Causal Abstraction Via Emergence for Predicting Bilateral Trade

Aruna Jammalamadaka AJAMMALAMADAKA@HRL.COM
Dana Warmsley DMWARMSLEY@HRL.COM
Tsai-Ching Lu TLU@HRL.COM
Information Systems and Sciences Laboratory
HRL Laboratories, LLC
Malibu, CA 90265, USA

Editor: Sisi Ma, Erich Kummerfeld

Abstract

Causal abstraction is key in finding efficient representations of noisy and complex systems, for decision-making and prediction of future system states. Hand-crafted causal abstractions, although accurate and interpretable, can be costly to construct and cannot generalize to large, novel datasets. In this paper, we explore the information-theoretic concept of causal emergence, its correspondence to recent definitions of causal abstraction, and the properties of emergent representations that enable more accurate state predictions and semantic interpretations. Using the bilateral trade network as a case study, we enumerate the conditions under which trade agreements exhibit causal emergence properties, and show that causally emergent representations are indeed able to provide better prediction capability than original trade network representations in a variety of cases.

Keywords: Causal Emergence, Causal Abstraction, Bilateral Trade, Trade Prediction

1. Introduction

Abstract representations of space, time, and information enable us to make predictions, and therefore good decisions, when faced with complex scenarios. Causal abstraction, in particular, is required to construct optimal decision-making policies. Its goal is to summarize a given system while maintaining all causal information in the form of causal dependencies and intervention effects. Causal abstraction can be performed using expert knowledge or physical properties, however these methods do not easily generalize to new observational data sets. In this paper, we focus on quantitative and automated methods; namely we map recent definitions of causal abstraction by Beckers and Halpern (2019) to networks of causal variables, using the information-theoretic concept of causal emergence (Hoel et al., 2013). Causal emergence allows us to reveal a data-driven abstract representation of the observed bilateral trade network, which provides better (albeit coarser granularity) prediction of future trade, and which we hypothesize could be useful for assessment of economic policies (interventions).

Given a set of causally related variables, *causal abstraction* seeks to find a lower-dimensional representation of those variables such that the causal information (i.e., effects of interventions) is preserved. Recent theoretical papers on causal abstraction have carefully defined what it means for a causal system to be abstracted (Beckers and Halpern, 2019; Chalupka et al., 2017), or approximately abstracted (Beckers et al., 2019). These papers focus on a single pairwise causal relationship. Specifically, a high-dimensional pair of cause and effect "micro-variables," which can be mapped to a lower dimensional pair of "macro-variables" by abstracting away the non-causal details of the system. *Causal emergence* (Hoel et al., 2013), on the other hand, is an information-theoretic measure defined on the state progression of networks. The set of all node values (variables) in the network represents the state of a system, and transitions between possible states are represented as a transition probability matrix (TPM). Similar to causal abstraction, the method seeks to find a "macro-node" representation of the network which can accurately approximate the TPM of the original "micro-node" network. However, causal emergence seeks to improve the causal representation, where that improvement is based on the determinism and degeneracy of state transitions. These key properties are used to characterize the causal information of the system, turning the search for macro-nodes into an optimization over micro-node groupings which increase the determinism and/or decrease degeneracy of the resulting network. The concept of causal emergence has also been adapted by Klein and Hoel (2020) to provide macro-node groupings for purely observed, non-causal graphs. Section 2 discusses causal emergence and its relationship to causal abstraction and state prediction in further detail. To the best of our knowledge, we are the first to draw this connection between the two ideas.

