Jonas Vestby, Andreas Forø Tollefsen, Halvard Buhaug

Climate and International Migration Flows: A Sensitivity Analysis of Gravity Model Specifications

Deliverable 2.6

QuantMig has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 870299.

History of changes

Version	Date	Changes
1.0	14 June 2022	Issued for Consortium Review
1.1	30 June 2022	First version submitted as official deliverable to the EC

Suggested citation

Vestby J, Tollefsen A F, and Buhaug H (2022) Climate and International Migration Flows: A Sensitivity Analysis of Gravity Model Specifications. QuantMig Project Deliverable D2.6. Oslo: PRIO.

Dissemination level

PU Public

Acknowledgments

We are grateful to Jakub Bijak, Michael Boissonneault, Marta Bivand Erdal, and Vally Koubi for reviewing the internal draft. This document reflects the authors' view, and the Research Executive Agency of the European Commission are not responsible for any use that may be made of the information it contains.

Cover photo: iStockphoto.com/Guenter Guni

Table of Contents

List of Tables

List of Figures

1. Introduction

Over the past decade, academics and policy actors alike have raised awareness to climate change as a potential major driver of human mobility (e.g., Foresight, 2011; Hunter & Simon, 2022; IPCC, 2022; McLeman, 2014; Piguet, 2010; Rigaud et al., 2018). Despite undisputable connections between extreme weather events and short-term displacement, empirical evidence linking broader categories of climatic conditions with migratory outcomes remains mixed (Beine & Jeusette, 2021; Cattaneo et al., 2019; Ferris, 2020; Hoffmann et al., 2020), contrasting the widespread but contentious notion of "climate refugees" (Boas et al., 2019).

There are probably both substantive and methodological reasons for the diverging results in the literature. Theoretically, the influence of shifting climatological and environmental conditions on people's motivation to move is thought to be highly dependent on both individual- and society-level contextual factors, many of which are hard to quantify and incorporate in statistical models. For example, whereas adverse climatic changes generally constitute a greater challenge to the livelihoods and wellbeing of marginalized populations, thus potentially increasing *motivation* to relocate, privileged people more often possess skills, networks, and resources that increase the *opportunity* for migration. Many of the people most vulnerable to environmental hazards, for whom mobility could be an especially effective coping strategy, simply lack the means to move in response to a household shock (Adams, 2016; Black, Bennett, et al., 2011). Although several notable conceptual works (e.g., Adger et al., 2015; Black, Adger, et al., 2011; Gemenne et al., 2014) have greatly influenced academic thinking on the topic, there still exists no single, unified, theoretical framework that adequately captures the complex, context-dependent climate-migration relationship (Cattaneo et al., 2019).

Aside from theoretical opacity, lack of convergence of findings in the quantitative literature is also likely due to heterogeneity in operationalization of core concepts. Considering the **outcome** variable, migration can be broken down into distinct analytical categories reflecting degrees of human agency (sometimes dichotomized as voluntary vs. forced migration; the latter encompassing refugees, asylum seekers, and displaced persons), purpose (protection; family reunification; education / labor opportunities; return migration), destination (internal vs. international; rural-rural vs. rural-urban migration), and various temporal characteristics (e.g., temporary vs. permanent migration). The measurement of migration likewise varies by capturing (inter alia) migrant stocks, migration inflows, migration outflows, bilateral migration flows, net migration, or urbanization rate.

Regarding the **treatment** variable of interest – climate – a variety of theoretically plausible indicators have been explored, including levels, first-difference changes, and deviations from long-term means in temperatures, evapotranspiration, and precipitation. Several studies also explore the occurrence of extreme weather events, e.g., floods and droughts.

The chosen **level of analysis** in the literature varies between individual-level, survey-based analysis and aggregate studies of migration between subnational entities (administrative units of grid cells) or between countries, relying on data observed at monthly, yearly, 5-yearly, or decadal time intervals. Relatedly, studies vary in terms of geographic scope; while some are global or cover countries across all world regions, other rely on data from specific regions or single countries.

Lastly, heterogeneity in findings might be partly attributable to differences in **estimation method**. The most common approach is econometric regression analysis aiming at estimating causal effects of climatic exposure (either directly or indirectly, considering the global average effect and exploring how it varies across contexts), although a few recent studies focus on building models with high predictive accuracy and explore the predictive contribution of adding climate information (e.g., Clement et al., 2021; Kiossou et al., 2020; Schutte et al., 2021). Each approach comes with a distinct set of assumptions that must be considered when inferring implications of the results for the posed research question.

1.1 Our Approach: Sensitivity Analysis

Considering the plethora of reasonable empirical approaches to study whether and how climate influences human mobility, we cannot provide a coherent, comprehensive, and inclusive analysis that covers all or many of the alternative analytical choices outlined above here. Instead, we zoom in on a specific mobility category – international migration – and consider a dominant modelling approach, namely the **gravity model**, to explore the sensitivity of the estimated effect of climatic exposure across specifications.[1] Although most migration occurs within state borders (Rigaud et al., 2018), climate change and associated physical process are expected to amplify international as well as internal migration in the future (IPCC, 2022), increasing political salience around human habitability and mobility in vulnerable regions of the world. Importantly, data quality is judged to be considerably higher for international migration, although we stress that results reported here do not necessarily speak to the robustness of climatic drivers on internal migration.

A large number of published articles have attempted to identify a causal effect of a climatic treatment on international migration flows using the gravity model (e.g., Abel et al., 2019; Backhaus et al., 2015; Beine & Parsons, 2015; Cai et al., 2016; Coniglio & Pesce, 2015). They report varyingly that temperature and/or precipitation either is not associated with migration outcomes, is positively and significantly associated with migration, or is positively and significantly affecting migration under particular conditions (e.g., in interaction with agricultural dependencies).

We employ **a sensitivity analysis approach** to explore and discuss various estimation choices, such as the log- linear vs Poisson specification, various fixed effects specifications (in origin, destination, time, paired), the variance-covariance estimator (robust and clustered standard errors), spatial dependencies, and temporal autocorrelation. We also consider recent developments in the econometrics of causal analysis using fixed effects and the extent to which these ideas travel to the directed dyadic case (e.g., Baker et al., 2022; Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfœuille, 2020; Jakiela, 2021). While these tests jointly capture a broad set of alternative, seemingly plausible research design issues, estimated results might also vary across different subgroups of international migration and across different (unobserved) strata of countries, which we are unable to evaluate in detail here.

Most empirical studies of international migration rely on various subsets of global migration flows, typically restricted to inflows to OECD countries. **The migration flow data is itself an estimate** based on migration stock data. The most common approach is to calculate the first-differenced stock (with drop negative or reverse negative flows). Azose & Raftery (2019) argue that these estimations are estimations of the lower bound, and develop a "Pseudo-Bayesian" approach to estimate actual flows. The differences between the Pseudo-Bayesian approach and the minimization approaches are substantial (Abel & Cohen, 2019) and the Pseudo-Bayesian approach

[1] The decision to assess the consistency and rigor of a climate effect on international migration across a broad set of specifications, as opposed to seeking to estimate such an effect directly and via armed conflict through a specific model design, explains the change of the title for this report (originally, the deliverable was titled "Climate change, armed conflict, and migration").

leads to better model fits compared to "bronze standard data" [2] (Abel & Cohen, 2019; Azose & Raftery, 2019). We explore through sensitivity analysis how estimates change across different, commonly applied methods for estimating migration flow, for different spatial subsets, and when accounting for the measurement uncertainty reported by Azose & Raftery (2019).

We find that estimation of individual parameter effects in count, dyadic, fixed effects models of international migration often relies on **heroic assumptions** (LeSage & Pace, 2008; Sellner et al., 2013); no published article sufficiently accounts for all modeling challenges, and results are quite sensitive to different approaches and input data. While it is tempting to build global systems models of bilateral migration where everything is included, we should also acknowledge that isolating any part of that system is bordering on impossible. In a recent critique of the gravity model approach to international migration, Beyer, Schewe, & Lotze-Campen (2022) argue that these models are poorly suited to predict future trends in international migration dynamics and to capture temporal dynamics in bilateral migration flows. Instead, these models "describe spatial patterns of international migration very well" (p.1). Models aimed at predicting migration flows indicate that climatic variables are very poor predictors (Kiossou et al., 2020). This suggests that any causal climate effect is likely to be highly context-sensitive and not the main driving force of bilateral migration flow, which has important implications for how causal models should be specified.

Not only does the estimation rely on heroic assumptions, but it also relies on a set of assumptions we are not able to simultaneously account for in single models. For instance, there is currently no fixed effects Poisson estimator able to account for spatial autocorrelation in dyadic data. Sellner, Fischer & Koch (2013) describe a spatial autoregressive gravity Poisson model, and Glaser, Jung & Schweikert (2022) describe fixed effects estimation with spatially dependent count data. Griffith, Chun & Li (2019) describe an approach using eigenvector spatial filtering to account for what they call network autocorrelation, but we have been unable to implement their algorithm on international migration data[3]. In the absence of a principled approach to accounting for spatio-temporal dependencies, we explore how estimates are sensitive to the inclusion of covariates that capture spatio-temporal (or network) relations.

2. The Gravity Model of International Migration

The most common modelling approach to estimate the climate effects on international migration flow globally is to employ the gravity model. The model was first described by Newton's *Mathematical Principles of Natural Philosophy* in 1687. George K. Zipf argued that commuting flows in the US were proportionate to the ratio of the population sizes in the origin and destination communities on the transportation distance, calling it the $\frac{P_1 P_2}{D}$ hypothesis (Zipf, 1946). The gravity model of trade was discussed by Walter Isard (1954) and later also for migration flows (Isard, 1966). In its simplest form, the gravity model can be written as

[2] Azose and Raftery (2019) define "bronze standard" as data "somewhat inferior to "gold standard" data but still sufficient for validation" (p. 121). The "bronze standard" datasets they use are flow estimates from the IMEM project (migration flows among 31 countries in the European Union and the European Free Trade Association) (Raymer et al., 2013) and migration flows to and from the OECD countries (OECD, 2015).

[3] The issue is computational, as both the stepwise regression and the need to create variables based on a Kronecker-product of the fixed effects does not scale well to situations with a high number of fixed effects.

$$F_{ij} = \frac{P_i^{\alpha} P_j^{\beta}}{f(r_{ij})}, i \neq j,$$

where F_{ij} is the directed flow from i to j, $P_i^{\alpha} P_j^{\beta}$ are the positive factors in the origin and destination creating a gravitational pull (such as population size in i and j) with adjustable exponents, while $f(r_{ij})$ is a resistance term capturing factors that limit bilateral flow (such as the distance between i and j). By log-transforming both sides (and acknowledging model errors), we can get a linear equation

$$\log(F_{ij}) = \alpha \log(P_i) + \beta \log(P_j) - \log(f(r_{ij})) + e_{ij}, i \neq j.$$

While Zipf's original gravity model was an empirical claim about the relationship between migration flow, population sizes, and the geographical distance between dyads, the more general form stated above can use multiple covariates with their own parameters. The core idea is to use a log-log directed dyad elasticity specification with information about both the origin, the destination, and factors related to their relationship that could affect the size of migration flow. Indeed, Zipf (1946) added more covariates by showing how the mode of transportation (highway, railway, and airway) did modify the relationship he was interested in (it was less clear for air-travel than highway travel).

The error term e_{ij} in the stochastic gravity model turns out to complicate things vastly, as we can only expect to get consistent estimates of model parameters under very strict assumptions. Particularly, the errors must be independent and homoscedastic. For international migration, they are neither. It can be useful to think theoretically about why that is the case.

In Newton's gravity model, the outcomes of the gravitational pulls are just the combined vectors, and the outcomes do not affect parameters in the system (e.g., the mass of an object does not change when getting closer to another object). In migration, the gravitational flows are aggregated descriptions of individual behavior, and the outcome for an individual can only be discrete. I.e., either you stay, or you move.

People in i that potentially could join the flow F_{ij1} can also be part of any other flow $F_{ij}, j \neq 1$ whose resistance term $f(r_{ij})$ is low enough. Changes in positive factors anywhere in the system where R_{ij} is low enough can therefore affect the propensity of a person in that part of the system for joining one flow rather than another. This suggests that observations cannot be considered spatially independent (Fischer & Griffith, 2008; LeSage & Pace, 2008; Sellner et al., 2013).

Furthermore, both gravitational pulls and resistances are probably a complex function of past flows F_{ij}, not only within a specific dyad but probably also informed by other dyads (but probably most prominently dyads involving either i or j). Carling, Czaika & Erdal (2020) argue that one of the defining characteristics of migration flow is that it has temporal autocorrelations due to, e.g., *chain migration*. Furthermore, restrictive migration policies might come as a function of past migration and is known to have a deterrent effect on migrant inflows (Carling et al., 2020). However, such policies might divert the flows to other countries instead – showcasing one of many possible interactions and dependencies between temporal and spatial processes. The temporal aspects of real-world migration has been explored in the gravity model setting through the notion of network autocorrelation (Chun, 2008; Chun & Griffith, 2011; Griffith et al., 2019). However, it is still fair to say that the interest in temporal patterns and dependencies in migration (and their consequences) has been less pronounced than the interest in modelling spatial patterns (Beyer et al., 2022).

Heteroscedasticity can occur for many reasons, but the main reason for it to occur in international migration is that migration flow counts are heavy tailed, even after log-transformation (Bijak, 2011). This tends to result in models that predict better in relative terms when the flow is small than when it is large (where it tends to underpredict). The outliers of this stochastic model then must be accounted for using covariates, which is difficult to do because the number of observations with large migration flows are small and reasons for very large population movements possibly are quite idiosyncratic. Controlling for population - or using per capita flows or a population offset - and other important explanatory factors of migration such as distance could possibly reduce heteroscedasticity, but not alleviate it entirely.

