

# Data Extraction Form Energy

Alternative to EPPI to extract data from papers for both the cells "effect of electrification on formal employment" and "effect of emissions trading scheme on carbon emissions"

\* Required

## General Information

To be filled in once per study

1. ID of person extracting data \*

2. Full Name of First Publication Author \*

3. Year Of Publication \*

Number must be between 1980 ~ 2021

4. Item ID internal \*

*You can find this in EPPI (search by first author for speed)*

Number must be between 10000000 ~ 99999999

5. Publication type \*

- Peer-reviewed journal or journal
- Report/working paper/grey literature
- Can't tell
- I've already answered the general questions about this paper and would like to code additional outcomes

6. Is the sampling frame clearly defined? \*

*What is the defined population from which the sample is drawn? Do the authors say who could potentially be chosen as sample?*

- Yes
- No
- Can't tell

7. Does the sampling frame include at least 1000 beneficiaries or covers an administrative area larger than a village? \*

- Yes
- No
- Can't tell

8. Is the sample randomly drawn from this sampling frame? \*

- Yes
- No
- Can't tell

## Intervention Arm

Enter this once for each relevant intervention arm

### 9. Type of Intervention \*

- Electrification
- Emissions Trading Scheme

### 10. Timing of the implementation of the intervention \*

- One time
- 1 to 12 months
- 1 to 3 years
- More than 3 years
- Can't tell

### 11. Country, Region, \*

### 12. Target Population Gender \*

*Multiple selection possible*

- Female
- Male

### 13. Target Population \*

- Individual
- Household
- Villages/Communities
- Subnational (district/state/county/etc.)
- National

### 14. Target Population Age \*

- Young Adults (18-35)
- Adults (36-65)
- Elderly (65+)
- Everyone
- Not specified

### 15. Target Population Living Environment \*

- Rural
- Urban
- Can't tell

### 16. Target population specific restrictions (e.g. pregnant women)

17. Nature of comparison group \*

*Method of assignment to treatment or control*

- Randomized experiment
- Quasi-experiment, prospective assignment
- Quasi-experiment, ex-post assignment
- DiD
- Non-random
- Can't tell

18. Nature of comparison Group \*

*Unit of assignment (if applicable)*

- Individual
- Cluster
- Can't tell

19. Outcome timing \*

*Time since the start of the intervention implementation*

- 1 to 3 years
- More than 3 years
- Can't tell

20. Number of relevant outcomes \*

*Either a measure of formal employment (Or general employment including formal employment) or of GHG emissions. If there are more than 4, you will have to start the form again from the beginning for outcomes 5-8*

- 1
  - 2
  - 3
  - 4
  -
- Other

## Outcome

Enter this form once for each relevant outcome

### 21. Description of the outcome \*

*This should be short, and only to differentiate between the other outcomes, e.g. effect of electrification on female employment, or outcome India / outcome South Africa if a paper differentiates outcomes by region.*

### 22. Direction of the Effect \*

- Effect favours treatment
- Effect favours comparison
- Can't tell

### 23. What intervention did the comparison group receive

*If applicable*

- No treatment
- Treatment as usual
- Alternative intervention
- Other
- Can't tell

### 24. Were there any differences in the measurement of this outcome between the treatment group and the comparison group? \*

- Yes
- No
- Can't tell

### 25. Is the Effect statistically significant? \*

- Yes
- No
- Can't tell

### 26. Significance level \*

- <0.10
- <0.05
- <0.01
- <0.001

### 27. Number of observations in the regression \*

The value must be a number

### 28. Number of observations in the regression from treatment group \*

The value must be a number

### 29. Number of observations in the regression from control group \*

The value must be a number

30. Overall Sample Size

The value must be a number

31. Treatment Sample Size

The value must be a number

32. Control Sample Size

The value must be a number

33. Is there clustering \*

*Are observations grouped into clusters, e.g. households clustered in villages. Search for whether authors mention clusters.*

- Yes
- No

34. Number of Clusters \*

The value must be a number

35. Nature of the outcome measure \*

- Continuous
- Discrete
- Dichotomous
- Categorical

36. Econometric model used \*

*E.g. DiD, propensity-score matched OLS, Fixed Effects regression, probit model. Search in the Methods/Methodology section for description*

37. List the significant control variables used, separated by semi-colons \*

*You can also list on what pages you find them. We just need the names, not the values of their coefficients or anything like that.*

38. Number of Control Variables; Explanation. \*

*First type the number; add a semi-colon, then add the explanation. This is the number used in the regression of the outcome we're interested in, including those that aren't noted with coefficients. Also list the pages where you find them. How do you calculate the number?*

39. Is the Intervention Variable a Dummy Variable \*

Yes

No

40. Sample Mean of Y (outcome variable)

The value must be a number

41. Standard Deviation of Y (outcome variable)

The value must be a number

42. Sample Mean of X

*X is the intervention variable, e.g. electricity access*

The value must be a number

43. Standard Deviation of X

*X is the intervention variable, e.g. electricity access*

44. The coefficient of the intervention variable (beta) is \*

Standardized

Non-Standardized

45. The standardized beta-coefficient is \*

The value must be a number

46. The non-standardized beta-coefficient is \*

The value must be a number

47. The variation in the intervention coefficient beta is captured by \*

*Comment from Arne: Look out for standard error and also tick when other statistical information on the intervention coefficient are given*

A standard error

A t-statistic value

A p-value

Other

48. The standard error is for \*

Standardized beta

Non-Standardized beta

49. The value of the standard error of standardized beta is \*

The value must be a number

50. The value of the standard error of non-standardized beta is \*

The value must be a number

51. Are the standard errors clustered \*

Yes

No

52. At what level are the standard errors clustered? \*

53. Are the standard errors robust? \*

Yes

No

54. The t-statistic is for \*

Standardized beta

Non-Standardized beta

55. The value of the t-statistic of standardized beta is \*

The value must be a number

56. The value of the t-statistic of non-standardized beta is \*

The value must be a number

57. The p-value is for \*

Standardized beta

Non-Standardized beta

58. The value of the p-value of standardized beta is \*

The value must be a number

59. The value of the p-value of non-standardized beta is \*

The value must be a number

60. What are the Degrees of Freedom for the regression?

*Only answer if the number is given*

61. What (if any) is the type of  $R^2$  reported? \*

- Non-adjusted  $R^2$
- Adjusted  $R^2$
- Pseudo  $R^2$
- There is no  $R^2$  stated for the regression of interest

62. What is the value of the non-adjusted  $R^2$ ? \*

The value must be a number

63. What is the value of the adjusted  $R^2$ ? \*

The value must be a number

64. What is the value of the pseudo  $R^2$ ? \*

The value must be a number

65. What calculation method was used to get Pseudo  $R^2$ ?

66. Commentary on methods (if multiple methods are selected)

67. Data Type \*

- Panel
- Cross-Section
- Time-Series

68. Is there another relevant outcome you haven't coded yet? \*

*Choose yes if there's another outcome you need to code. Choose no if you've finished this study.*

- Yes
- No