Inferring causal models from macro-economic data is a difficult task, due to unaccounted for confounding factors which threaten the validity of causal estimates, limits in data availability such as selection-biased samples (Hünermund and Bareinboim, 2019), and the general oversimplification of models related to the complex interconnectivity of the global economy. Many studies focus on pairs of cause-effect relationships (Doremus et al., 2019) using variants of Granger causal tests (Granger, 1969) or structural vector autoregressions (SVAR) (Moneta et al., 2011), which force researchers to assume a specific set of potential causal variables a priori (Hall et al., 2019; Su et al., 2019). Such models are abstractions which embed researcher's assumptions about the system, and then fit parameters which may turn out to be statistically significant or predictive, but do not necessarily prove their assumptions. One extreme case of such hand-crafted abstractions are Computable General Equilibrium (CGE) models (Aguiar et al., 2019), where relationships between all economic variables are predefined by simple equations, allowing researchers to study the effects of economic shocks within a simulated setting. These models, while criticized for their use in analyzing the North American Free Trade Agreement (Stanford, 1993), are still used by entities such as the World Bank to determine the causal impact of exogenous shocks to economic systems (Ianchovichina et al., 2016). Here, the shock variables are not truly exogenous, since one can only administer shocks to predefined endogenous variables. Causal emergence allows us to avoid many of the pitfalls associated with these approaches to economic modeling by deriving causal macro-variables in an automated and data-driven way.

In this study, we focus on discovering a macro-node abstraction of the bilateral trade network using causal emergence. This abstract representation, which maintains the causal information of the noisier network of micro-nodes, should indicate the appropriate level of detail for policy makers to optimize their predictions regarding the effects of their interventions on the market. With this goal in mind, we test the ability of causal emergence to characterize specific trade agreement dynamics, and more generally to improve the prediction of future economic states (bilateral trade). To our knowledge, this is the first paper to outline the correspondence between causal abstraction and causal emergence, and the first application of causal emergence to prediction of a real-world time-evolving causal network. This paper attempts to answer the following:

- How does emergence map to abstraction and should emergent representations impact prediction accuracy? (Section 2)
- Do trade agreements between countries constitute appropriate causal macro-nodes? (Section 3.2)
- Does the emergent representation of the bilateral trade network provide better prediction capability than the micro-node representation? (Section 3.3)

Section 2 presents the ideas and implementation details of causal emergence. Section 3.1 describes the economic dataset used in this study. Section 3 presents the experiments and results and Section 4 contains conclusions and next steps.

2. Causal Emergence

As opposed to the aforementioned recent definitions of causal abstraction which do not explicitly mention abstracting causal networks of micro-nodes to macro-nodes, causal emergence aims to do just this. It relies on finding coarse-grained macro mechanisms which are more effective (more deterministic and/or less degenerate) than the underlying micro mechanisms. The concept of "emergence" assumes that in some specific cases "the whole is greater than the sum of its parts", and is contrary to that of "supervenience" (Hoel et al., 2013), which is the idea that the properties of micro-level systems necessarily determine all properties of macro-level systems.

Causal emergence relies on the assumption that the effective information (EI) contained in a causal network's connectivity is a measure of its causal information, and can be characterized by the uncertainty in the out-weights and in-weights of its nodes. Finding a causally emergent representation then translates to finding macro-nodes of a network, either in space, time, or both, by grouping nodes which have similar causal effect on other nodes in the network. Specifically, following the definition of Klein and Hoel (2020), EI is computed as

$$EI = H(< W_i^{out} >) - < H(W_i^{out}) > \tag{1}$$

where W_i^{out} is a vector of edge weights w_{ij} from node v_i to each other node v_j and $\sum_j w_{ij} = 1$, $<>$ represents the averaging function, and function H represents Shannon entropy (Shannon, 1948). The first term $H(< W_i^{out} >)$ characterizes the extent to which uncertainty is distributed across the entire network, thereby offering an understanding of how degenerate the network is. If the average of the outweights from all nodes i to any given node j is particularly high (resulting in low entropy, high degeneracy), it will be difficult to determine the previous node in the path (the "cause" from the "effect"). The second term uses Shannon entropy $H(W_i^{out})$ to measure the uncertainty of a node's output, and averages this value across nodes to characterize the level of determinism. Nodes with low entropy have outweight distributions concentrated on one or few nodes, making it easier to determine the next node in the path (the "effect" from the"cause"). As a measure of causation, EI captures how effectively (deterministically and uniquely) causes produce effects in the system, and how selectively causes can be identified from effects. EI is a general measure for causal interactions because it uses perturbations to capture the effectiveness/selectivity of the mechanisms of a system in relation to the size of its state space. EI is maximal for systems that are deterministic and not degenerate, and decreases with noise (causal divergence) and/or degeneracy (causal convergence). Although a corresponding increase in prediction accuracy of the macro-node system representation is not explicitly addressed in the literature, the optimization of these key properties of causality indicate that if one can find a macro-level grouping of nodes with a higher EI, it should intuitively lead to a better prediction model.