2.1 Alternatives to the Gravity Model

An alternative to estimating the linear causal effect of covariates on international migration is to evaluate the predictive performance of models with and without covariate information. A benefit with this approach is that we are attempting something simpler than isolating the causal effect of a single variable. Rather, we ask whether knowing covariate information enables us to predict subsequent outcomes. Prediction comes with its own set of issues, such as proper measurement of predictive performance and avoid using post hoc information in model training. However, these are arguably easier problems than causal inference. Prediction is useful even if we assume that we have correctly estimated the causal effect, if the inclusion of the variable does not improve predictions of the outcome, then either the effect has changed over time or it is not substantial. We should generally demand that variables both contribute to predicting the outcome and showcase the expected causal effect in attempts at isolating it (Ward et al., 2010).

Recent research show that machine learned models (e.g., fully connected artificial neural networks and Long Short-Term Memory models) predict international migration better than the log-linear gravity model with origin, destination, and year fixed effects (Golenvaux et al., 2020). More generally, there is large development and much creativity in modelling approaches that aim at predicting migration outcomes (see, e.g., Lenormand et al., 2016; Robinson & Dilkina, 2018; Simini et al., 2012).

One general point worth highlighting is that there has been interesting developments in how to measure predictive performance, with a harmonic mean ($^{2ab}/_{a+b}$) alternative in the Common Part of Commuters (CPC)[4] devised for migration flow networks (Lenormand et al., 2016). Errors measured with the harmonic mean tends to be less sensitive to outliers and more sensitive to small values. The CPC is also a coefficient between 0 and 1, meaning that performances can be compared across different data-samples (just as R^2 is a relative alternative (and proportional) to the mean square error). However, in our opinion, the most interesting part of the CPC is that it can easily be extended to measure how well a prediction captures the topological structure of a network ("Common Part of Links"), i.e., the extent to which the prediction correctly predicts 0 flows and >0 flows[5], and the observed commuting *distance* patterns ("Common Part of Commuters According to the Distance"). Due to the zero-inflated and heavy tailed distributions in international migration flow data, we would also think that probabilistic approaches would be useful (e.g., Czado et al., 2009).

It can be shown that different types of models are better optimized towards different types of prediction errors, and a sole focus on the mean square errors might be misplaced (Golenvaux et al., 2020; Robinson & Dilkina, 2018). The focus should probably also be different across applications and

[4] More precisely, CPC is a "Sørensen Index" for network flow matrices, which measures the similarity of two matrices with positive numbers. $CPC = {2\sum_{i,j=1}^{n}\min(T_{ij},\widetilde{T_{ij}})}\Big/{\sum_{i,j=1}^{n}T_{ij} + \sum_{i,j=1}^{n}\widetilde{T_{ij}}}$

[5] A Sørensen Index with Boolean values is a F1-score, meaning that the Common Part of Links is a F1-score for the classification of >0 flow links.

levels of analysis. E.g., when predicting international migration, we might be more interested in the large flows or large changes in flows ("outliers") than small and stable flows.

While parameter inference is difficult in machine-learned models, there are some options available to explore variable contributions and effects (Molnar, 2019). This is done in Kiossou et al. (2020), who explore feature importance and Partial dependence plots[6] (PDP) for input-variables in an artificial neural network (ANN) model of international migration flow. They find that drought (SPEI) and disaster data, when added to such prediction models, contribute little to improve predictive performance. Exploring the PDP, they find that if anything, drought events are predicted in the model to decrease international migration flow (although the effect is negligible). Their ANN model is substantially better at predicting migration flow than the Poisson gravity model.

3. Modeling Decisions Explored in Sensitivity Analysis

We have identified several issues regarding the international migration flow data, causal identification with panel data using fixed effects, and practical estimation that has been argued in previous literature is likely to affect parameter estimates and standard errors. In the following, we describe these issues and how we explore these issues using sensitivity analysis.

The sensitivity analysis is also motivated by the fact that published literature that aims to estimate the effect of climate variability on international migration flow have made slightly different choices. One aim with the sensitivity analysis is therefore to explore whether these choices can explain the different results. Table 1 provides an overview of the main specifications used in this literature.

Table 1. Main specifications in gravity models of climate variability and international migration

Article	Outcome	Spatial Scope	FE type	Climate variable	Method	Std. err.
Backhaus et al. 2015	log(flow + 1)	19 OECD	O-D + Y	T + P	OLS	HCE
Beine and Parsons 2015	Per cap. mig. rate	Global (Özden et al. 2011)	O + Y	ΔT + ΔP	PPML Poisson	HCE
Abel et al. 2019	log(flow | flow > 0)	Asylum-seeking flow (UNHCR)	No	SPEI-12	Bivariate probit + censored outcome GMM	O+D
Coniglio and Pesce 2015	Migration flow	29 OECD	O + D-Y	ΔT + ΔP	PPML Poisson	D
Cai et al. 2016	log(per cap. mig. rate + 1)	42 mostly OECD	O-D + ttO + ttD	T + P	OLS	O / O+D

Note: FE type can be (O)rigin, (D)estination, (Y)ear. O-D is the paired version. ttO is a linear time-trend in the origin. Climate variable can be (T)emperature, (P)recipitation, or climate anomalies

[6] See https://christophm.github.io/interpretable-ml-book/pdp.html.

e.g., ΔT (somewhat differently calculated in the articles) and SPEI. Standard errors can be heteroscedasticity consistent/"robust" (HCE), clustered on (O)rigin/(D)estination or both (O+D).

3.1 Data on International Migration Flows

We do not have a complete set of data on international migration flows. Indeed, we only have somewhat reliable estimates of migrant inflows for a small number of destination countries (mainly within the OECD). Some approaches choose to use this data to estimate effects. However, this limits the scope to emigration to rich OECD countries and ignore changes in pull factors from other possible destinations. We do have mostly reliable global migrant stock data through population censuses asking people about their country of birth (and the current country of residence). Some approaches proxy bilateral migration flow through a first-differencing approach to the stock data (either censoring negative estimates at zero or reversing the negative estimates). Abel (2013) and Abel & Cohen (2019) compare these approaches to other possible ways to estimate bilateral migration flow, and concludes that the current best approach is the Pseudo-Bayesian approach developed by Azose & Raftery (2019) through validating the estimates against “bronze standard” flow data.

Published research attempting to quantify the influence of climate on international migration commonly estimate this effect using the first difference of migration stocks, and often using only inflows to (various subsets of) OECD countries. Beine & Parsons (2015) note that the first difference approach they use can be defended if return migration constitutes a small share of total migration flow. However, Azose & Raftery (2019) reveal that return migration likely takes up a large share of migration flow, with the first difference reverse negative approach to accounting for return migration being a low estimate. When return migration is expected to be large, it is also problematic to use net migration as the outcome variable, since it is not possible to differentiate between a case where emigration goes down and a case where emigration stays at the same level, but with increased returns. With the exception of Beine & Parsons (2015), who estimate models using both the drop negative and the reverse negative differenced stock approach, none of the published articles that estimate the climate effects on migration flows have questioned whether measurement errors in the outcome affects our ability to estimate this effect.

The Pseudo-Bayesian demographic accounting approach comes with substantial measurement uncertainty. Azose & Raftery (2019) judge the uncertainty bounds of their estimates to be “roughly correspond[ing] to an 80% confidence interval whose lower and upper bounds differ by a factor of 5.4” (SI, p.12) (assuming that patterns from inflows to Europe can be generalized to the rest of the world). Measurement error, like any missing data issue, is benign if it is missing completely at random (and accounted for when estimating standard errors) but can introduce biases if measurement errors are systematic. Since we know that migration inflow statistics are more reliable for OECD countries than other destinations, we also know that the measurement errors are not missing completely at random.

With the data from Abel & Cohen (2019) and estimates of measurement uncertainty from Azose & Raftery (2019), we are now able to conduct a broader sensitivity analysis of the climate estimates as a function of how migration flows are estimated and which subsets of country pairs and years are included. We estimate models using all six approaches described in Abel & Cohen (2019): Stock differencing drop negative, stock differencing reverse negative, migration rates, demographic accounting closed minimization, demographic accounting open minimization, and pseudo Bayesian demographic accounting. We estimate the uncertainty of the effect of temperature and precipitation on international migration flow when accounting for the measurement uncertainty described in Azose & Raftery (2019) assuming these errors are missing completely at random. And we calculate how sensitive estimates are to the choice of spatial scope: All directed dyads, only OECD inflow dyads, and only non-OECD inflow dyads.

3.2 Causal Identification in the Gravity Model Using Fixed Effects

The causal identification design that motivates the use of a fixed effects gravity model to estimate the effect of temperature and precipitation on international migration flows is the difference-in-difference (DiD) design. In the canonical DiD case, we have two periods (pre- and post-treatment) and two groups (those that receive treatment and those that do not). First, we calculate the difference in outcome for each group across time (post – pre). Then we calculate the difference in the two differences in outcomes across each group (treated - control)[7]. Assuming parallel trends and non-interference, it can be shown that we can estimate the average treatment effect on the treated (ATT) using this design (Angrist & Pischke, 2008; Baker et al., 2022). We do not estimate the Average Treatment Effect (i.e., the expected treatment effect for the whole population) because we are comparing the treated units with a counter-factual that the treated units did not get treated. We therefore only estimate the expected treatment effect for the units that got treatment using DiD.

The DiD design can be generalized to multiple units and multiple time periods using the two-way fixed effects form (henceforth TWFE). However, when treatments are "staggered" (i.e., not applied at the same time), it can be shown that DiD estimates are likely biased when effects vary over time for each unit (Baker et al., 2022). The literature does contain some suggestions for how to diagnose this issue (Jakiela, 2021)[8] and possible solutions (e.g., Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun & Abraham, 2021). The implications of these insights are probably widespread across many social-science fields applying fixed effects estimators since in most studies of society we should expect effects to vary over time. Here, we approach the issue by attempting to diagnose whether the basic fixed effects approach is likely to bias our estimates.

The fixed effects gravity model is remote from the idealized DiD model with multiple units and time periods from which we can reasonably derive the ATT. The most obvious challenge is that our units of observations are directed flows between an origin and a destination, while treatments are given to the spatial units (that are both origins and destinations at the same time). Treatments on units will very likely interfere with other units, violating the non-interference assumption. E.g., treatments that cause migration flow from A to B also influence return migration from B to A or make flow from B to C more likely (when those moving from A to B are in transition to C).

The assumption that observations in a gravity model are mutually independent has been called "heroic" (LeSage & Pace, 2008; Sellner et al., 2013). Practically, it has been shown that both coefficients and standard errors are sensitive to accounting for spatial autoregression, e.g., in the case of US State migration flows (LeSage & Pace, 2008). None of the models estimating the effects of climatic factors attempt to model spatial dependencies other than through the inclusion of covariates.

Ignoring treatment-interference, our treatment (5-year averages in climatic conditions) is "given" continuously to all origins and destinations but with varying absolutes and variance. The staggered DiD TWFE effect estimate is a "weighted average of all possible two-group/two-period DiD estimators in the data" (Goodman-Bacon, 2021). This means that a naïve implementation of

[7] https://andrewcbaker.netlify.app/2019/09/25/difference-in-differences-methodology/ provides a visual explanation of this.

[8] We draw extensively from the code and discussion by Andrew Heiss in his blog post on the subject: https://www.andrewheiss.com/blog/2021/08/25/twfe-diagnostics/

the TWFE leads to comparisons not only between treated and untreated units, but between treated and already-treated units. Goodman-Bacon (2021) shows that "when already-treated units act as controls, *changes* in their treatment effects over time get subtracted from the DiD estimate" (p.3). It is possible to explore both treatment homogeneity and when comparisons between treated and already-treated occurs (through the notion of negative treatment weights) (Jakiela, 2021). While we diagnose these issues, we do not know whether there are complications to the fact that we are working with dyadic data and have not seen any direct discussions of these issues in the gravity case.

Another entry-point to motivate the gravity fixed effects estimator is to think of it as a theoretical model, and deal with identification issues through considering conditional independence assumption and potential omitted variable bias (Angrist & Pischke, 2008). Variables that affect the outcome and correlate with the treatment are confounders that, if not controlled for, lead to omitted variable bias. Here, the argument for using origin and destination fixed effects is to account for (fixed across time) "multilateral resistance terms" (Anderson & van Wincoop, 2003). It has been further argued that using destination-time fixed effects "will completely account for any multilateral resistances in receiving countries" (Beine & Parsons, 2015, p. 732).

In the literature, these discussions commonly are more general modelling considerations, instead of discussing the potential for confounding for any specific treatment. It is quite possible, however, that absolute temperature and precipitation are correlated with the ease of international migration (i.e., "multilateral resistance"). It is less obvious to us that a country-standardized treatment such as the standardized precipitation-evapotranspiration index (SPEI) would be correlated. It might, then, be more reasonable to not apply fixed effects when estimating effects of SPEI like in Abel et al. (2019) (also see Vestby, 2019), than when estimating effects of temperature and precipitation.

While motivating the causal inference with the DiD design provides us with leverage over some potential issues with our causal estimator, and an idea of what kind of effect we are estimating, the standard DiD TWFE literature does not provide a clear recommendation for which fixed effects we need to add in the directed dyad setting. Should they follow the units that treatment is being applied to, or the units of observation? And which units are treatment applied to? If we follow the treatment on the origin, then if we want to estimate the effect of temperature in the origin, perhaps origin and year fixed effects would be sufficient? However, if temperature in the origin partially correlates with the multilateral resistance terms in destinations when origin-averages are controlled for (perhaps due to shocks in destinations correlated with origin temperatures making travel more difficult), then maybe we would need to add destination (or destination-year) fixed effects too. If we follow the units of observation, we will argue that we need origin-destination pair plus year fixed effects (or alternatively origin-destination plus destination-year pairs if we are worried about multilateral resistances in destinations varying over time).

