**Technical Appendix**

**Disaggregation of emissions variance**

Table 1 shows the results of a likelihood ratio (LR) test on the emissions data that compares the proposed null two-level model (excluding any predictors) with the equivalent single-level regression.[[1]](#footnote-1) Both of the variance components are significant at the 0.001 level. Strikingly, 91.4% of total variance occurs between countries, while the remaining 8.6% is due to longitudinal variation between countries. This is strong evidence that observations of annual emissions levels from the same country are correlated, suggesting that single-level regression are likely to significantly underestimate standard errors and lead to erroneous conclusions about the drivers of emissions behaviour.

|  |  |
| --- | --- |
| Parameter | Estimate |
| *Fixed Effects* | |
| Constant | 159.41 (51.03)\*\* |
| *Random effects* | |
| Country variance | 386476.10 (44959.00)\*\*\* |
| Country-year variance | 36340.79 (898.47)\*\*\* |
| LR test | 7580.50\*\*\* |

Table 1: Results of the null single and two-level models

Note: Coefficient entries are maximum likelihood estimates with estimates standard errors in parentheses.

\*Significant at 5% (p<0.05)

\*\*Significant at 1% (p<0.01)

\*\*\*Significant at 0.1% (p<0.001)

**Control Variables**

GDP denotes per capita GDP in a given year. By providing a snapshot of average living standards, it provides a good indication of economic development, which is generally positively associated with emissions levels. Per capita GDP data are obtained from the World Bank WDI figures. EXPORTDIV is a continuous variable that measures a country’s reliance on the export of a narrow unprocessed raw material base as a proportion of GDP. Structuralist scholars (e.g. Betsill et al. 2006; Roberts and Parks 2007,2010) argue that advanced economies, which tend to possess diversified export sectors, find it easier to reduce emissions than developing countries, whose incomes are often dependent on high-emissions activity. Export diversification data come from the International Monetary Fund’s Export Diversification Index, which assigns a score from zero (low export diversity) to seven (high export diversity). If legal enthusiasts are right, then the presence of binding mitigation targets under the global climate regime should provide a strong impetus for emissions reduction (von Stein 2008). ANNEX is a binary variable that is coded one for annex parties, which possess quantifiable emissions targets under the KP, and zero for non-annex parties, which have only a (non-binding) political commitment to help global mitigation over the first commitment period. Fossil fuel dependency (FFDEP) denotes the proportion of national income that comes from fossil fuels – either as an export or as a component of production. Some authors (e.g. Sunstein 2009; Vezirgiannidou 2008) have argued that countries with lower fossil fuel dependency should exhibit more mitigation as they stand to gain from their ability to reduce emissions at relatively lower opportunity costs than countries that are highly dependent. FFDEP data comes from the World Bank WDI database records of the percentage of a country’s GDP that is accrued from fossil fuels. Fossil fuel values from the WDI were centred for each year by subtracting the world mean fossil fuel dependency level in a given year from a country’s dependency score in that year to capture the *relative* dependency level in relation to other countries. Technological advances are expected to raise efficiency levels and thus be associated with lower emissions (Stern 2004). TECH captures reductions in emissions levels that are attributable to technological development by measuring the annual level of renewable energy consumption in a country as a percentage of final energy consumption using data from the WDI database. POP captures the demographic demand for emissions by using data from the United Nations Population Division. EMDECOUP captures sustainable policy initiatives that are designed to divorce economic development from emissions. In accordance with the Organisation for Economic Co-operation and Development’s *Indicators to Measure Decoupling of Environmental Pressure from Economic Growth (2005)*, emissions decoupling is measured by calculating the annual ratio of change in CO2 emissions relative to change in GDP. Scores are inverted to aid interpretation so that higher positive values denote more effective decoupling than smaller positive values while negative scores suggest the very opposite of decoupling.

**Testing strategy**

For all models, the subsets ‘i’ and ‘j’ refer to the country-year and country levels.

*Single-level model (OLS)*

This model serves as a reference for evaluating the results of the two-levelled models described below.

*Random intercept model (RIM)*

This equation begins modelling the hierarchical data structure by differentiating between observations from different countries.[[2]](#footnote-2) It is possible to determine whether the results of the OLS are robust to country-level clustering by comparing the coefficient signs, sizes and p-values with the results of the RIM.

where is the overall mean emissions level for all observations across all countries and is the mean emissions level for all observations from country j.

*Random coefficient model (RCM)*

The last stage of the testing strategy is to check for signs of heterogeneous democracy effects by setting up a RCM which allows for the slope of democracy to vary between countries. A significant random effect coefficient, u1j, would indicate that the effect of democracy on emissions varies significantly between countries.

