Data Storm Chasers: Proposal

1 Background

Many lower middle income countries confront a significant challenge posed by climate change (CC): achieving inclusive economic development and improved quality of life in the face of increasingly challenging environmental conditions to which they must adapt. There are likely to be many economic impacts of CC in a country such as India, with obvious ramifications for health and wellbeing. Flooding can cause loss of property and livelihoods; droughts can disrupt crop cycles; extreme conditions can cause reduce daytime labour hours, and so on. But just as important will be less-direct economic effects of CC via public policy. Adapting to CC may involve investment in new technologies, social support, fiscal measures and changes to international trading arrangements, etc - all of which are likely to have ramifications for quality of life.

Changes in trading arrangements point us to another set of indirect impacts. The effects of CC in any one country - and especially an LMIC - will be partly shaped by the policies of other countries, and how they respond to CC. For example, the growing commitment of advanced economies to decarbonise may present LMICs with both challenges and opportunities. On the one hand, the transition to electric vehicles in advanced economies is likely to result in vehicle markets in some LMICs being flooded with cheap but dirty internal combustion vehicles. On the other hand, a country such as India may have lucrative opportunities to invest in low-carbon technologies, bringing economic and thence health benefits

2 Our Prototype

For the purposes of transparency and robustness, we have used the tried, tested and trusted approach of linear regression modelling, as opposed to more "black-box" ML techniques. We use the following econometric model to estimate climate effects on the economies of LMICs and how they can be moderated by importing from a G7 country.

$$Y_{it} = \alpha_i + h(t) + \beta_1 X_{it} + \beta_2 q_i X_{it} + \beta_3 q_{it} + e_{it}$$

where Y_{it} is an economic (or health) variable, α_i is a country random (grouping) effect, $h(t)$ is a temporal trend, X_{it} is a climate variable, q_{it} is a dummy indicating imports from any G7 country, the β values are fitted coefficients and e_{it} are residuals. In the platform, the user can pick their economic variable and one or more climate variables from a list. Our prototype is readily extended to allow causal inference by obtaining proportions of explained variance from models of health against (1) economy, (2) climate and (3) both economy and climate, and obtaining a mediated R^2 estimate as $(1) - (3) + (2)$ (Fairchild et al., 2009).

3 Plans for Future Work

Guided by the global nature of our analyses, our one-year project aims to extend the functionality and develop the accessibility of our prototype, guided by potential users from LMIC

contexts and in liaison with UK policymakers. This section details (1) the model, (2) the platform and the (3) practicalities of project management.

3.1 The Model

In order to uncover more nuanced health impacts of CC via the economy, we will extend the econometric model from Section 2 by allowing the incorporation of (i) non-local climate and economy effects, (ii) both recent and forecast climate variables and (iii) micro-level economic and climate data. Formally, let $\mathbf{C}$ denote a vector of true climate parameters (e.g. temperature expectations), and put $\mathbf{c}_{i\tau}$ as $\mathbf{C}$'s realisation at time τ and region i and $\mathbf{b}_{i\tau}$ as $\mathbf{C}$'s forecasts for spatial region i, starting at time τ. Given a vector of economic or health outcomes Y (e.g, unemployment, respiratory illnesses), we consider the effects of $\mathbf{C}$ on Y at time period τ and region i as

$$Y_{i\tau}(\mathbf{C}, \mathbf{E}) = Y[local(\mathbf{C}, \mathbf{E}); remote(\mathbf{C})]$$

where $local(\cdot); remote(\cdot)$ are are functions dependent on local and remote weather $(\mathbf{c}(\mathbf{C}), \mathbf{b}(\mathbf{C}))$, resp., and local economic $(\mathbf{E})$ input data. As a linear model we then have

$$Y_{i\tau} = \alpha_i + h(\tau) + \underbrace{\sum_{t=\tau_0}^{\tau-1} \beta_{1t}\mathbf{c}_{it} + \beta_2 \mathbf{b}_{i\tau} + \beta_3 \mathbf{E}}_{\text{local}} + \underbrace{\sum_{i_n \neq i} Q_{i_n} \sum_{j \in \{i, i_n\}} \left[\beta_{4j}\mathbf{c}_j^* + \beta_{5j}\mathbf{b}_{j\tau}\right]}_{\text{remote}} + e_{i\tau}$$

where $\mathbf{c}_j^*$ represents a time trend of climate parameters at region i, α_i is the country effects, $h_i(\tau)$ represents non-linear factors, Q_{i_n} is a bilateral continuous variable (Climate Finance "flows", mass migration patterns, trade) and $e_{i\tau}$ is the error. In parallel to the proposed linear model we also propose to: better estimate the uncertainty by bootstrapping and implementing causal ML frameworks; and allow quantile regression so distributional effects can be explored. We will pre-register our study and post plans on Open Science Framework.

3.2 The Platform

On top of allowing the user to specify the econometric model, our platform will do the following.

- Display a global map with basic climate change layers built-in and easily updated.
- Enable spatial extent to be manipulated graphically while re-running analyses.
- Allow users to upload directed graphs for remote causality (e.g. climate finance flows).
- Provide flow diagrams (e.g. Sankey) to visualise bilateral influences.
- Offer a grid-level GDP indicator using night-light measurements (Martinez, 2022) and WMO data for climate (with spatial autocorrelation in the error term as necessary).
- Display timeline of data sets with feasible lags and any misalignments.
- Connect to data sources using API and simple gamification.

The tool will be fully documented in a user guide detailing the methods used, for complete transparency.

3.3 Practicalities

We envisage a co-designed research and development programme that can be further developed in a future project. Our one-year pilot scheme brings policymakers from two LMICs together with those from the UK international development office to shape a tool for understanding and tackling international climate - economy - health links. The core activities are as follows:

- Run two in-person workshops to engage policymakers: in Bangladesh and Argentina, each with guests from the UK civil service (e.g. DfID).
- Run online steering group meetings with team and policymakers
- Dissemination: a research paper, a policy brief, blogging (*The Conversation*).
- Seeking funding for extending the programme.

Work Allocation: The allocation of responsibilities is as follows:

- Liaise with policymakers and dissemination work; **Liaison RA** (to recruit)
- Line-manage Liaison RA: **Richard Gunton** Co-I; Liaison support:**Sakib Anwar** Co-I
- Develop coding and GUI: **Postdoc RA** (to recruit)
- Line-manage Postdoc RA: **Jorge Bruno**, Co-I; Coding support: **Paolo Pareti**, Co-I
- Oversee whole project and timekeeping: **Wioleta Kijewska**, PI.

The PI and Co-Is will give around 4 hours/week. All team members will contribute to writing a research paper. An outline schedule is shown in the Gantt chart below.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Liaise with LMIC policymakers												
LMIC workshops												
Extend functionality												
Improve GUI												
Write paper												

Equality, Diversity and Inclusion: Our diverse team (5 nationalities) will recruit RAs on the basis of merit, pressing our HR department for fair salaries. The project begins with the Liaison RA engaging with policymaking stakeholders to form an international steering group. The Postdoc RA joins after 3 months so the LMIC steering group can influence the appointment and the postholder's agenda.

References

1. Fairchild, A. J., MacKinnon, D. P., Taborga, M. P., & Taylor, A. B. (2009). R 2 effect-size measures for mediation analysis. *Behavior research methods*, *41*(2), 486–498.
2. Martinez, L. R. (2022). How much should we trust the dictator's GDP growth estimates? *Journal of Political Economy*, *130*(10), 2731–2769.