We are unable to provide a principled suggestion as to which of these setups are most correct. We do observe that the literature also does not agree on the specification. Backhaus et al. (2015) use the origin-destination pair plus year fixed effects, Beine & Parsons (2015) use origin and destination fixed effects, Coniglio & Pesce (2015) use origin plus destination-year fixed effects, while Cai et al. (2016) use origin-destination pair fixed effects plus a control for time-trends in origin and destination. Abel et al. (2019) do not add fixed effects, but their climatic treatment variable is SPEI[9].

[9] It should be noted that Abel et al. (2019) estimate three simultaneous equations to explore the mediation effect of SPEI on international migration flow through conflict occurrence. This approach does not attempt to estimate the ATT or the ATE, but the Local Average Treatment Effect (LATE). The LATE is the average

Since treatments are given to countries (or they are averaged at the country level) and not to the directed dyads, we will need to cluster our standard errors. It is not clear on what we should cluster, however, since it is not entirely clear at what level treatments are "given". The treatment is given to both the origin and to the destination, and climatic conditions (temperature more than precipitation) can be spatially dependent across vast distances (possibly covering multiple countries or even the whole globe), so that clustering on both origin, destination, and time are all reasonable approaches. A more principled approach to standard errors in dyadic data is the Dyadic-Robust t-statistic, or to estimate standard errors through bootstrapping (Aronow et al., 2017; Fafchamps & Gubert, 2007; Tabord-Meehan, 2019). Bergé (2018), considering a gravity model of international trade, use two-way clustering of standard errors on the origin and destination. Both Backhaus et al. (2015) and Beine & Parsons (2015) report heteroscedasticity-consistent (HCE or "robust") standard errors. Coniglio & Pesce (2015) specify standard errors clustered on destination. Cai et al. (2016) include standard errors clustered on origin in the main specification and clustered on the origin-destination pair as robustness check. Abel et al. (2019) cluster on origin and destination.

3.3 Practical Estimation

While the models we draw our intuition from can be estimated using a simple OLS regression, the practical issues of calculating the treatment effect in the gravity-setting quickly escalates. First, international migration flow is count data, and in the gravity model, we are supposed to log-transform both the outcome and the predictors. Both zero-flow observations and estimation issues concerning how the outcome is log-transformed can lead to bias. Second, we are not estimating just a couple of fixed effects, we are estimating hundreds - or thousands when we use pairs. This causes computational issues that must be addressed. Third, our treatments are given simultaneously to the origin and destination and not independently to each observation. We therefore need to deal with standard errors that should be clustered potentially in complex nested ways. Fourth, we need to account for spatio-temporal dependencies.

The main reason for log-transforming our outcome is that we have more confidence in being able to build an additive linear model of the log-transformed outcome than the pure outcome. The reason is that the outcome has a heavy right tail. While we should expect our predictive error to be small in absolute numbers for small flows, we also expect it to increase as the flows become larger. In such settings, it makes more sense to work with a relative scale and relative errors, which is what you get when you log-transform the outcome.

One issue with the log-linear approach is that heteroscedastic errors lead to biased estimators. It can be shown that since the expected value of a logged error depends on higher-order moments of the error distribution, heteroscedasticity in the model error results in bias of log-linear regression parameter estimates (Cohn et al., 2022; Silva & Tenreyro, 2006). The bias can even cause the sign of coefficients to be wrong, and the bias can be even worse in fixed effects models (Cohn et al., 2022). The simple fix is to use a Poisson model where instead of estimating $E(\log(y)|X)$ as in the linear case, we estimate $\log(E(y|X))$.

treatment effect for the compliers of the instrument, i.e., those who take the treatment if and only if they were assigned to the treatment group (Angrist & Imbens, 1995). In the case of Abel et al. (2019), that means that the effect is only estimated for the population of dyadic flows where SPEI increase the likelihood of conflict occurrence. Non-compliers are either "always-takers" – those that always have conflict irrespective of SPEI, "never-takers" – those that never have conflict irrespective of SPEI, or "defiers" – those where SPEI decrease the likelihood of conflict occurrence.

An alternative is to scale the outcome to produce a rate. However, as Cohn, Liu & Wardlaw (2022) point out, this would require "a suitable scaling variable that captures the potential exposure of an observation to the outcome". Whether the total population in a country is a suitable scaling factor, or if only a subset of the total population ("potential migrants") should be used as a scaling factor is not clear. For instance, if countries with most population close to borders of other countries have a larger share of their population as "potential migrants", using the per capita migration rate might bias results. Just using the Poisson model is the sounder solution.

Another issue with the log-linear approach is that we need to offset zero valued outcomes as the log of zero is undefined. The offset can in principle be any positive number, but a common choice is to add 1, as the log of 1 is 0. However, this arbitrary choice affect estimates whenever there are any nonlinear relationships among covariates (which is a likely case), and particularly when the mean of the outcome distribution is low and the share of zero observations is large (Cohn et al., 2022). For directed dyadic international migration, 51% of the observations are zero, the mean is 2087, and the max is 3 309 139. Moreover, when we offset the outcome, the coefficient estimates cannot be interpreted as (semi-)elasticities "or any other quantity likely to be of interest" (Cohn et al., 2022, p. 9). Again, using Poisson regression will solve these issues.

The three first issues (mentioned in the first paragraph of this section) are dealt with via Laurent Bergé's **fixest** package in R (Bergé, 2018). He has implemented a concentrated likelihood approach that can handle any number of fixed effects (and any number of sets/clusters of fixed effects) both for the linear and the Poisson case. The package also handles estimation of nested clustered standard errors.

The fixed effects regression functions in **fixest** have not implemented ways to account for spatio-temporal dependencies other than through covariates, however. The closest solution we found that would account for such dependencies was the eigenvector spatial filtering approach to network autocorrelation (Chun, 2008; Chun & Griffith, 2011; Griffith et al., 2019). While this filtering approach in principle works in conjunction with the **fixest** approach, it scales poorly with the number of spatial units since we must add the squared-number-of-spatial-units numbers of covariates and run a stepwise regression afterwards. With more than 100 spatial units, that means more than 10 000 covariates. With almost 200 000 observations, we were unable to fit these models. There are plausibly Bayesian MCMC approaches (possibly with random instead of fixed effects) that could work. We tried to estimate random effects Poisson gravity models using **brms** (which can in principle model spatio-temporal autocorrelation too), but these did not converge. We know of monadic models able to estimate Poisson models with fixed effects and spatial autocorrelation (Glaser et al., 2022), but have not seen an implementation of this in R (or any other modeling language). We are unsure whether this approach can be used in the dyadic case.

4. Data

4.1 Variables

Outcome: We use the estimates of international bilateral migration flow from Abel & Cohen (2019). They provide estimates for each origin-destination dyad for 200 countries in 5-year intervals from 1990-2019. They provide estimates using six different methods: stock differencing drop negative ($\boldsymbol{M_{od}^{SDd}}$), stock differencing reverse negative ($\boldsymbol{M_{od}^{SDr}}$), migration flow ($\boldsymbol{M_{od}^{SDr}}$), demographic accounting (DA) open minimization ($\boldsymbol{M_{od}^{DAo}}$), DA closed minimization ($\boldsymbol{M_{od}^{DAc}}$), and DA pseudo-Bayesian ($\boldsymbol{M_{od}^{PB}}$), the latter being the approach suggested by Azose & Raftery (2019). Unless otherwise noted, we use $\boldsymbol{M_{od}^{PB}}$.

Treatment: Most studies of the effects of climate on migration flows are exploring the effects of absolute temperature and precipitation, temperature and precipitation where country averages are subtracted ("de-meaning"), or first-differences in temperature and precipitation, alternatively exploring possible moderating contexts (i.e., treatment heterogeneity) by interacting this effect with some variable (such as agricultural dependency or absolute temperature/precipitation). Temperature and precipitation anomalies, for instance using the Standardized Precipitation-Evapotranspiration Index (SPEI), are also used in a few studies. When adding fixed effects, temperature and precipitation are demeaned, but they retain the country level variance characteristics. Estimating treatment effects across units with varying variance characteristics can bias the fixed effects estimator because it could lead to treatment heterogeneity (Gibbons et al., 2019), which could be a reason to standardize the treatment within each country. Temperature and precipitation data comes from CRU TS-401 (Harris et al., 2020).

It is not necessarily clear why we would want to log-transform the predictors in the model (other than sticking to the original gravity model specification where all coefficient effects are elasticities). A naïve log-transformation of temperature is problematic because it is a continuous variable that crosses zero. One approach is to first normalize temperature to become a variable between 0 and 1 where 0 is the coldest global observation and 1 is the warmest global observation, and then to log-transform it (adding 1 to the variable before log-transformation to avoid the log of zero). Temperatures do not vary much across the globe, it is not extreme-distributed, so it is not clear why we would want to cast its effects on a relative scale, however. Furthermore, if the theory is that the negative effect of temperature is due to some threshold effect, then assigning treatment status whenever we see effects above said threshold might be a better approach. Log-transformation of precipitation might be more reasonable since it has a long right-tail distribution. Here, it might make more sense that effects are linear on the relative scale of the predictor, than on the absolute scale. However, again, if the effect of interest is that at some threshold, e.g., there is either too little precipitation to support plant growth, or too much precipitation to avoid flooding (of plants, of infrastructure, of houses), and that threshold varies across space and time, a global linear model would at best be a very basic approximation of the theoretical idea.

We normalize temperature and precipitation and log-transform them with an offset of 1.

$$lTn_{ot} = log\left(\frac{T_{ot} - \min(T_o)}{\max(T_o) - \min(T_o)} + 1\right) = \log(Tn_{ot} + 1)$$

We denote the normalized, offset, and log-transformed temperature in the origin $\boldsymbol{lTn_{ot}}$ and $\boldsymbol{lTn_{dt}}$ in the destination, and $\boldsymbol{lPn_{ot}}$ and $\boldsymbol{lPn_{dt}}$ likewise for the case of precipitation.

As briefly discussed in the introduction, we have many other alternatives to how we operationalize our climate treatment, either using transformations of temperature and precipitation (e.g., in combination such as in SPEI), or extreme weather events such as cyclones, heatwaves, etc. Another alternative again would be to use climate disaster data. Some of these climatic exposures could arguably affect migration stronger than deviances in 5-year country average temperatures and precipitation.

One reason to use deviances in temperature and precipitation as our treatment (as we do when we add fixed effects) is that it is arguably less endogenous to the social system than extreme weather events and natural disasters (although variances in temperatures and precipitation might pick up differences in how societies have adapted to varying types of climate exposure). It is therefore arguably less difficult to estimate causal effects for such variables. At the same time, it might be argued that the *treatment* each unit is exposed to - and for which we are estimating the average effect *of* - when using precipitation and temperature is highly heterogeneous. E.g., a 5-degree deviation in temperature going from 10 degrees to 15 degrees is different than going from 35 degrees to 40 degrees. Our main reason for using temperature and precipitation here, however, is that most published studies using the gravity model to estimate effects of climate on international migration flow use these variables. Our aim here is to explore the sensitivity of such estimates to varying econometric specifications and assumptions about measurement quality.

Controls: The main controls we use are fixed effects on origin, destination, time, origin-destination pairs, and destination-time pairs. Our time dummies are dummies for the 5-year intervals of our data. In most specifications, we do not add any other controls. However, we estimate the effects of temperature in the origin, temperature in the destination, precipitation in the origin, and precipitation in the destination simultaneously[10].

To further explore the effects of spatio-temporal dependencies, we add covariates that are all log-transformed with an offset of 1.

We add the population in the origin ($\boldsymbol{POP_o}$) and destination ($\boldsymbol{POP_d}$) from WDI (SP.POP.TOTL). We calculate dyadic distances between country polygons from the cShapes dataset (Weidmann et al., 2010). While distances do change somewhat over time due to border-changes, we used values from 2015 (no changes since 2011) for simplicity ($\boldsymbol{D_{od}}$). Dyadic trade data was taken from Correlates of War Trade 4.0 (Barbieri et al., 2009). We use the "smoothtotrade" variable ($\boldsymbol{TF_{do}^{od}}$) which is the sum of the trade flow from both directions smoothed over time. We attempt to control for spatial autocorrelation by controlling for the sum of migration flow (using the Pseudo-Bayesian Demographic Accounting method) into the destination minus the flow from the origin ($\boldsymbol{M_d^{in}}$), the sum of migration flow from the origin minus flow to the destination ($\boldsymbol{M_o^{out}}$), the sum of $\boldsymbol{M_d^{in}}$ for the first-order neighbors of the destination ($\boldsymbol{M_{\dot{d}}^{in}}$), and the sum of $\boldsymbol{M_o^{out}}$ for the first-order neighbors of the origin ($\boldsymbol{M_{\dot{o}}^{out}}$). We seek to control for temporal autocorrelation using the lagged dependent variable ($\boldsymbol{M_{od,t-1}^{PB}}$).

4.2 Descriptive Statistics

Using dyadic data on international migration flows, we analyze the impact of climate on migration, accounting for climatic conditions on both the origin and destination side of the dyad. This section presents the descriptive statistics of the main variables included in the analysis. We proceed with describing the patterns in the data, primarily focusing on the distinction between migration to and from OECD and non-OECD countries, where inflows to OECD countries have, due to data availability, been the primary focus of existing climate-migration analysis. Next, we explore the relationship between migration flows and differences in temperature in origin and destination countries, highlighting the fact that most people migrate to countries with relatively similar temperature levels. Last, we highlight the fact that most migration also trigger return migration, where people travel back to their country of origin, a notion that is often overlooked in the climate-migration literature.

In this paper we do not limit our empirical approach to looking at climatic conditions only in the origin country, often seen as a push factor, but we also account for climatic conditions in the destination country. We estimate the dyadic flows of migrants, including data at the origin and destination side of the dyad. Table 2 presents the general descriptive statistics of the dependent

[10] The sensitivity estimates for the temperature and precipitation in the destination are for the most part not reported.

variable and main independent variables.