In the original definition of Hoel et al. (2013), the maximum EI is found by perturbing each system through the entire set of possible causal states ("counterfactuals," in the general sense of alternative possibilities) and evaluating the resulting effects using EI. Not only is this combinatorial state-space expensive to estimate for networks of more than a few nodes, but it assumes access to interventional data, i.e. one must be able to set the network to any given state and observe the resulting TPM. In Klein and Hoel (2020) the concept of causal emergence is adapted for purely observed static networks, and an intervention effect is estimated by dropping a random walker on a particular node in the graph (system state) and observing its path. The random walker's path over nodes represents the causal process of various subsequent system states being realized. The accuracy of the macro-node representation in this case is determined by whether or not random walkers behave consistently between the abstracted network (G_{macro}) and the original network (G_{micro}), where inconsistency is defined as the Kullback-Leibler (KL) divergence between the expected distribution of random walkers on G_{macro} and G_{micro} given identical initial conditions.

Note that the causal emergence process differs from the community detection problem of graph theory in that community detection is focused on subgraphs that have more in-group connectivity than out-group, whereas macro-nodes represent subgraphs that possess a viable summary statistic in terms of their behavior in the network, and therefore macro-nodes can exist over a range of connectivity patterns. Additionally, after finding appropriate subgraphs, macro-nodes are a recasting of the network itself (Griebenow et al., 2019).

2.1. Mapping Emergence to Abstraction

Klein and Hoel (2020)'s definition of emergence can be mapped to Beckers et al. (2019)'s definition of approximate causal abstraction. Following the terminology in Beckers et al. (2019), the function τ which maps micro to macro-variables in causal emergence is simply an aggregation of system states (nodes), either in time or (network) space. A *τ-abstraction* is one for which τ induces a natural mapping between interventions in the two spaces, which in the case of causal emergence is captured by the four methods of aggregating micro-node edge weights into macro-node edge-weights outlined in Klein and Hoel (2020). This is also a *strong τ-abstraction* since once the new edge weights are calculated all interventions on the macro-level model are allowed, and more specifically it is a *constructive τ-abstraction* since the individual macro-variables can be computed from non-overlapping subsets of micro-variables. Finally, it is a *constructive* $\tau - \alpha$ *approximate abstraction* since the interventional effects between G_{micro} and G_{macro} differ by at most α, where α is the value of the KL divergence between random walker distribution over micro-nodes vs macro-nodes. By mapping causal emergence to the definitions from causal abstraction theory, we further validate its usefulness in causal inference.

2.2. Implementation Details

Since the EI of a network is dependent on the network's size N, we use the normalized form of EI known as "effectiveness" to better compare graphs of different sizes (e.g. before and after creating a new macronode). Effectiveness ranges from 0.0 to 1.0 and is computed as

$$effectiveness = \frac{EI}{log_2(N)} \tag{2}$$

This differs from the publicly available implementation provided by Klein and Hoel (2020), which can lead to counterintuitive results because it uses EI.

Another known issue with the available implementation is that the use of "out-weights" alone skews results for smaller networks which contain nodes that do not have any emanating edges. These nodes are ignored in the computation of EI and effectiveness, resulting in the counterintuitive result shown in Figure 1. Due to the size of our trade network we do not implement any adjustment for this here, but we believe it should be considered in future work.

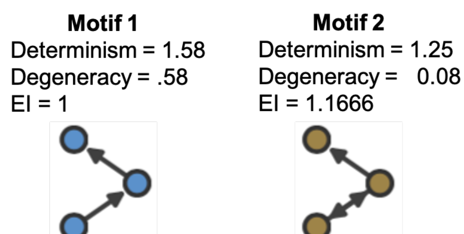

Figure 1: EI (and effectiveness) for Motif 2 are higher than for Motif 1 despite Motif 1 being a perfect causal chain with theoretical determinism = 1.0 and degeneracy = 0.0 according to Hoel et al. (2013).

