the [0.05] level of significance, indicating that there is statistically significant evidence of a relationship between dependent and the independent variables.

(T) Significance Test:

[4] My model does not incorporate the two variables of forest area and forest rate. The reason is that there is a mutual transformation relationship between the forest rate and our target variable agricultural land rate. The presence of both in one regression will lead to collinear problems.
I tried to add Forest area to the model to reflect the effect of forest carbon absorption and release, but the t-test was not significant. So, it was dropped.

H0: The coefficient of $X_{it} = 0$.

H1: The coefficient is not zero.

Given a t-value of [47.12] and a p-value of [<2.22e-16], we can reject H0 at the [0.05] level of significance, indicating that there is statistically significant evidence of a relationship between [ln(GHG)] and [ln(Ergpc)] holding the effects of other independent variables constant. The standard error of coefficients of ln(Ergpc) indicates the average error in estimating it is [0.023].

Given a t-value of [-5.177] and a p-value of [4.503e-07], we can reject H0 at the [0.05] level of significance, indicating that there is statistically significant evidence of a relationship between [ln(GHG)] and [Cleanerg_rate] holding the effects of other independent variables constant. The standard error of coefficients of Cleanerg_rate indicates the average error in estimating it is [0.0018].

Given a t-value of [4.557] and a p-value of [7.972e-06], we can reject H0 at the [0.05] level of significance, indicating that there is statistically significant evidence of a relationship between [ln(GHG)] and [Agland_rate] holding the effects of other independent variables constant. The standard error of coefficients of Agland_rate indicates the average error in estimating it is [0.0016].

(This is the final model, having already dealt with the issues of NMC and Autocorrelation. See the technical Appendix for details.)

【Final Regression Equation】

ln(Greenhouse Gases emission)=β_i + 1.06*ln(Energy use per capita) – 0.009*(Alternative and nuclear energy % of total energy use) + 0.0074*(Agricultural land % of land area)[5]

```
> fixef(plm_2FX)
    AUS     BRA     CHN     DEU     FRA     IND     JPN     KOR     THA      UK     USA     VNM
3.6074 5.9772 7.6216 4.6262 4.2196 7.4450 5.2377 4.2565 4.5428 4.2132 5.8597 5.2504
```

Figure 2. Intercept value based on fixed effect

The F-statistic for the overall model is significant, which means the model is valid, its independent variables do relate to the dependent variable. And each independent variables' individual t-test result shows that the variable relates to the dependent variable. So, the coefficients of each of our variables are valid. The model's R^2=0.9218 that means 92.18% of the variation in dependent variable can be explained by independent variables. And its adjR^2=0.9104 that means 91.04% of the variation in dependent variable can be explained by independent variables after adjusting for the number of independent variables. Therefore, the explanatory power of the model is very strong.

From the regression equation, we can know:

1. Holding other variables constant, a 1 % increase in Energy use per capita will have a 1.06% change in Greenhouse gases emission on average.

2. Holding other variables constant, a 1 unit increase in alternative and nuclear energy using rate will have a -0.9% change in Greenhouse gases emission on average.

3. Holding other variables constant, a 1 unit increase in proportion of agricultural land will have a 0.74% change in Greenhouse gases emission on average.

5. Conclusion

By using multiple regression model, I prove the initial hypothesis that the expansion of agricultural land will lead to an increase in greenhouse gas emissions for a country. And my regression model allows us to quantify the impact of agricultural land expansion on greenhouse gas emissions growth. Holding a country's current per capita energy consumption and clean energy use constant, for every one percent increase in agricultural land as a percentage of all land would increase 0.74% greenhouse gases emissions on average. This conclusion can help the government make better decisions on how to arrange the country's land use.

However, the proportion of agricultural land is only a rough proxy. I recommend that scholars continue to study the specific effects of different scenarios of agricultural land expansion on

[5] lnGDPpc is not included in the final model, it was removed because of the multicollinearity problem.

greenhouse gas production. For example, the expansion of agricultural land through deforestation, the expansion of agricultural land through reclamation of waste land, and the expansion of agricultural land through transforming urban land.

References

[1] Damodar N. Gujarati, Dawn C. Porter. (2009) Basic Econometrics Fifth Edition Published by McGraw-Hill/Irwin

[2] IPCC (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Stocker, T.F., D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press

[3] IPCC (2007) Climate Change 2007: AR4 Climate Change 2007: Mitigation of Climate Change

[4] Jonathan Kropko, Robert Kubinec. 2020 Interpretation and identification of within-unit and cross-sectional variation in panel data models

[5] Methane as a Greenhouse Gas: Why the EPA Should Regulate Emissions from Animal Feeding Operations and Concentrated Animal Feeding Operations Under the Clean Air Act Author(s): JOHN VERHEUL Source: Natural Resources Journal , Spring 2011, Vol. 51, No. 1 (Spring 2011), pp. 163- 187

[6] Smith, P., D. Martino, Z. Cai, D. Gwary, H. Janzen, P. Kumar, B. McCarl, S. Ogle, F. O'Mara, C. Rice, B. Scholes, O. Sirotenko, 2007: Agriculture. In Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[7] Wooldridge JM (2002). Econometric Analysis of Cross-Section and Panel Data. MIT Press.

[8] Yves Croissant, Giovanni Millo.(2008) Panel Data Econometrics in R: The plm Package. Journal of Statistical Software