Table 2. Descriptive statistics

Variable	Mean	Median	Min	Max	NotNA	PropNA
M_{od}^{PB}	2087	0	0	3309139	234048	0
Tn_o	0.7	0.798	0	1	189729	0.189
Pn_o	0.3	0.271	0	1	189729	0.189
POP_o	33037800.5	6813200	62152	1379860000	224399	0.041
D_{od}	6427705.2	5869881.9	0	19160168	167686	0.284
TF_{do}^{od}	540.8	0.4	-9	618167	91805	0.608
Tn_d	0.7	0.798	0	1	189729	0.189
Pn_d	0.3	0.271	0	1	189729	0.189

Most of global migration flows within continents and between neighboring countries (Moore & Shellman, 2007). Simple calculation linking our dyadic flow data with a spatial contiguity matrix of countries, we find that almost 35% of the world's international migrants move to first-order contiguous countries. However, many migrate to countries farther away. A popular notion is that most international migration flows go to OECD countries (Neumayer, 2004). However, this could be a result of data availability, or simply the empirical focus. As Figure 1 shows, most international migration occurs between non-OECD countries, and most immigrants to OECD countries arrive from other OECD countries. This aligns well with the fact that in 2017, 85 % of all migrants were hosted in developing countries (UNHCR, 2018), suggesting that most international migration is primarily dominated by south-south flows.

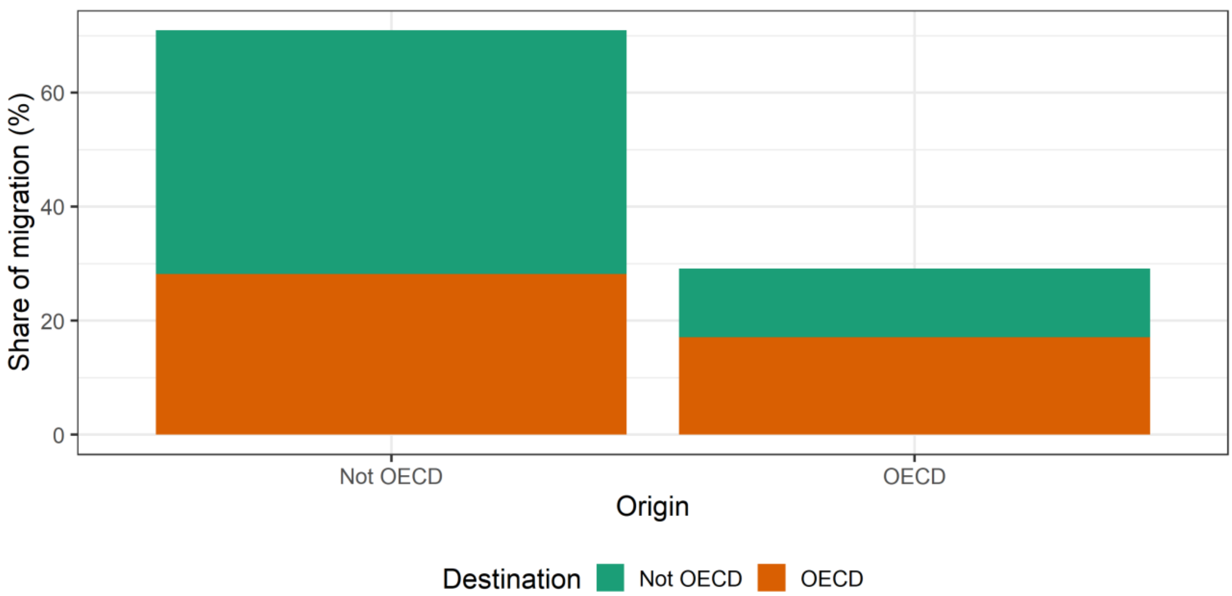

Figure 1. Global migration shares by OECD status for origin and destination country

Since international migration most commonly takes place between contiguous neighbors, that also suggests that most migrants go to countries with similar climatic conditions. Figure 2 shows the distribution of migration flow across temperature differences between origin and destination countries. 62% of migrants move to countries with average temperatures that deviate less than 5 degrees Celsius from their country of origin.

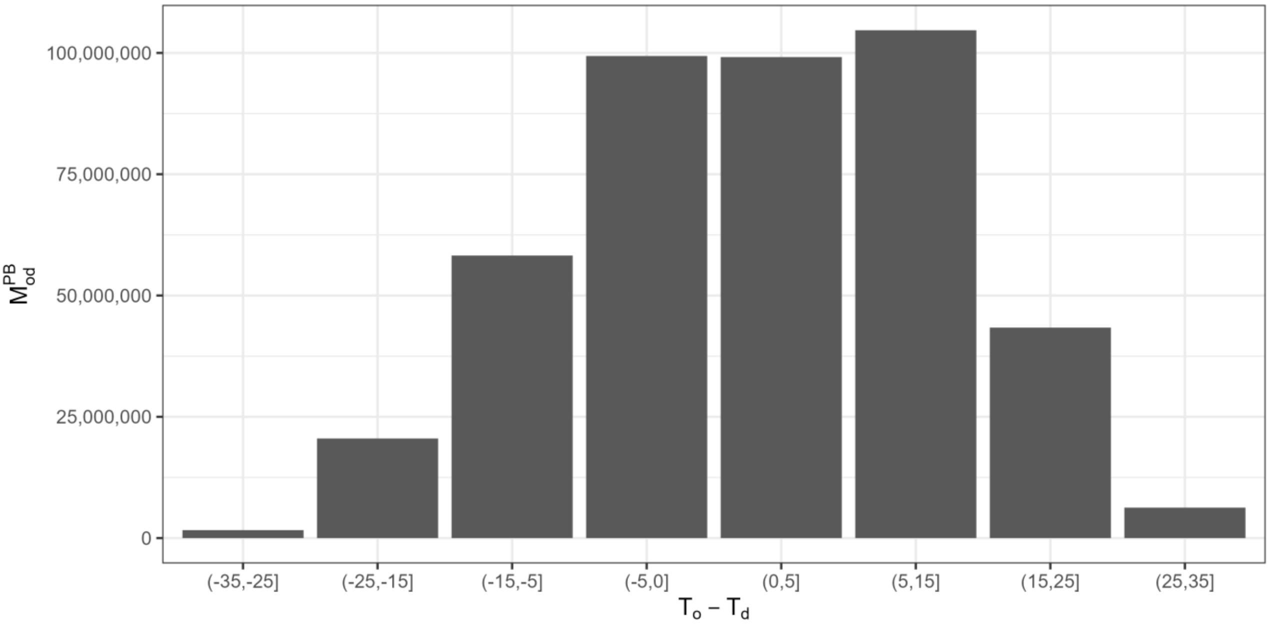

Figure 2. Migration flow against the difference in temperature in origin and destination countries

Migration flows are not constant, nor unidirectional, but migrants leaving are also often returning home, either permanently or to work, visit or maintain property or land in their original location. Migrants are also largely affected by conditions in the destination country. This aligns closely with one of Ravenstein's (1885) laws of migration which stipulates that every migration flow produces a compensating counter-current. Figure 3 shows how migration flows between dyads in 2010 relate to the reverse dyadic flow in 2015. Flows out usually also trigger flows back. Previous research has given little attention to conditions in the destination country. In this paper we account for conditions in the destination country by including temperature and precipitation in the destination.

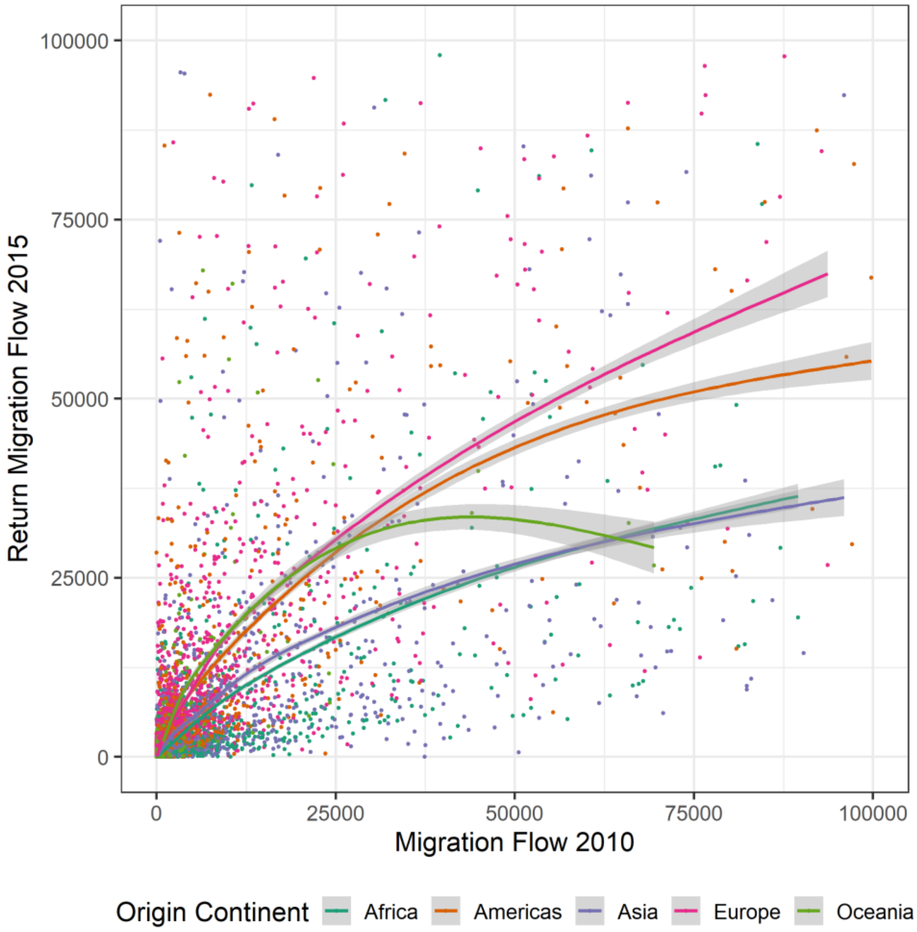

Figure 3. Directed dyadic migration flow in 2010 and the reverse dyadic migration flow in 2015

5. Sensitivity Analysis

In this section, we report the results from a variety of sensitivity tests such as exploring the impact of taking measurement uncertainties into account, the impact of econometric choices such as various fixed effects, ways to calculate standard errors, and the log-linear versus Poisson approaches, and the impact of adding covariates that account for (some) spatial and temporal dependencies. Additionally, we diagnose assumptions in the DiD TWFE model when we have staggered treatments.

5.1 Measurement Uncertainty of the Outcome

5.1.1 Temperature

Figure 4 shows estimates of the effect of temperature ($\boldsymbol{lTn_o}$) on international migration flow across the six different methods to estimate international migration flow from Abel & Cohen (2019). The figure reveals significant variation in estimates depending on how international migration flow is defined. The largest positive effects are estimated when using the stock differencing drop negative or demographic accounting open approaches, while migration rates and the Pseudo-Bayesian demographic accounting approaches yields estimates closest to zero.

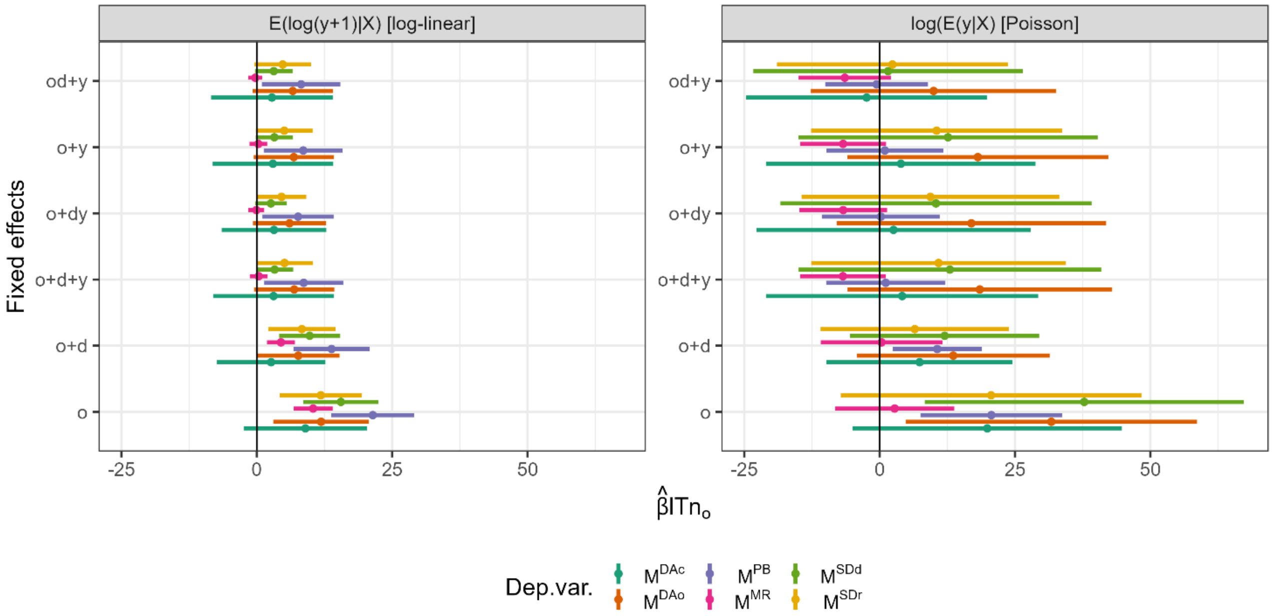

Figure 4 Sensitivity of temperature estimate on method of calculating international migration flow

When we only estimate effects with flows to OECD destinations (Figure 5), the effect tends to be larger than when we use all dyads.