3. Experiments and Results

In this section, we describe our exploration regarding emergence of trade network macro-nodes. Given our bilateral trade network (described in section 3.1), we seek to understand whether trade organizations and agreements can be considered macro-nodes, and whether the measure for causal emergence naturally groups countries at the macro-level in a way that positively affects prediction of future trade. Since we do not have direct intervention capabilities on the bilateral trade network as outlined in Hoel et al. (2013), we employ Klein and Hoel (2020)'s formulation for identifying causal emergence on general networks.

Section 3.2 focuses on select trade agreements which naturally partition the bilateral trade network into macro-nodes (groups of countries) which are likely to exhibit causal emergence based on their network structure, according to the findings of Klein and Hoel (2020). We examine causal emergence on those trade agreements and compare it to that of randomly selected sets of countries. In Section 3.3, we test our hypothesis that the macro-node representation of a system should achieve better prediction accuracy on future states of the system by comparing the prediction capabilities of network features on bilateral trade using a traditional regression approach.

3.1. Bilateral Trade Network

We make use of the BACI International Trade Database (Gaulier and Zignago, 2010) in order to re-construct the bilateral trade network over a span of several years. Trade networks constructed from this data are ones in which each country is represented by a node, and the directed, weighted edges between country nodes represent the export volume from the origin country to the destination country. Although this is not an explicitly causal network, the aggregation of products traded between countries represents the flow of trade and can be thought of as a causal process in which resources transition from state to state, or country to country. The resulting macro-nodes have implications regarding trade policies and agreements as described in section 3.2.

The database provides bilateral trade data that is disaggregated into the Harmonized System (HS Rev 1992, 4-digit level). The data consists of import and export values for various origin-product-destination triplets, and has been cleaned to ensure consistency among importer and exporter trade reports for each given product (Feenstra et al. (2005), Jun et al. (2019)). We filter our data following the work of Jun et al. (2019), focusing on the years 2000-2017. Our analyses include countries with a population of at least 1.2 million and trade volume of at least one billion (US dollars). As in Jun et al. (2019), we also remove data related to Iraq, Chad and Macau.

3.2. Emergence of Trade Agreement Macro-Nodes

Klein and Hoel (2020) outlines a number of useful conditions under which the opportunity for causal emergence, or an increase in effectiveness, exists. In general, effectiveness increases when nodes that have the same causal effect on the rest of the network are grouped together, thereby increasing determinism and decreasing degeneracy. With regard to network structure, this corresponds to networks which have clusters (groups of nodes) that are either a) bipartite-like with connections mostly existing between clusters or b) clique-like with connections mostly within clusters *and* where the relative sizes of clusters differ (size asymmetry). These conditions create a high level of uncertainty or noise in the network, since a network with size asymmetry has a less evenly distributed $< W_{out}^{i} >$ and therefore high degeneracy, and a network with very low (bipartite) or very high (clique-like) within cluster connections has low determinism. This uncertainty or noise can more likely be reduced by an emergent macro-node representation.

We consider two trade agreements as potential macro-nodes. The first is the Organization of the Petroleum Exporting Countries (OPEC), a trade organization created in 1960 by oil-producing countries that coordinates the petroleum policies of its members. We include only countries that were members of OPEC for the year span of 2000-2017 - Algeria, Congo, Ecuador, Iran, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, United Arab Emirates, and Venezuela. In the context of the bilateral trade network, it seems intuitive that the countries within OPEC have a similar causal effect on the rest of the network. In addition,

Table 1: Effectiveness values (and percent change in effectiveness) resulting from manual grouping of trade agreements into macro-nodes. Effectiveness of bilateral oil trade network in which OPEC countries are grouped is higher due to the coordination of policies and shared causal effect on the rest of the world.

