Spring 6-1-2021

Improving existing methods for calculating embodied carbon emissions in trade through feature discovery: an information theoretic approach

Sam Morton
Dartmouth College, Samuel.H.Morton.21@dartmouth.edu

1]Sam Morton 2]Rajarshi Das 3]Temiloluwa Prioleau [1]45 Lebanon Street, Hanover, NH 03755 [2]Westchester County, NY [3]Dartmouth College, NH

Embodied emissions in trade, greenhouse gas emissions, information theory, Leontief algebra

Please provide an abstract of no more than 300 words. Your abstract should explain the main contributions of your article, and should not contain any material that is not included in the main text.

DISCOVERY OF NATURAL ANALOGS TO SUPPLEMENT ANTICIPATED GLOBAL EMBODIED EMISSIONS IN TRADE: AN INFORMATION THEORETIC PERSPECTIVE

An Undergraduate Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

in

by

[

DARTMOUTH COLLEGE

Hanover, New Hampshire

June 8, 2021

Examining Committee:

, Chair

Dean of the Guarini School of Graduate and Advanced Studies

Introduction

An increasingly global world economy has widened spatial ranges between consumers and the production of goods. Lower production costs in developing nations have allowed production in developed economies to shift point sources of pollution overseas while simultaneously lowering costs. Though this practice reduces local pollution in the developed nation, it may increase the cumulative carbon footprint of a product. Products must now travel further to reach shelves, and are now often produced in regions with less stringent environmental standards.

Rising concern regarding the risks posed by global climate change have inspired interest in studying anthropogenic carbon sources and how their effects can be mitigated in the future. One such carbon source is international trade. Though international trade is widely known to be economically beneficial to participating nations, it may also be a significant carbon source. The increasingly interconnected and global economy of the twenty-first century has also increased interest in the impact of trade on overall emissions. This has prompted research into methods for quantifying the contribution of international trade to overall carbon emissions. Given that carbon dioxide (CO_2) is the primary anthropogenic greenhouse gas by volume, embodied emissions in trade measurements often focus primarily on the embodied emissions of CO_2. This quantification of overall carbon emissions is known as the embodied emissions in trade (EET). Understanding EET has the potential to allow both researchers and policymakers to better account for the contribution of international trade to overall emissions. Modern methods for calculating embodied emissions in trade can yield insights into trends in overall EET of nations, as well as indicate which sectors and trading partners contribute the most or least to overall emissions.

Previous work in this field has found that international trade accounts for 13-14% of total emissions from the world's largest economies [1]. For other nations, EET accounted for up to 50% of total emissions [1].

Typical analysis of EET invokes linear algebraic models developed by Leontief known as single-region and multi-region input-output models, or SRIO and MRIO models respectively. SRIO models are less sophisticated than MRIO models, but require fewer data. However, MRIO models are more complex and allow for more sophisticated simulations of international trade.

EET studies, and in particular those implementing MRIO models, frequently rely on data published by the global MRIO database, which has published releases in 2013 and 2016, each of which include data for several dozen countries and are likewise divided into sectors. The resulting MRIO data therefore takes the form of an NxN grid, where N is equal to the number of countries included in the dataset multiplied by the number of sectors in the data.

Though datasets such as the Global MRIO Database or Eora26 are highly useful for researchers and the development of augmentations of the base Leontief model, they are limited in both their frequency and temporal scope. Offering data on yearly basis with the most recent releases spanning no further than 2015, these data limit the ability of researchers to develop real-time models for present or future EET projection or calculate EET values for time periods less than a year. It therefore may be worthwhile for researchers to consider alternative data sources with larger numbers of features, higher frequency, and greater temporal range. Variables which contain enough information regarding EET could serve as proxies and facilitate predictive models to estimate true EET.

This paper seeks so uncover variables related to the EET of a given country which are available in higher frequency and greater temporal range than is possible with existing MRIO datasets. To assess the information regarding EET contained within a given variable, linear measures such as Pearson correlation were applied. To measure nonlinear dynamics between select variables and EET, information theoretic methods such as mutual information, conditional mutual information, and transfer entropy were leveraged.

Related work

The subject of EET has received cross-disciplinary interest from the fields of economics, statistics, and environmental science, and thus has enjoyed abundant attention in the literature.

Materials and Methods

Data

Several data sources were leveraged for the purposes of this study. Country-wise EET calculation was completed using the Global MRIO Database 2013 release. EET generation via an MRIO model requires two data inputs, both of which are included in the MRIO Database release: international input-output data and environmental accounts data. As both datasets were provided by the same organization, and therefore both required data sources were broken down into analogous sectors which allowed for efficient feature engineering and generation of an EET dataset. NEED TO PROVIDE SPECIFICS ABOUT THE SIZE AND SHAPE OF THE DATA

Research into variables which are highly related to EET focused on two datasets: (1) select features taken from the U.N. Comtrade Database and (2) select features taken from the World Bank Indicators database. The U.N. Comtrade database offers publicly available international commodities and services trade data for most countries. Data can be either monthly or yearly from 2010 through the present day, or yearly stretching from the 1980s to the present day. This study focused on U.S. commodities imports and exports data between 1995 and 2009 as this corresponded to the data available for EET calculation. Between both U.S. imports and exports, this international trade dataset contained XXX features, all of which were compared to the generated EET data.

Additional features were selected to supplement the generated EET data U.N. Comtrade international trade data. Environmental indicators NO_2, Fossil Fuel Consumption, Forest Cover, Total Greenhouse Gas Emissions, CO_2 Emissions, CO_2 Emissions from Liquid Fuel, and CH_4 were taken from the World Bank public database. Each feature provides yearly values and were selected between 1995 and 2009, the temporal range of this study.

Some LaTeX Examples

Use section and subsection commands to organize your document. LaTeX handles all the formatting and numbering automatically. Use ref and label commands for cross-references.

Figures and Tables

Use the table and tabular commands for basic tables — see Table 1, for example. You can upload a figure (JPEG, PNG or PDF) using the project menu. To include it in your document, use the includegraphics command as in the code for Figure 1 below.

[width=0.7]frog

Figure 1: An example image of a frog.

Item	Quantity
Candles	4
Fork handles	?

Table 1: An example table.

Citations

LaTeX formats citations and references automatically using the bibliography records in your .bib file, which you can edit via the project menu. Use the cite command for an inline citation, like [**?**], and the citep command for a citation in parentheses lees2010theoretical.

Mathematics

LaTeX is great at typesetting mathematics. Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $E[X_i] = \mu$ and $Var[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n} \sum_i^n X_i$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

Lists

You can make lists with automatic numbering …

[noitemsep]Like this, and like this.

…or bullet points …

[noitemsep]Like this, and like this.

. . . or with words and descriptions . . .

Word Definition

Concept Explanation

Idea Text

Acknowledgments

Additional information can be given in the template, such as to not include funder information in the acknowledgments section.

Bibliography

[1] N. Ahmad and A. Wyckoff, "Carbon dioxide emissions embodied in international trade of goods," 2003.