Figure 6 shows a simulation where we add a random Student-t distributed measurement error with 7 degrees of freedom and a scaling factor of $\sigma = 0.59$ (Azose & Raftery, 2019 SI p.12) to the outcome. We assume this error is completely at random. It shows that the standard error of the temperature estimate is underestimated when we are not accounting for measurement errors, and quite substantially so. Since we add the error completely at random, we should not expect this approach to bias the estimate. However, if measurement error is correlated with temperature, then it could also induce bias. Since temperature is correlated with socio-economic development and bureaucratic capacity, it is not unlikely that measurement errors and temperature correlate.

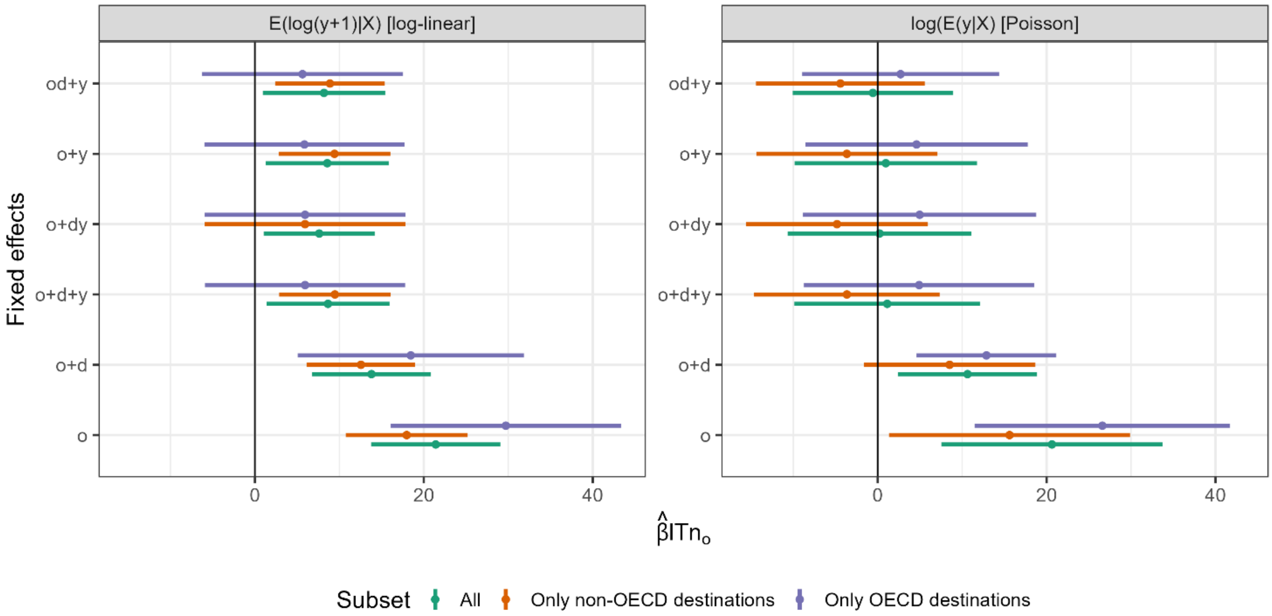

Figure 5. Sensitivity of temperature estimate on spatial subset

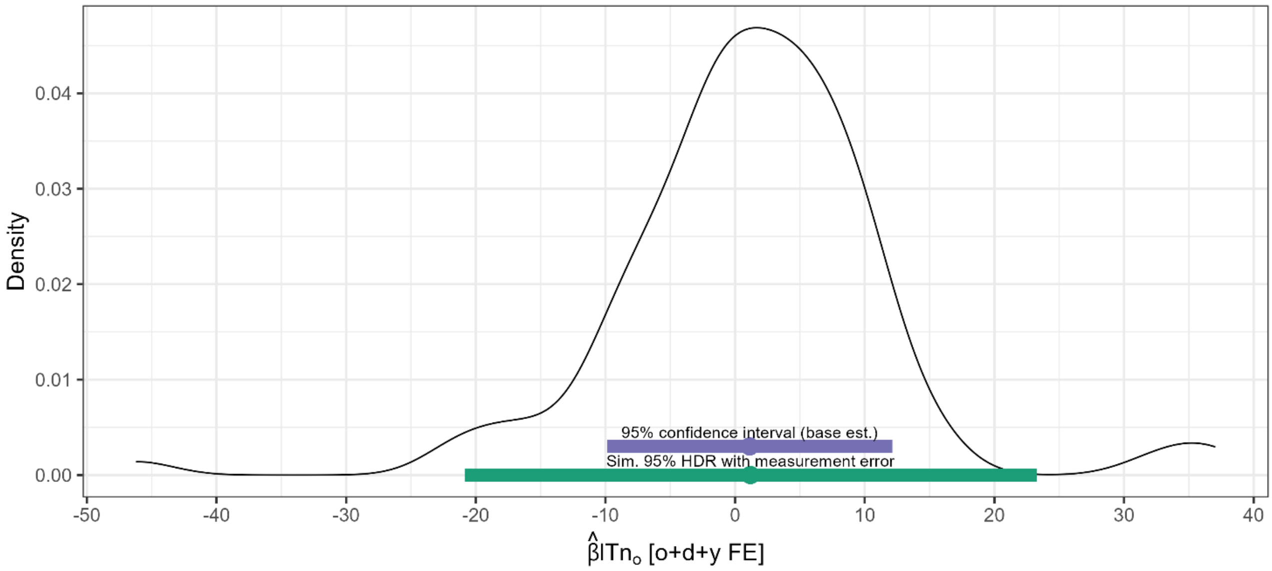

Figure 6. Sensitivity of temperature estimate based on a simulation where measurement uncertainties in the demographic accounting pseudo-Bayesian estimate of migration flow are accounted for assuming measurement errors are completely random

5.1.2 Precipitation

Figure 7 reveals large heterogeneity in results of the effect of precipitation on international migration flow depending on how migration flow is estimated. The stock differencing approaches lead to negative estimates, while the migration rate approach leads to positive estimates. The demographic accounting approaches are in between. It is also worthwhile to see that uncertainty estimates are much larger in the Poisson approach than for the log-linear approach.

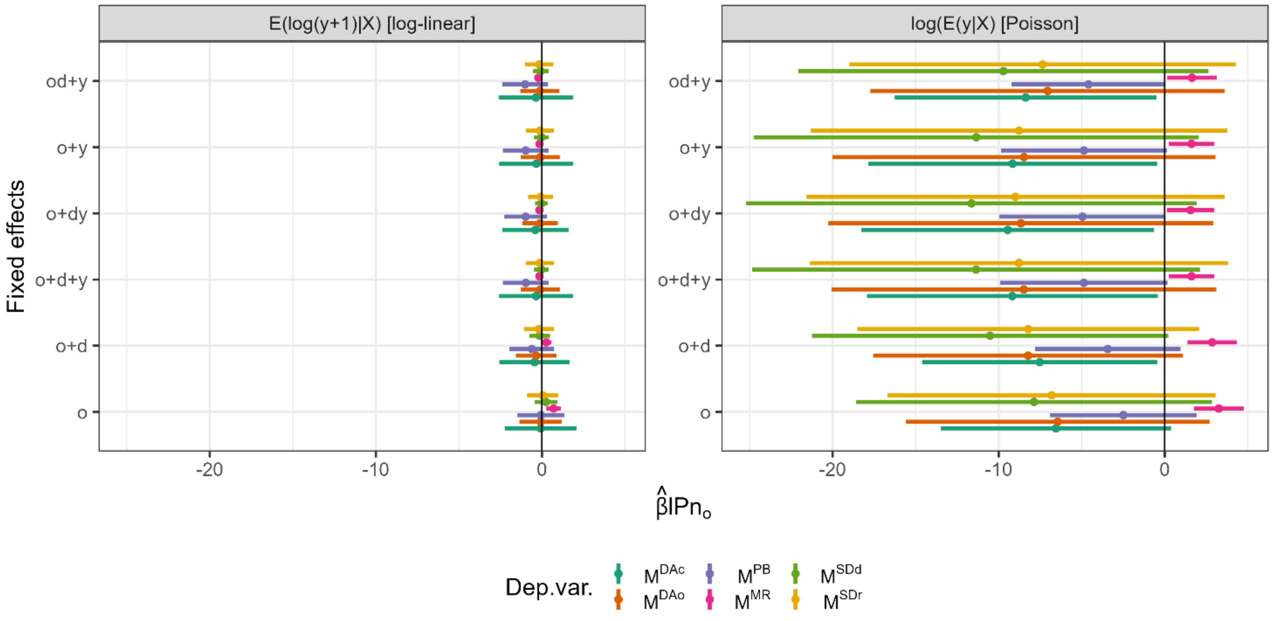

Figure 7. Sensitivity of precipitation estimate on method of calculating international migration flow

Figure 8 shows that the differences in estimates across spatial subsets are negligible for precipitation. Figure 9 reveals that the naïve standard error estimates are smaller than what we would have when accounting for measurement uncertainties (even assuming these are missing completely at random). The difference in estimates is smaller for precipitation than for temperature, however.

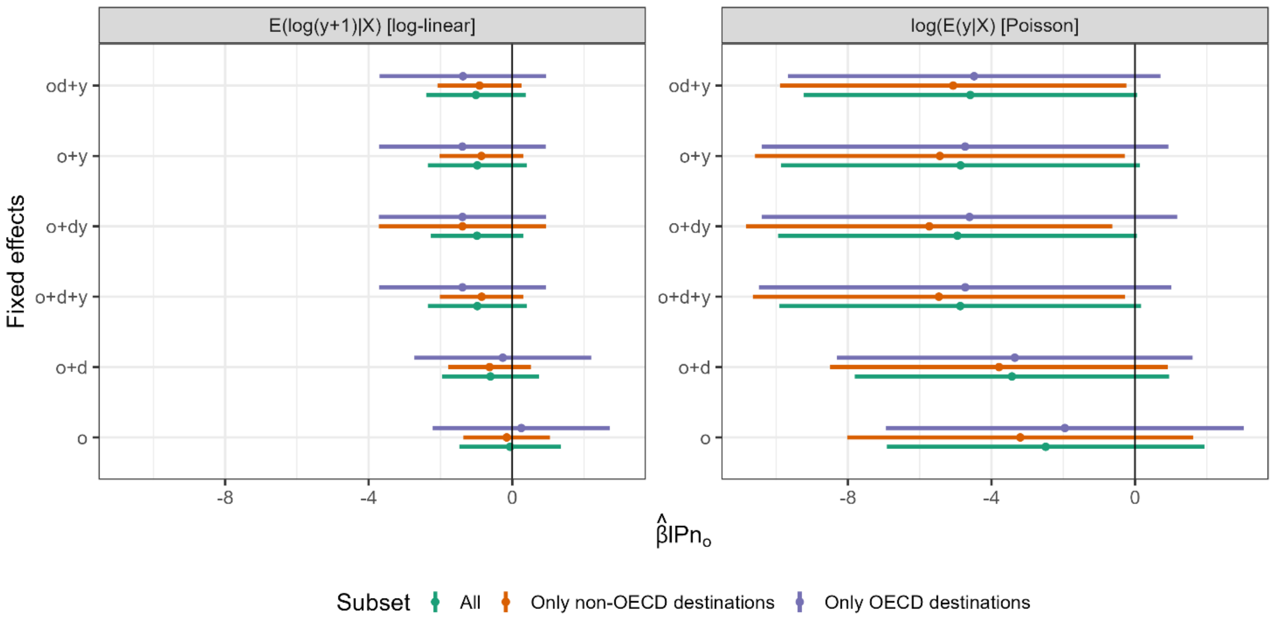

Figure 8. Sensitivity of precipitation estimate on spatial subset

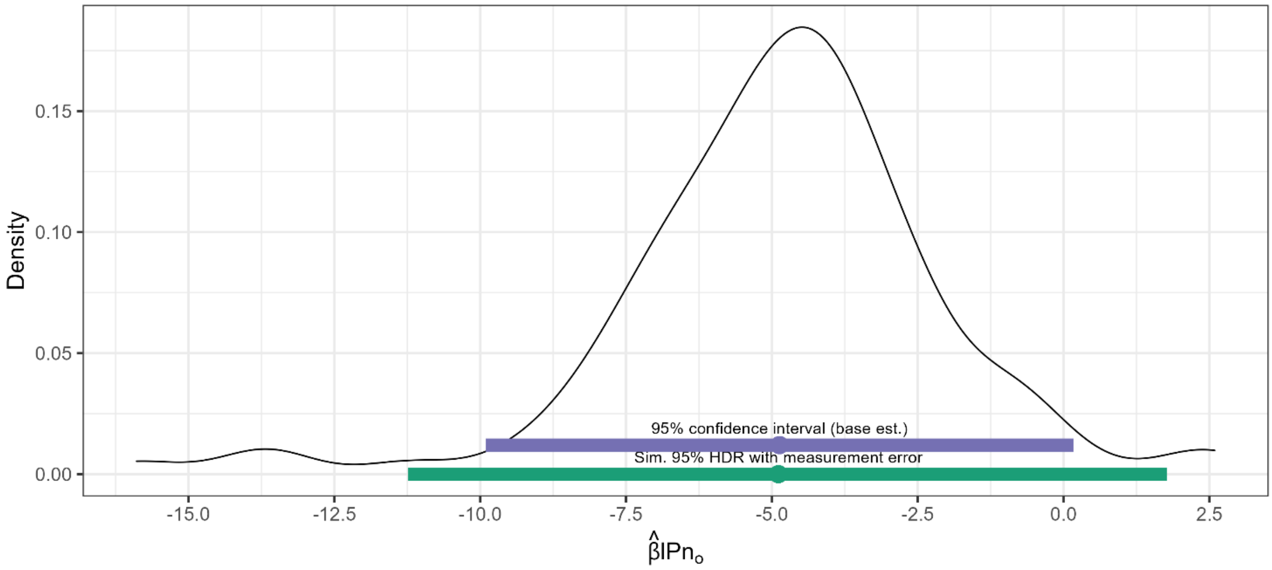

Figure 9. Sensitivity of precipitation estimate based on a simulation where measurement uncertainties in the demographic accounting pseudo-Bayesian estimate of migration flow are accounted for assuming measurement errors are completely random

5.2 Log-linear vs. Poisson, Fixed Effects, and Standard Errors

5.2.1 Temperature

In Figure 10 we see best estimates and 95% confidence intervals of the effect of temperature on international migration flow in log-linear models (left) and Poisson models (right) when we vary the fixed effects specification (y-axis) and how we calculate standard errors (colors). We see that the log-linear specification always produces positive and significant effects when origin fixed effects are included. Since we know that the log-linear model is biased when errors are heteroscedastic, one might conclude that this biases the effect upwards when we compare estimates with the Poisson approach. When estimating the effect of temperature in the origin, it seems adding origin fixed effects is the most impactful choice. Adding year fixed effects shifts the estimate towards zero. Adding destination fixed effects does not seem to have any large effect on the estimate. When using Poisson, there are minor differences between the clustering approach and robust standard errors. In the log-linear specification, two-way clustering yields much larger standard errors than the robust approach.