	2000	**2005**	**2010**	**2015**
Original Network	.521	.529	.512	.484
OPEC	.546 (+4.8%)	.549 (+3.8%)	.535 (+4.5%)	.503 (+3.9%)
NAFTA	.515 (−1.2%)	.516 (−2.5%)	.508 (−0.8%)	.478 (−1.2%)
Random Grouping	.509 (−2.1%)	.516 (−2.5%)	.498 (−2.7%)	.457 (−5.6%)

a macro-node including OPEC would be both asymmetric (since OPEC is a small subset of total countries which trades widely with the rest of the world) and bipartite-like (since member countries mainly export crude oil to non-OPEC nations and there is minimal trade of oil between those importing countries). Therefore, we hypothesize that causal emergence will be significant in this context. The second is the North American Free Trade Agreement (NAFTA), created in 1994, which is a treaty entered into by the United States, Canada, and Mexico that eliminates all tariff and non-tariff barriers of trade and investment between member countries. Typical items traded include fruits, vegetables, textiles, vehicles and auto parts, and mineral fuels. While NAFTA resulted in increased trade *within* its member countries, it did not make any stipulations for trade with the rest of the world. As such, NAFTA countries are likely to differ in their trading patterns with outside countries and are therefore not likely to have the same causal effect on the rest of the network. By grouping them, we actually lose information related to those patterns. Therefore, a macro-node containing NAFTA countries is unlikely to show causal emergence because it would have significant between and within cluster connections.

Since OPEC is a crude oil trading organization and crude oil is also one of the top products traded between NAFTA countries, we restrict our focus to the crude oil trade network, i.e., directed, weighted edges between two countries represent the amount of crude oil exported from one country to the other. Using this network, we tested the hypothesis that grouping OPEC countries by trade agreement membership should increase the effectiveness of the oil trade network, while grouping NAFTA countries should offer limited, if any, improvement. We compare the effectiveness of the original oil trade network G_{micro} to that of three modified G_{macro} trade networks with OPEC members grouped, NAFTA members grouped, and a randomly chosen set of countries (with no formal trade agreement) grouped into one macro-node. The modified networks are said to exhibit *causal emergence* if the effectiveness of the network increases after grouping nodes, and *causal reduction* if effectiveness decreases.

Table 1 shows the effectiveness values and percent change in effectiveness from manually grouping macro-nodes in the networks corresponding to four different years. The results presented in the "Random Grouping" row are an average over 20 trials, each with a different set of randomly chosen countries. The results confirm our hypotheses that grouping OPEC countries into a macro-node results in causal emergence whereas grouping NAFTA results in causal reduction. Finally, as expected, randomly grouped countries show the greatest causal reduction since these groups have no formal organization.

We also ran the greedy algorithm implementation on the entire crude oil trade network to see whether we could find any semantic meaning or economic interpretation of the resulting macro-nodes. While the algorithm was able to find abstractions for each G_{micro} that exhibited causal emergence, the macro-nodes found were not entirely intuitive. Each country within NAFTA belongs to separate macro-nodes, as expected based on our results from the manual grouping above. However, OPEC countries were also spread across multiple macro-nodes. While most of the OPEC countries fell into one macro-node, the remainder were either placed in other macro-nodes or remained micro-nodes. In addition, the macro-node that contained the largest number of OPEC countries also contained a large number of non-OPEC countries. These results, combined with the manual grouping results above, indicate that while the OPEC trade agreement is a driving causal force in the network, there are more complicated trade patterns remaining to be considered.

3.3. Prediction Capability of Emergent Representations

Following Jun et al. (2019), we formulate a regression model for predicting bilateral trade volume at future time points. Since the goal of Jun et al. (2019) was to determine whether product or geographic neighbor relatedness played a significant role in predicting bilateral trade flows between countries, their regression model relied on independent variables which included three measures of relatedness (product, exporter, and importer) as well as standard gravity model-related measures such as gross domestic product, geographic distance between origin and destination, and language proximity.

Since our goal is to determine whether the macro-node representation of the trade network provides a better basis for prediction of future export volume between two countries, we instead focus solely on independent variables related to network structure features. By doing so, we avoid having to make ad-hoc choices for the mappings between micro and macro-node gravity features which may not make sense. For example, since importer and exporter relatedness are partially based on geographic distance between two countries, we would need to define the geographic location of a macro-node, which is an aggregate of multiple countries. Averaging over pairwise distances between countries in the macro-node, or taking the centroid of all country locations are not desirable approaches if the countries in the macro-node are globally spread out. Network features, however, have been proven to impact the evolution of trade (Kosowska-Stamirowska, 2020) and are straightforward to compute on both the G_{micro} and G_{macro} networks. As shown by the abbreviations in Figures 2 and 3, the network features we use as independent variables in the regression model include number of common neighbors, in-degree, out-degree, closeness centrality (Freeman, 1979), betweenness centrality (Brandes, 2001), PageRank centrality (Page et al., 1999), and eccentricity of both the origin and destination countries. Using the louvain clustering algorithm (Blondel et al., 2008), we also include a variable indicating whether or not the origin and destination countries belong to the same community.