**Additional findings of the RIM**

While not the primary focus of this article, comparing the results of the first two models provides some insights into the effects of the other putative drivers of emissions and their robustness to country-level clustering. In the flat regression, per capita GDP was positively associated with emissions, although the effect was statistically insignificant. In the RIM, both the effect size and significance level undergo striking changes; a one-point increase in per capita GDP is associated with a 1.62 Mt increase in emissions, suggesting that economic development is incompatible with mitigation for the average country. This finding is significant at the 0.05 level. Export diversity, on the other hand, goes from having a highly significant emissions-boosting effect in the first model to exerting only a negligible and statistically insignificant effect on emissions in the new model, suggesting the estimate in the single-level model is not robust to country-level clustering. In both models emissions decoupling has a small positive effect on emissions, although neither effect is statistically significant. Strikingly, as with the democracy coefficient, the coefficients of the four remaining drivers undergo sign changes, signalling the existence of cluster-confounding. Annex membership and fossil fuel diversity go from having inhibitory to emissions-boosting effects. In the first model, annex parties to the KP are found to emit 21MtCO2 more than non-annex parties on average. This situation is reversed in the second model; a given country is likely to emit 162 MtCO2 more emissions when it is an annex party than it would otherwise emit without annex status. Similarly, model one suggests that moving to a country with 1% higher reliance on fossil fuel incomes is associated with a 1.72Mt decline in emissions. However, model two indicates that a 1% increase in fossil fuel dependency within the same country is associated with an 0.95Mt increase in emissions. Yet it is important not to exaggerate the significance of these results as the coefficients of both drivers are statistically insignificant in both models. Technology and population, on the other hand, exhibit cluster confounding in the opposite direction and are of statistical significance. In the first model, a country that derives 1% more of its total energy consumption from renewable technology tends to exhibit a 2.53 Mt increase in emissions. Yet modelling country-level clustering reverses the direction of the effect in accordance with expectations; in the RIM, one-point increase in the ratio of renewable energy consumption within the same country is associated with a 2.22 Mt decline in emissions, suggesting compatibility with mitigation. This finding is significant at the 0.001 level. Similarly, a country with a 1% higher population growth rate is associated with 0.17 MtCO2 higher emissions in the OLS, although the effect is not statistically significant. However, the RIM shows that a one-point increase in population growth rates in the same country is associated with a 2.22 Mt decline in emissions. The finding is significant at the 0.01 level.

## Regional Groupings in the Multilateral Climate Negotiations

*All countries with a (\*) symbol are excluded due to missing data.*

**Annex Parties**

1. **The European Union (EU)**

Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus\*, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the UK

1. **The Umbrella Group (UG)**

Australia, Canada, Iceland, Japan, New Zealand, Norway, the Russian Federation and the US

**Non Annex Parties**

1. **Emerging Economies (BASICs)**

Brazil, South Africa, India and China

1. **Middle-Income Developing Countries (MIDCs)**

Argentina, Bahrain, Bolivia, Botswana, Chile, Colombia, Costa Rica, Cote d'Ivore, Egypt, El Salvador, Georgia, Grenada, Guatemala, Honduras, Indonesia, Israel, Jordan, Kenya, Lebanon, Malaysia, Mongolia, Morocco, Nicaragua, Nigeria, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Sri Lanka, Syria, Thailand, Tunisia, Turkey,\* Uruguay, Venezuela, Viet Nam and Zimbabwe

1. **Least Developed Countries and the Alliance of Small Island States (LDCs and AOSIS)**

Afghanistan\*, Angola\*, Antigua and Barbuda, Bahamas, Bangladesh, Barbados, Belize, Benin, Bhutan, Burkina Faso, Burundi\*, Cambodia, Central African Republic, Cape Verde, Chad, Comoros, Cook Islands, Cuba, Democratic Republic of the Congo\*, Djibouti, Dominica, Dominican Republic\*, Equatorial Guinea\*, Eritrea\*, Ethiopia, Fiji, Gambia, Guinea, Guinea Bissau, Grenada, Guyana, Haiti, Jamaica, Kiribati, Loa People’s Democratic Republic, Lesotho, Liberia\*, Madagascar\*, Malawi, Maldives, Mali, Marshall Islands\*, Mauritania, Mauritius, Micronesia\*, Mozambique, Myanmar, Nauru, Nepal, Niger, Niue\*, Palau\*, Papua New Guinea, Rwanda\*, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines\*, Samoa, Sao Tome and Principe\*, Senegal, Seychelles, Sierra Leone, Singapore\*, Solomon Islands, Somalia\*, South Sudan\*, Sudan, Suriname\*, Timor-Leste\*, Togo\*, Tonga\*, Trinidad and Tobago, Tuvalu\*, United Republic of Tanzania, Uganda, Vanuatu, Yemen and Zambia

1. **Petroleum Exporting Countries (OPEC)**

Algeria, Indonesia, Iran, Kuwait, Libya\*, Nigeria, Qatar, Saudi Arabia, United Arab Emirates and Venezuela

1. **Central Asia, Caucasus, Albania and Moldova (CACAM)**

Albania, Armenia, Azerbaijan, Kazakhstan, Kyrgyzstan\*, Moldova\*, Tajikistan, Turkmenistan and Uzbekistan

1. The LR value is the probability of obtaining the observed values (emissions data for the sample) if that model were true. [↑](#footnote-ref-1)
2. All multilevel models described in this article were fitted using Stata’s xtmixed command. [↑](#footnote-ref-2)