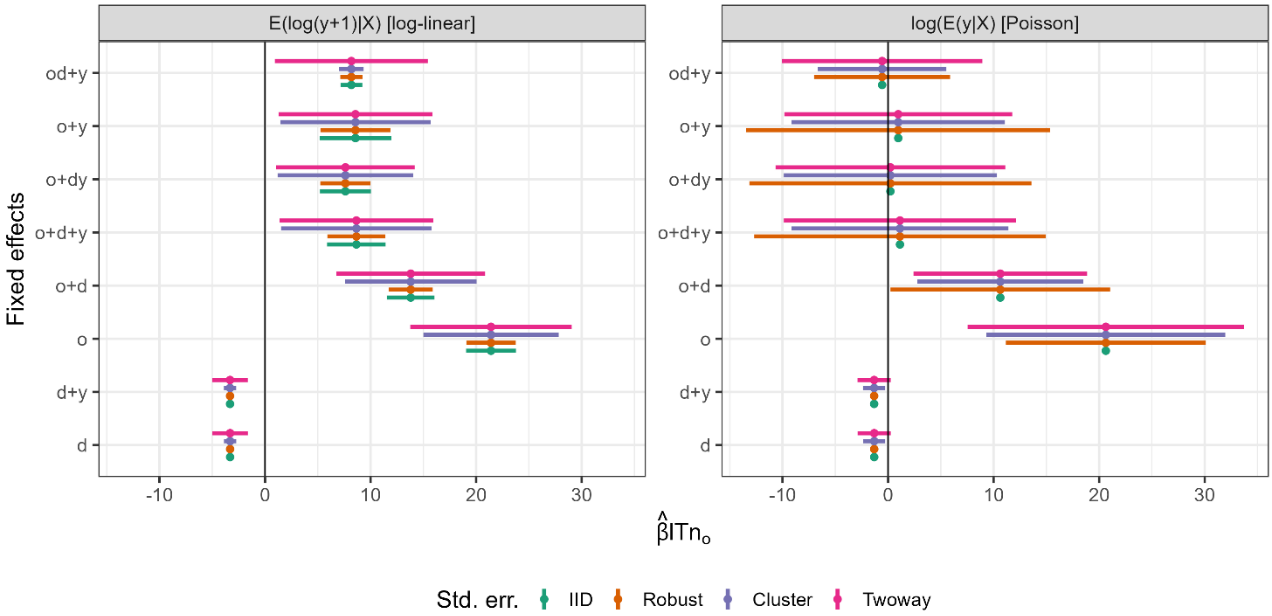

Figure 10. Sensitivity of temperature estimate on modelling choices: log-linear vs Poisson, types of fixed effects, and ways to calculate standard errors

5.2.2 Precipitation

Estimates of precipitation on migration flow looks to be similarly biased upwards in the log-linear approach when we compare with the Poisson models (Figure 11). Unlike the estimate for temperature, which seems to move toward zero when adding controls, the estimate for precipitation is quite consistently on the negative side (although with large uncertainty bounds).

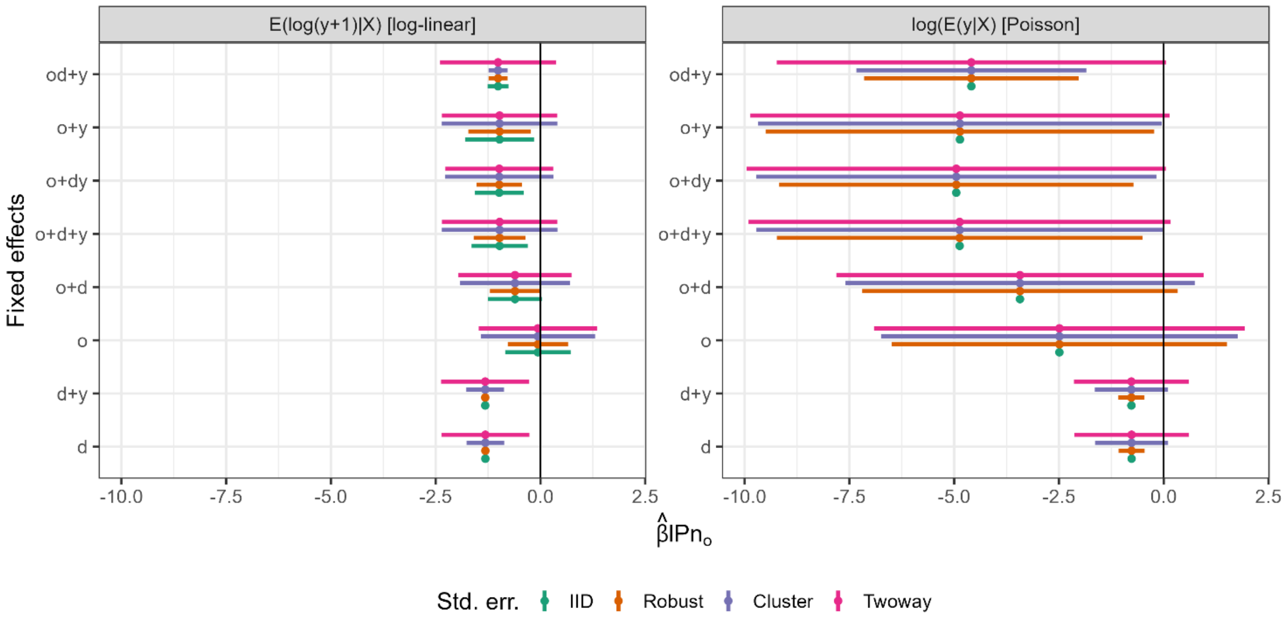

Figure 11. Sensitivity of precipitation estimate on modelling choices: log-linear vs Poisson, types of fixed effects, and ways to calculate standard errors

5.3 Negative Treatment Weights and Treatment Heterogeneity

5.3.1 Temperature

When defining "treated" units as those with origin temperatures that deviate more than a standard deviation from the origin mean temperature in any direction, we calculate that 53% of the treated units have negative treatment weights.[11] The scatterplot in Figure 12 of the residualized treatment against the residualized outcome show a slight tendency for the tails of the residualized treatment to have larger residualized outcomes. The linear hypothesis test reveals a significant interaction between $\boldsymbol{lTn_o}$ and residualized treatment on the the residualized outcome (F = 8.09, Pr(>F) = 0.0045). This means that we cannot reject the hypothesis that there is treatment heterogeneity. When there is treatment heterogeneity and negative treatment weights, the fixed effects estimator of the ATT is biased (Jakiela, 2021).

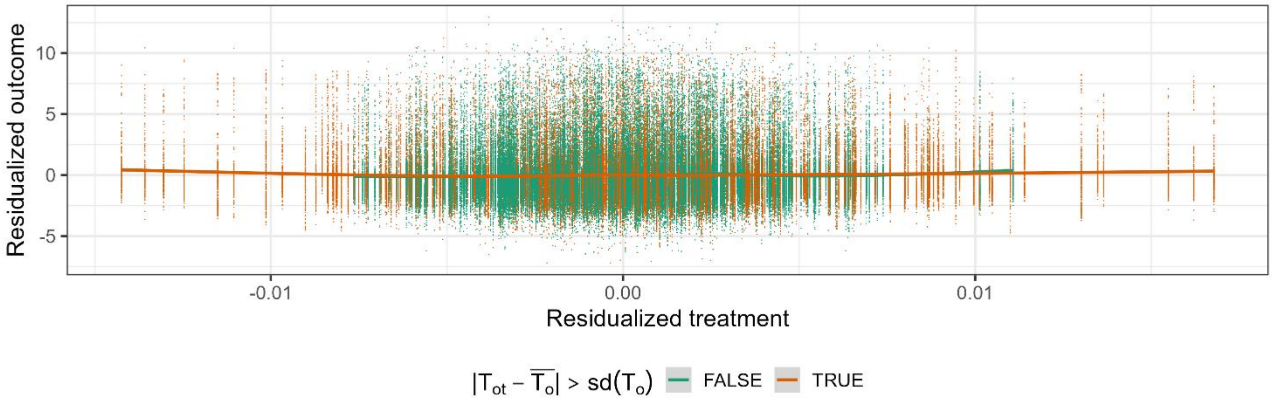

Figure 12. Test of the treatment homogeneity assumption in fixed effects estimation of the ATT for the temperature effect on migration flow. Effects are homogeneous when the relationship between the residualized treatment and outcome is equal for the treated and control groups

An alternative way to test the treatment homogeneity assumption is to subset the data, as we should see the same effect estimate for different subsets if the assumption is correct. Figure 8 has already shown us that effects are on average different in OECD inflows than for non-OECD inflows. Figure 13 shows that dropping both the earliest (1990) and latest (2015) time-periods tend to lead to lower estimates, particularly when using log-linear models. The difference is negligible in the Poisson specification.

[11] An inherent issue with climate variables is that they are treatments that are applied continuously to all units of observations at all times. Indeed, to make any sense, we need to specify when units are more or less treated by climate. Since we are using a fixed effects approach on absolute temperature in the origin, thinking about treated units as those experiencing extreme deviations from the normal origin temperatures might be a reasonable approach.

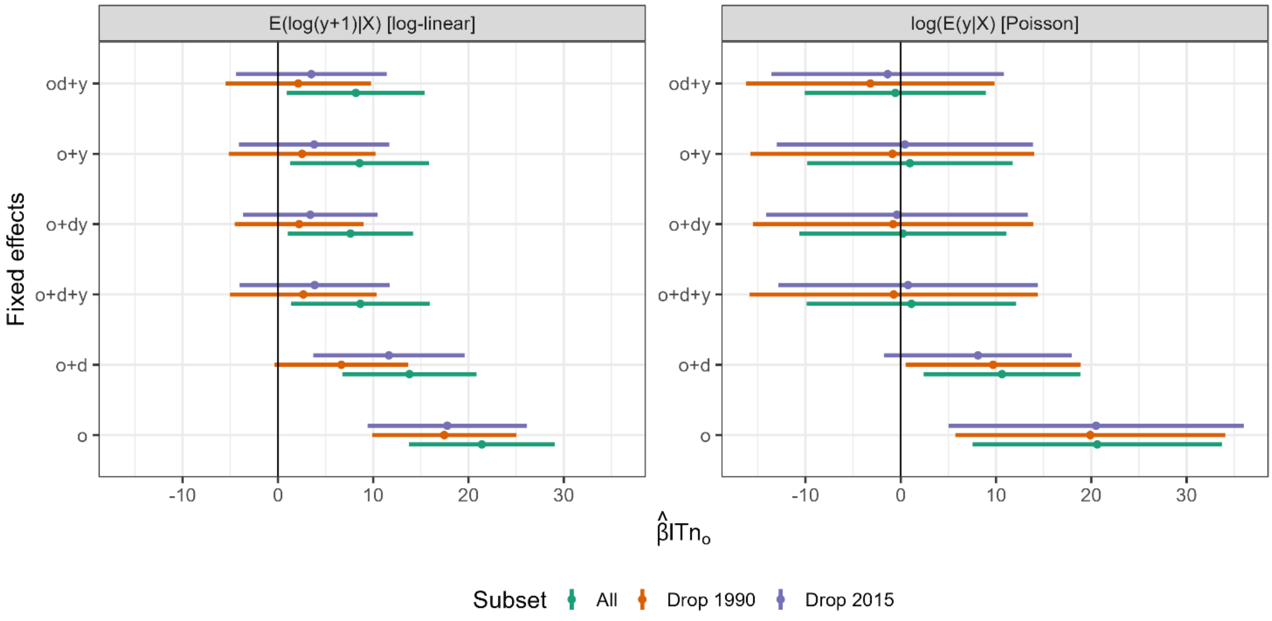

Figure 13. Sensitivity of temperature estimate on temporal subset

5.3.2 Precipitation

Unlike for temperature, there is not a significant linear interaction term between $\boldsymbol{lPn_o}$ and the residualized treatment on the residualized outcome (F = 3.53, Pr(>F) = 0.06), meaning that we can reject the thesis of treatment heterogeneity. We also see this in Figure 14, where there is a less tendency that the residualized outcome is above zero when the residualized treatment is deviating from zero.