As in Jun et al. (2019), we split the results into three time periods based on the Global Financial Crisis (GFC) starting in 2007. These correspond to 2000–2006 (pre-financial crisis), 2007–2012 (crisis period), and 2013–2017 (recovery period). We also add a regression model for the full time period of the dataset (2000-2017). In order to check whether our results are product-dependent, we construct three different views of the trade network by summing the export values of all products, oil products (crude oil, refined petroleum, and petroleum gas), and fruit products (bananas, tropical and citrus fruits, grapes, melons, apples, pears, pitted and other fruits) between countries to create the edge weight values. For

Table 2: Resulting R^2 coefficients for regression models built using the time period of the entire dataset, 2000–2006 (pre-financial crisis), 2007–2012 (crisis period), and 2013–2017 (recovery period), and using either all products, only oil products, or only fruit products. In all cases we see that the regression model fit is better for the G_{macro} representation.

Product Type	**Years**	**Prediction Horizon**	$\boldsymbol{R^2, G_{micro}}$	$\boldsymbol{R^2, G_{macro}}$
All Products	2000 – 2017	1 year	.578	.686
		2 years	.574	.693
	2000 – 2006	1 year	.577	.677
		2 years	.574	.707
	2007 – 2012	1 year	.577	.682
		2 years	.574	.674
	2013 – 2017	1 year	.589	.715
		2 years	.592	.733
Oil Products	2000 – 2017	1 year	.244	.447
		2 years	.238	.458
	2000 – 2006	1 year	.248	.458
		2 years	.245	.460
	2007 – 2012	1 year	.252	.440
		2 years	.246	.455
	2013 – 2017	1 year	.237	.461
		2 years	**.240**	**.483**
Fruit Products	2000 – 2017	1 year	.310	.487
		2 years	.306	.474
	2000 – 2006	1 year	.313	.461
		2 years	.314	.456
	2007 – 2012	1 year	.311	.522
		2 years	.302	.502
	2013 – 2017	1 year	.298	.510
		2 years	.299	.492

each time periods and view, we compare the quality of the regression model of the original trade network G_{micro} to that of the abstracted trade network G_{macro} by running the causal emergence algorithm to automatically identify macro-nodes in the network, recalculating network features based on these macro-nodes, and re-fitting the regression model.

We find that for all twelve cases, G_{macro} provides a better R-squared coefficient than G_{micro}, indicating a better prediction model, and supporting our hypothesis that causal emergence increases the predictive ability of the emergent network. These results are shown in Table 2. The biggest improvement in model fit is found using the recovery period data for only oil-related products with a 2-year prediction horizon, followed by a 1-year prediction horizon. We hypothesize that this may be due to the relative stability of other sectors with respect to oil prices experienced during the post-GFC of 2014-2016 (Mensi, 2019). Figures 2 and 3 show the regression results for that largest improvement (Table 2, bold).