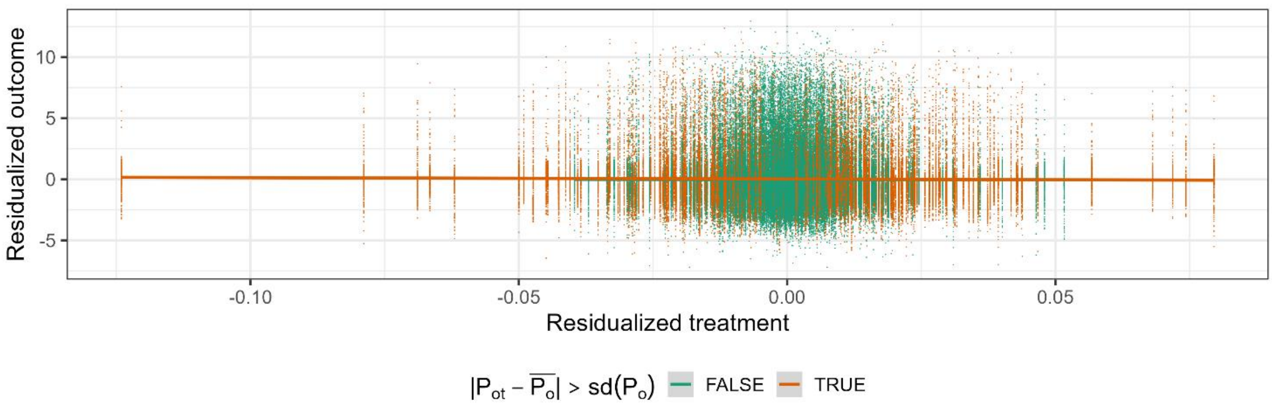

Figure 14. Test of the treatment homogeneity assumption in fixed effects estimation of the ATT for the precipitation effect on migration flow. Effects are homogeneous when the relationship between the residualized treatment and outcome is equal for the treated and control groups

Looking at different subsets, there is less variation in estimates across spatial subsets (Figure 8), but still some when dropping years. Interestingly, as can be seen in Figure 15, estimates are biased downwards when dropping 2015 in the log-linear model but biased upwards in the Poisson model. Dropping 1990 bias the estimate upwards in both the log-linear and Poisson approaches.

The differences in estimates across subsets are still smaller for precipitation, and consistent with a homogeneous treatment effect.

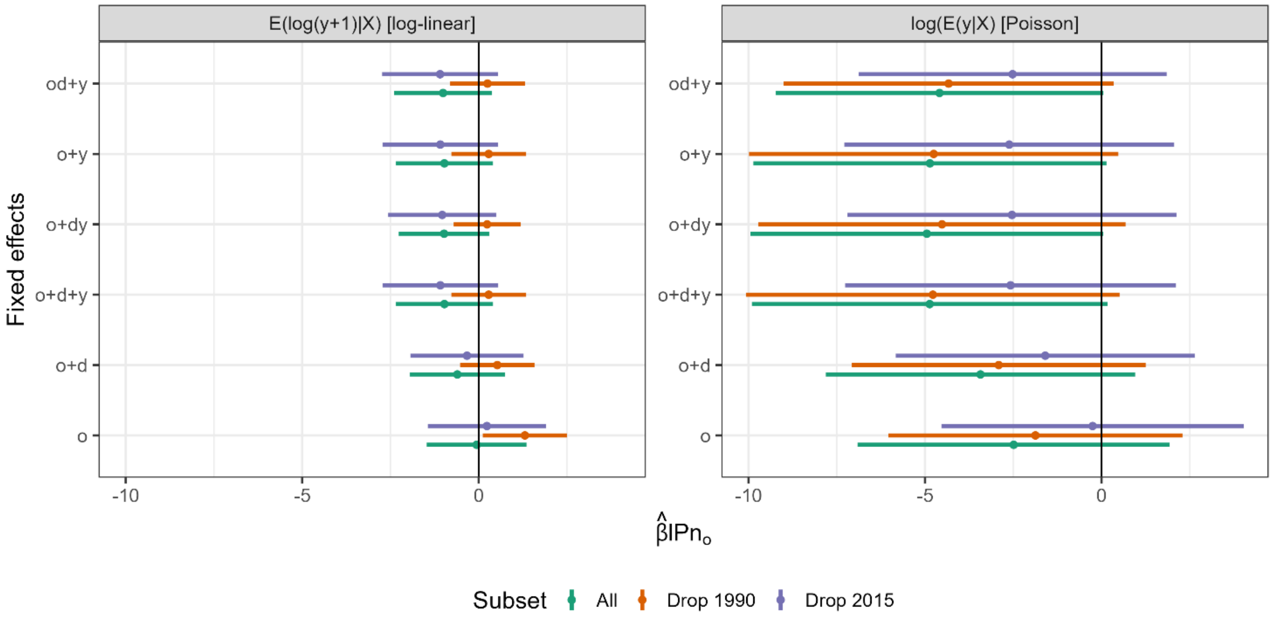

Figure 15. Sensitivity of precipitation estimate on temporal subset

5.4 Adding Covariates (Spatio-Temporal Dependencies)

Table 3 documents the sensitivity of estimates of temperature and precipitation (in origin and in destination) using a Poisson model with origin, destination, and year fixed effects when adding covariates in an attempt to account for spatial and temporal dependencies. The general result is that estimates are sensitive to the covariate specification. The single covariate that changes climate estimates most is the lag dependent variable ($\boldsymbol{M_{od,t-1}^{PB}}$). Adding information on migration flow from neighbors of the origin to the destination ($\boldsymbol{M_o^{out}}$) and into neighbors of the destination from the origin ($\boldsymbol{M_d^{in}}$) seems to affect estimates. One reason here could of course be that neighboring countries are treated with the same climatic treatments as the origin and destination (of the unit of observation) (i.e., violating the non-interference assumption in the DiD approach). Alternatively, flows to a destination is affected by flows to neighboring countries, or flows from an origin is affected by flows from neighboring countries (also possibly violating the non-interference assumption).

Table 3. Sensitivity of temperature and precipitation estimates when adding covariates

	O+D+Y-1	O+D+Y-2	O+D+Y-3	O+D+Y-4	O+D+Y-5	O+D+Y-6	O+D+Y-7
lTn_{ot}	1.26	0.01	3.71	5.19	4.62	3.87	10.11
	(5.57)	(4.94)	(4.91)	(6.43)	(6.46)	(6.73)	(5.73)
lPn_{dt}	-5.03	-4.83*	-4.43	-5.81*	-6.25*	-6.40*	-2.03
	(2.61)	(2.45)	(2.35)	(2.60)	(2.62)	(2.64)	(2.19)
lTn_{dt}	6.53	5.76	6.28	7.46	9.53	10.42	10.36
	(8.74)	(8.00)	(7.43)	(8.86)	(7.69)	(7.42)	(6.63)
lPn_{dt}	-2.46	-2.03	-1.97	-2.45	-1.93	-1.97	-3.85
	(2.32)	(2.06)	(2.05)	(2.32)	(2.24)	(2.20)	(2.66)
$log(D_{od} + 1)$		-0.21***	-0.21***	-0.20***	-0.20***	-0.19***	-0.03***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
$log(POP_o + 1)$			1.00***	1.36***	1.38***	1.45***	1.06*
			(0.23)	(0.33)	(0.33)	(0.34)	(0.52)
$log(POP_d + 1)$			-0.22	-0.17	-0.23	-0.29	-1.07***
			(0.28)	(0.33)	(0.32)	(0.33)	(0.23)
$log(M_d^{in} + 1)$				-0.24	-0.26	-0.24	0.17*
				(0.17)	(0.17)	(0.17)	(0.08)
$log(M_o^{out} + 1)$				-0.33	-0.34	-0.34	0.21*
				(0.21)	(0.22)	(0.22)	(0.09)
$log(M_d^{in} + 1)$					0.03	0.03	0.01*
					(0.02)	(0.01)	(0.01)
$log(M_o^{out} + 1)$					-0.01	-0.01	-0.00
					(0.01)	(0.01)	(0.01)
$log(TF_{do}^{od} + 1)$						0.05	0.01
						(0.03)	(0.01)
$log(M_{od,t-1}^{PB} + 1)$							0.79***
							(0.04)
Num. obs.	141,392	141,392	141,392	141,392	141,392	130,919	109,186
Num. groups: orig	157	157	157	157	157	157	157
Num. groups: dest	157	157	157	157	157	157	157
Num. groups: year	6	6	6	6	6	6	5
Pseudo R^2	0.49	0.67	0.67	0.68	0.68	0.68	0.92

We also note here that the estimates of precipitation and temperature in the origin ($\boldsymbol{lTn_o}$ and $\boldsymbol{lPn_o}$) and in the destination ($\boldsymbol{lTn_d}$ and $\boldsymbol{lTn_d}$) are quite similar. None published article have reported effects of the climate in the *destination* on international migration flows. One reason could be that the theoretical arguments for such effects are less clear than from the origin. If we should have less reason to expect such effects, then the results above should make us suspicious.

6. Conclusion

As can be surmised from the key take-aways below, the sensitivity analysis shows that we likely know less about the effect of temperature and precipitation on international migration flow than what published literature has led us to believe. Our epistemic uncertainties are larger than what has been reported and there are likely other issues that we do not bring up here (Bijak & Czaika, 2020).

While the key issue in this report has been on our ability to estimate the causal effect of climate variability exposure on international migration flows, it can be useful to think about what kind of treatment exactly we are trying to estimate the causal effect of, and if it makes sense to estimate such an effect. To conclude, we therefore discuss some issues regarding this approach more generally.

As both climate variances and absolutes (average conditions) vary across countries, so do the "treatment" the countries are exposed to. Since treatments vary, we should expect treatment effects to vary too. Another source of treatment heterogeneity is that we should expect variance in climatic exposure across individuals in countries over 5-year periods, and a third source is the varying effect such exposure has on the migration choice depending on the contextual factors and individual circumstances. The average effect of these varying exposures might not be particularly relevant for any case where climate exposure did affect the migration choice.

Increasing spatio-temporal resolutions of empirical data could be seen as a useful approach. However, there are limitations to this approach, particularly that climates and weather events are dependent across space and time (so that the actual "number" of independent treatments are limited). High temporal resolution could also limit our ability to identify compound climatic exposure. Higher resolutions also put higher demands on the quality of migration data – quality that does not generally exist for historical data.

If most international migration is to nearby areas or countries with similar climates, which issues are migration supposed to solve for those who choose to migrate due to climatic exposure? One possible answer here is that they seek access to areas or livelihoods less vulnerable to more specific climate induced natural hazards (e.g., erosion, landslides, floods, droughts, extreme temperatures, tropical/convective storms). However, if that is the case it would make more sense to try to estimate a causal effect of such natural hazards rather than of average temperatures or temperature deviations.

This report has focused on international migration. The gravity model has also been used to estimate internal migration flow, most prominently in the World Bank-commissioned Groundswell Reports (Clement et al., 2021; Rigaud et al., 2018). It is reasonable to assume that climate exposure is more likely to affect proximate mobility within countries than cross-border migration (McLeman, 2014). However, estimating causal effects for internal migration might be even more difficult than for international migration. First, just estimating internal migration flows puts large demands on data quality and is partly driven by strong assumptions such as constant birth- and death-rates within countries or within urban and rural parts of countries (Alessandrini et al., 2020; CIESIN, 2011). Furthermore, the quality of subnational population data has been debated (see e.g., Thomson et al., 2021). Second, it is probably even more difficult to account for non-interference in climatic treatments in a domestic setting. While we have been unable to ascertain exactly the econometric approach used in the Groundswell Reports, we assume that the identification issues we report here are also relevant for internal migration flow.

Rather than aiming for a global model of the relationship between climatic exposure and international (or internal) migration flow, which would have to account for varying measurement quality and spatio-temporal dependencies and where the global effect is likely to be small, we believe a more fruitful approach is to collect high-quality data on specific cases where the theoretical and empirical link between climatic exposure and migrant behavior is more directly apparent. Additionally, since the possibility for causal inference is contested, we should complement causal inference analysis with evaluations of predictive performance with and without climate information (e.g., Kiossou et al., 2020; Schutte et al., 2021).

KEY TAKE-AWAYS FROM THE SENSITIVITY ANALYSIS:

- Methodological choices have a substantial impact on the estimated effects of temperature and precipitation on international migration flows. Methodological choices are likely an important cause for the variance of estimated effects reported in published literature.
- Measurement uncertainties – unaccounted for in published literature – have substantial impact on our estimates. We know less about the magnitude and consistency of climate effects on international migration than what published estimates lead us to believe.
- It is not clear that we are able to identify causal effects of climatic exposure on international migration flow using the gravity model with fixed effects. Particularly, the assumption of treatment non-interference seems to be violated. No published article directly controls for spatio-temporal/network dependencies. We are unable to find an econometric approach that does this in an adequate fashion.
- Evaluating predictive performance is a recommended supplement or alternative to causal inference, particularly in observed studies where assumptions in causal inference are difficult to evaluate. Published results exploring the predictive power of climatic variables on international migration flow show that such variables are poor predictors.
- Most international migration occurs between countries that have fairly similar climatic conditions (e.g., neighboring countries).

References

Abel, G. J. (2013). Estimating global migration flow tables using place of birth data. *Demographic Research*, *28*, 505–546.

Abel, G. J., Brottrager, M., Crespo Cuaresma, J., & Muttarak, R. (2019). Climate, conflict and forced migration. *Global Environmental Change*, *54*, 239–249. https://doi.org/10.1016/j.gloenvcha.2018.12.003

Abel, G. J., & Cohen, J. E. (2019). Bilateral international migration flow estimates for 200 countries. *Scientific Data*, *6*(1), 82. https://doi.org/10.1038/s41597-019-0089-3

Adams, H. (2016). Why populations persist: Mobility, place attachment and climate change. *Population and Environment*, *37*(4), 429–448. https://doi.org/10.1007/s11111-015-0246-3

Adger, W. N., Arnell, N. W., Black, R., Dercon, S., Geddes, A., & Thomas, D. S. G. (2015). Focus on environmental risks and migration: Causes and consequences. *Environmental Research Letters*, *10*(6), 060201. https://doi.org/10.1088/1748-9326/10/6/060201

Alessandrini, A., Ghio, D., & Migali, S. (2020). *Estimating net migration at high spatial resolution*.