4. Conclusion

The goal of this paper is to explore the concept of causal emergence as it relates to definitions of causal abstraction, the properties needed for accurate state predictions, and the semantic meaning behind emergent macro-nodes of the bilateral trade network. To our knowledge, we are the first to explore these connections between causal abstraction, causal emergence, and prediction. In Section 2, we established the correspondence between definitions of causal abstraction and causal emergence and discussed how causally abstract (or emergent) representations should theoretically impact prediction accuracy. In Section 3.2, we outlined cases where causal emergence of trade agreement macro-nodes do or do not

```
                            OLS Regression Results
==============================================================================
Dep. Variable:           export_val_2   R-squared:                       0.240
Model:                            OLS   Adj. R-squared:                  0.239
Method:                 Least Squares   F-statistic:                     351.9
Date:                Thu, 22 Apr 2021   Prob (F-statistic):               0.00
Time:                        10:28:32   Log-Likelihood:                -17204.
No. Observations:               13420   AIC:                         3.443e+04
Df Residuals:                   13407   BIC:                         3.453e+04
Df Model:                          12
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.2073      0.010    -20.003      0.000      -0.228      -0.187
comm_neigh     0.2367      0.018     13.376      0.000       0.202       0.271
indeg_orig    -0.5540      0.094     -5.903      0.000      -0.738      -0.370
indeg_dest     0.4417      0.088      5.031      0.000       0.270       0.614
outdeg_orig   -0.1707      0.017     -9.864      0.000      -0.205      -0.137
outdeg_dest   -0.1201      0.024     -4.989      0.000      -0.167      -0.073
close_orig     0.6958      0.101      6.861      0.000       0.497       0.895
close_dest    -0.2393      0.097     -2.477      0.013      -0.429      -0.050
bet_orig       0.1400      0.024      5.882      0.000       0.093       0.187
bet_dest       0.0147      0.023      0.629      0.530      -0.031       0.061
pr_orig       -0.0328      0.011     -2.962      0.003      -0.055      -0.011
pr_dest        0.1909      0.012     16.310      0.000       0.168       0.214
same_comm      0.4659      0.016     29.110      0.000       0.434       0.497
==============================================================================
Omnibus:                      240.509   Durbin-Watson:                   1.470
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              156.397
Skew:                          -0.133   Prob(JB):                     1.09e-34
Kurtosis:                       2.543   Cond. No.                         46.9
==============================================================================
```

Figure 2: Regression Results for Predicting Future Bilateral Trade from G_{micro}, Oil Products only, 2013-2017

```
                            OLS Regression Results
==============================================================================
Dep. Variable:           export_val_2   R-squared:                       0.483
Model:                            OLS   Adj. R-squared:                  0.476
Method:                 Least Squares   F-statistic:                     75.95
Date:                Sat, 24 Apr 2021   Prob (F-statistic):          9.82e-131
Time:                        16:08:23   Log-Likelihood:                -1078.0
No. Observations:                 990   AIC:                             2182.
Df Residuals:                     977   BIC:                             2246.
Df Model:                          12
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1204      0.032     -3.806      0.000      -0.182      -0.058
comm_neigh     0.3328      0.047      7.130      0.000       0.241       0.424
indeg_orig     0.3057      0.183      1.674      0.094      -0.053       0.664
indeg_dest     0.0097      0.173      0.056      0.955      -0.329       0.349
outdeg_orig    0.2197      0.060      3.640      0.000       0.101       0.338
outdeg_dest    0.1440      0.073      1.964      0.050       0.000       0.288
close_orig    -0.2177      0.186     -1.167      0.243      -0.584       0.148
close_dest     0.0261      0.177      0.148      0.883      -0.321       0.373
bet_orig       0.0103      0.070      0.147      0.883      -0.127       0.147
bet_dest      -0.0395      0.073     -0.545      0.586      -0.182       0.103
pr_orig        0.1801      0.049      3.671      0.000       0.084       0.276
pr_dest        0.0836      0.053      1.586      0.113      -0.020       0.187
same_comm      0.2626      0.047      5.544      0.000       0.170       0.355
==============================================================================
Omnibus:                       15.650   Durbin-Watson:                   1.685
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               16.118
Skew:                          -0.303   Prob(JB):                     0.000316
Kurtosis:                       2.849   Cond. No.                         30.1
==============================================================================
```

Figure 3: Regression Results for Predicting Future Bilateral Trade from G_{macro}, Oil Products only, 2013-2017

occur. Finally, in Section 3.3, we showed that the automatically-identified macro-level network representations (via the greedy algorithm) are indeed able to provide better prediction capability than the original trade networks in a variety of cases. As such, we contend that this work is one step towards the broader goal of developing automated algorithms to find optimal and efficient abstract representations that maintain the faithfulness of causal relationships for improved prediction of future states from large, noisy, complex data sources. These automated algorithms overcome drawbacks associated with hand-crafted abstractions which, although easily interpreted and accurate, do not generalize to the larger and novel datasets commonly collected today. On the other hand, automated abstraction methods may produce uninterpretable results, as evidenced by the discrepancy between the expected and actual results of the greedy algorithm (described in Section 3.2). Our future work focuses on novel methods which narrow the gap between these two types of abstraction. The applications of such methods extend beyond economic trade networks to any domain in which data can be represented as a causal graph, such as vehicle health maintenance and digital twin systems, decision-making sciences, and social influence networks.

Acknowledgments

We would like to thank Dr. Brennan Klein for discussions clarifying his implementation of the causal emergence algorithm.

References

Angel Aguiar, Maksym Chepeliev, Erwin L Corong, Robert McDougall, and Dominique van der Mensbrugghe. The gtap data base: Version 10. *Journal of Global Economic Analysis*, 4(1):1–27, 2019.

Sander Beckers and Joseph Y Halpern. Abstracting causal models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 2678–2685, 2019.

Sander Beckers, Frederick Eberhardt, and Joseph Y Halpern. Approximate causal abstraction. In *Uncertainty in artificial intelligence: proceedings of the... conference. Conference on Uncertainty in Artificial Intelligence*, volume 2019. NIH Public Access, 2019.

Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

Krzysztof Chalupka, Frederick Eberhardt, and Pietro Perona. Causal feature learning: an overview. *Behaviormetrika*, 44(1):137–164, 2017.

Nicolas Doremus, Alessio Moneta, and Sebastiano Cattaruzzo. Cause-effect pairs in time series with a focus on econometrics. In *Cause Effect Pairs in Machine Learning*, pages 191–214. Springer, 2019.

Robert C Feenstra, Robert E Lipsey, Haiyan Deng, Alyson C Ma, and Hengyong Mo. World trade flows: 1962-2000. Technical report, National Bureau of Economic Research, 2005.

L Freeman. Centrality in networks: I. conceptual clarifications. social networks. *Social Network*, 1979.

Guillaume Gaulier and Soledad Zignago. Baci: international trade database at the product-level (the 1994-2007 version). 2010.

Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438, 1969.

Ross Griebenow, Brennan Klein, and Erik Hoel. Finding the right scale of a network: Efficient identification of causal emergence through spectral clustering. *arXiv preprint arXiv:1908.07565*, 2019.

Scott Wayne Hall, Michael Reed, et al. Us-brazil bilateral fuel ethanol trade. In *2019 Annual Meeting, February 2-5, 2019, Birmingham, Alabama*, number 284278. Southern Agricultural Economics Association, 2019.

Erik P Hoel, Larissa Albantakis, and Giulio Tononi. Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences*, 110(49): 19790–19795, 2013.

Paul Hünermund and Elias Bareinboim. Causal inference and data-fusion in econometrics. *arXiv preprint arXiv:1912.09104*, 2019.

Elena Ianchovichina, Shantayanan Devarajan, and Csilla Lakatos. *Lifting economic sanctions on Iran: global effects and strategic responses*. The World Bank, 2016.

Bogang Jun, Aamena Alshamsi, Jian Gao, and César A Hidalgo. Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade. *Journal of Evolutionary Economics*, pages 1–31, 2019.

Brennan Klein and Erik Hoel. The emergence of informative higher scales in complex networks. *Complexity*, 2020, 2020.

Zuzanna Kosowska-Stamirowska. Network effects govern the evolution of maritime trade. *Proceedings of the National Academy of Sciences*, 117(23):12719–12728, 2020.

Walid Mensi. Global financial crisis and co-movements between oil prices and sector stock markets in saudi arabia: A var based wavelet. *Borsa Istanbul Review*, 19(1):24–38, 2019. ISSN 2214-8450.

Alessio Moneta, Nadine Chlaß, Doris Entner, and Patrik Hoyer. Causal search in structural vector autoregressive models. In *NIPS Mini-Symposium on Causality in Time Series*, pages 95–114, 2011.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

Claude E Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

James Stanford. Continental economic integration: Modeling the impact on labor. *The Annals of the American Academy of Political and Social Science*, 526(1):92–110, 1993.

Chi-Wei Su, Yu Song, Ran Tao, and Lin-Na Hao. Does political conflict affect bilateral trade or vice versa? evidence from sino-us relations. *Economic Research-Ekonomska Istraživanja*, pages 1–20, 2019.