Publications Office of the European Union. http://dx.doi.org/10.2760/383386

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, *93*(1), 170–192. https://doi.org/10.1257/000282803321455214

Angrist, J. D., & Imbens, G. W. (1995). *Identification and Estimation of Local Average Treatment Effects* (Working Paper No. 118). National Bureau of Economic Research. https://doi.org/10.3386/t0118

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Aronow, P. M., Samii, C., & Assenova, V. A. (2017). Cluster–Robust Variance Estimation for Dyadic Data. *Political Analysis*, *23*(4), 564–577. https://doi.org/10.1093/pan/mpv018

Azose, J. J., & Raftery, A. E. (2019). Estimation of emigration, return migration, and transit migration between all pairs of countries. *Proceedings of the National Academy of Sciences*, *116*(1), 116–122. https://doi.org/10.1073/pnas.1722334116

Backhaus, A., Martinez-Zarzoso, I., & Muris, C. (2015). Do climate variations explain bilateral migration? A gravity model analysis. *IZA Journal of Migration*, *4*(1), 3. https://doi.org/10.1186/s40176-014-0026-3

Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, *144*(2), 370–395. https://doi.org/10.1016/j.jfineco.2022.01.004

Barbieri, K., Keshk, O. M. G., & Pollins, B. M. (2009). Trading Data: Evaluating our Assumptions and Coding Rules. *Conflict Management and Peace Science*, *26*(5), 471–491. https://doi.org/10.1177/0738894209343887

Beine, M., & Jeusette, L. (2021). A meta-analysis of the literature on climate change and migration. *Journal of Demographic Economics*, *87*(3), 293–344. https://doi.org/10.1017/dem.2019.22

Beine, M., & Parsons, C. (2015). Climatic Factors as Determinants of International Migration. *The*

Scandinavian Journal of Economics, *117*(2), 723–767. https://doi.org/10.1111/sjoe.12098

Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: The R package FENmlm. In *DEM Discussion Paper Series* (No. 18–13; DEM Discussion Paper Series). Department of Economics at the University of Luxembourg. https://ideas.repec.org/p/luc/wpaper/18-13.html

Beyer, R. M., Schewe, J., & Lotze-Campen, H. (2022). Gravity models do not explain, and cannot predict, international migration dynamics. *Humanities and Social Sciences Communications*, *9*(1), 1–10. https://doi.org/10.1057/s41599-022-01067-x

Bijak, J. (2011). *Forecasting International Migration in Europe: A Bayesian View*. Springer. https://link.springer.com/book/10.1007/978-90-481-8897-0

Bijak, J., & Czaika, M. (2020). *Assessing Uncertain Migration Futures—A Typology of the Unknown*. University of Southampton and Danube University Krems.

Black, R., Adger, W. N., Arnell, N. W., Dercon, S., Geddes, A., & Thomas, D. (2011). The effect of environmental change on human migration. *Global Environmental Change*, *21*, S3–S11.

Black, R., Bennett, S. R. G., Thomas, S. M., & Beddington, J. R. (2011). Migration as adaptation. *Nature*, *478*(7370), 447–449. https://doi.org/10.1038/478477a

Boas, I., Farbotko, C., Adams, H., Sterly, H., Bush, S., van der Geest, K., Wiegel, H., Ashraf, H., Baldwin, A., Bettini, G., Blondin, S., de Bruijn, M., Durand-Delacre, D., Fröhlich, C., Gioli, G., Guaita, L., Hut, E., Jarawura, F. X., Lamers, M., … Hulme, M. (2019). Climate migration myths. *Nature Climate Change*, *9*(12), 901–903. https://doi.org/10.1038/s41558-019-0633-3

Cai, R., Feng, S., Oppenheimer, M., & Pytlikova, M. (2016). Climate variability and international migration: The importance of the agricultural linkage. *Journal of Environmental Economics and Management*, *79*, 135–151. https://doi.org/10.1016/j.jeem.2016.06.005

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carling, J., Czaika, M., & Erdal, M. B. (2020). *Translating migration theory into empirical propositions.* Quantmig Project Deliverable 1.2.

Cattaneo, C., Beine, M., Fröhlich, C. J., Kniveton, D., Martinez-Zarzoso, I., Mastrorillo, M., Millock, K., Piguet, E., & Schraven, B. (2019). Human Migration in the Era of Climate Change. *Review of Environmental Economics and Policy*, *13*(2), 189–206. https://doi.org/10.1093/reep/rez008

Chun, Y. (2008). Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographical Systems*, *10*(4), 317–344. https://doi.org/10.1007/s10109-008-0068-2

Chun, Y., & Griffith, D. A. (2011). Modeling Network Autocorrelation in Space–Time Migration Flow Data: An Eigenvector Spatial Filtering Approach. *Annals of the Association of American Geographers*, *101*(3), 523–536. https://doi.org/10.1080/00045608.2011.561070

CIESIN. (2011). *Migration and Global Environmental Change: MR4 Estimating net migration by ecosystem and by decade: 1970-2010*. CIESIN. https://ciesin.columbia.edu/documents/foresight-2011-oct.pdf

Clement, V., Rigaud, K. K., de Sherbinin, A., Jones, B., Adamo, S., Schewe, J., Sadiq, N., & Shabahat, E. (2021). *Groundswell Part 2: Acting on Internal Climate Migration*. World Bank. https://openknowledge.worldbank.org/handle/10986/36248

Cohn, J. B., Liu, Z., & Wardlaw, M. (2022). *Regression with Skewed, Non-negative Outcome Variables in Finance* (SSRN Scholarly Paper No. 3800339). Social Science Research Network. https://doi.org/10.2139/ssrn.3800339

Coniglio, N. D., & Pesce, G. (2015). Climate variability and international migration: An empirical analysis. *Environment and Development Economics*, *20*(4), 434–468. https://doi.org/10.1017/S1355770X14000722

Czado, C., Gneiting, T., & Held, L. (2009). Predictive Model Assessment for Count Data. *Biometrics*, *65*(4), 1254–1261. https://doi.org/10.1111/j.1541-0420.2009.01191.x

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Fafchamps, M., & Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, *83*(2), 326–350. https://doi.org/10.1016/j.jdeveco.2006.05.005

Ferris, E. (2020). Research on climate change and migration where are we and where are we going? *Migration Studies*, *8*(4), 612–625. https://doi.org/10.1093/migration/mnaa028

Fischer, M. M., & Griffith, D. A. (2008). Modeling Spatial Autocorrelation in Spatial Interaction Data: An Application to Patent Citation Data in the European Union*. *Journal of Regional Science*, *48*(5), 969–989. https://doi.org/10.1111/j.1467-9787.2008.00572.x

Foresight. (2011). *Migration and Global Environmental Change: Future Challenges and Opportunities*. The Government Office for Science.

Gemenne, F., Barnett, J., Adger, W. N., & Dabelko, G. D. (2014). Climate and security: Evidence, emerging risks, and a new agenda. *Climatic Change*, *123*(1), 1–9. https://doi.org/10.1007/s10584-014-1074-7

Gibbons, C. E., Serrato, J. C. S., & Urbancic, M. B. (2019). Broken or Fixed Effects? *Journal of Econometric Methods*, *8*(1). https://doi.org/10.1515/jem-2017-0002

Glaser, S., Jung, R. C., & Schweikert, K. (2022). Spatial panel count data: Modeling and forecasting of urban crimes. *Journal of Spatial Econometrics*, *3*(1), 2. https://doi.org/10.1007/s43071-021-00019-y

Golenvaux, N., Alvarez, P. G., Kiossou, H. S., & Schaus, P. (2020). *An LSTM approach to Forecast Migration using Google Trends* (arXiv:2005.09902). arXiv. https://doi.org/10.48550/arXiv.2005.09902

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of*

Econometrics, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Griffith, D. A., Chun, Y., & Li, B. (Eds.). (2019). *Spatial Regression Analysis Using Eigenvector Spatial Filtering*. Academic Press. https://doi.org/10.1016/B978-0-12-815043-6.09990-0

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109. https://doi.org/10.1038/s41597-020-0453-3

Hoffmann, R., Dimitrova, A., Muttarak, R., Crespo Cuaresma, J., & Peisker, J. (2020). A meta-analysis of country-level studies on environmental change and migration. *Nature Climate Change*, *10*(10), 904–912. https://doi.org/10.1038/s41558-020-0898-6

Hunter, L. M., & Simon, D. H. (2022). Time to Mainstream the Environment into Migration Theory? *International Migration Review*, 01979183221074343. https://doi.org/10.1177/01979183221074343

IPCC. (2022). *Summary for Policymakers. Working Group II Contribution to the IPCC Sixth Assessment Report*. Intergovernmental Panel on Climate Change.

Isard, W. (1954). Location Theory and Trade Theory: Short-Run Analysis. *The Quarterly Journal of Economics*, *68*(2), 305–320. https://doi.org/10.2307/1884452

Isard, W. (1966). *Methods of Regional Analysis: An Introduction to Regional Science*. The MIT Press.

Jakiela, P. (2021). *Simple Diagnostics for Two-Way Fixed Effects* (arXiv:2103.13229). arXiv. https://doi.org/10.48550/arXiv.2103.13229

Kiossou, H. S., Schenk, Y., Docquier, F., Houndji, V. R., Nijssen, S., & Schaus, P. (2020). *Using an interpretable Machine Learning approach to study the drivers of International Migration* (arXiv:2006.03560). arXiv. https://doi.org/10.48550/arXiv.2006.03560

Lenormand, M., Bassolas, A., & Ramasco, J. J. (2016). Systematic comparison of trip distribution laws and models. *Journal of Transport Geography*, *51*, 158–169. https://doi.org/10.1016/j.jtrangeo.2015.12.008

LeSage, J. P., & Pace, R. K. (2008). Spatial Econometric Modeling of Origin-Destination Flows*. *Journal of Regional Science*, *48*(5), 941–967. https://doi.org/10.1111/j.1467-9787.2008.00573.x

McLeman, R. A. (2014). *Climate and human migration: Past experiences, future challenges*. Cambridge University Press.

Molnar, C. (2019). *Interpretable Machine Learning*. Christoph Molnar.

Moore, W. H., & Shellman, S. M. (2007). Whither will they go? A global study of refugees' destinations, 1965–1995. *International Studies Quarterly*, *51*(4), 811–834.

Neumayer, E. (2004). Asylum Destination Choice: What Makes Some West European Countries More Attractive Than Others? *European Union Politics*, *5*(2), 155–180. https://doi.org/10.1177/1465116504042444

OECD. (2015). *OECD International Migration Database*. Organization for Economic Co-Operation and Development.

Piguet, E. (2010). Linking climate change, environmental degradation, and migration: A methodological overview. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(4), 517–524. https://doi.org/10.1002/wcc.54

Ravenstein, E. G. (1885). The Laws of Migration. *Journal of the Statistical Society of London*, *48*(2), 167–235. https://doi.org/10.2307/2979181

Raymer, J., Wiśniowski, A., Forster, J. J., Smith, P. W. F., & Bijak, J. (2013). Integrated Modeling of European Migration. *Journal of the American Statistical Association*, *108*(503), 801–819. https://doi.org/10.1080/01621459.2013.789435

Rigaud, K. K., de Sherbinin, A., Jones, B., Bergmann, J., Clement, V., Ober, K., Schewe, J., Adamo, S., McCusker, B., Heuser, S., & Midgley, A. (2018). *Groundswell: Preparing for Internal Climate Migration*. World Bank. https://www.google.com/url?q=https://openknowledge.worldbank.org/bitstream/handle/10986/29461/WBG_ClimateChange_Final.pdf&sa=D&ust=1569585643138000&usg=AFQjCNH

kSEpG_vTFA5JsgrR4F3mEZK8ELQ

Robinson, C., & Dilkina, B. (2018). A Machine Learning Approach to Modeling Human Migration. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 1–8. https://doi.org/10.1145/3209811.3209868

Schutte, S., Vestby, J., Carling, J., & Buhaug, H. (2021). Climatic conditions are weak predictors of asylum migration. *Nature Communications*, *12*(1), 2067. https://doi.org/10.1038/s41467-021-22255-4

Sellner, R., Fischer, M. M., & Koch, M. (2013). A Spatial Autoregressive Poisson Gravity Model. *Geographical Analysis*, *45*(2), 180–201. https://doi.org/10.1111/gean.12007

Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*(4), 641–658. https://doi.org/10.1162/rest.88.4.641

Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96–100. https://doi.org/10.1038/nature10856

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Tabord-Meehan, M. (2019). Inference With Dyadic Data: Asymptotic Behavior of the Dyadic-Robust t-Statistic. *Journal of Business & Economic Statistics*, *37*(4), 671–680. https://doi.org/10.1080/07350015.2017.1409630

Thomson, D. R., Gaughan, A. E., Stevens, F. R., Yetman, G., Elias, P., & Chen, R. (2021). Evaluating the Accuracy of Gridded Population Estimates in Slums: A Case Study in Nigeria and Kenya. *Urban Science*, *5*(2), 48. https://doi.org/10.3390/urbansci5020048

UNHCR. (2018). *Global Trends: Forced Displacement in 2017*. United Nations High Commissioner for Refugees. https://www.unhcr.org/globaltrends2017/

Vestby, J. (2019). Climate variability and individual motivations for participating in political violence. *Global Environmental Change*, *56*, 114–123. https://doi.org/10.1016/j.gloenvcha.2019.04.001

Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research*, *47*(4), 363–375. https://doi.org/10.1177/0022343309356491

Weidmann, N. B., Kuse, D., & Gleditsch, K. S. (2010). The Geography of the International System: The CShapes Dataset. *International Interactions*, *36*(1), 86–106. https://doi.org/10.1080/03050620903554614

Zipf, G. K. (1946). The P1 P2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, *11*(6), 677–686. https://doi.org/10.2307/2087063

■