Supplemental Material of "Modeling Dynamic Transport Network with Matrix Factor Models: an Application to International Trade Flow"

Elynn Y. Chen[1] and Rong Chen[2]

[1]New York University, Stern School of Business
[2]Rutgers University, Department of Statistics

A International Trade Data: Exploratory Analysis

The dynamic trading network can be cast into a time series of relational matrices that record the ties (trading volumes) between the nodes (countries) in the network. The length of our network matrix time series is 408 months. At each time point, the observation is a square matrix whose rows and columns represent the same set of 24 countries. Each row (column) corresponds to an export (import) country. Each cell in the matrix contains the dollar trading volume that the exporting country exports to the importing country. The diagonal elements are undefined.

Figure 5 plots the time series of trading volumes in U.S. dollar among top 13 countries from January, 1982 to December, 2015 in our dataset. Each time series is normalized for ease of visualization. These 13 countries are representative of all countries and regions in our dataset. They falls into three major groups: Canada, Mexico, and United States compose the NAFTA group; France, Germany, Italy, Spain, and United Kingdom are in the EU group; Australia, China Mainland, Indian, Japan and Korea belong to the APEC group. Overall, all countries experienced rapid growth in trades along with the accelerating wave of globalization. The world saw largest collapse in the value of good traded in 2009 when the impact of the global financial crisis was at its worst. Some actually have not recovered yet. For example, we see that Spain's downturn in import has not recovered so far, though its export has mostly recovered. While the upward trends are shared among all countries, the pattern of trading are more alike among countries within the same group. For example, the exports time series of the five European countries resembles more to each other than to the exports time series of the Asian countries.

In order to illustrate the pattern of bilateral relationships, a set of four circular trading plots are shown in Figure 6. The direction of flow is indicated by the arrowhead. The size of the flow is shown by the width of the arrow at its base. Numbers on the outer section axis, used to read the size of trading flows, are in billions. Each plot is based on the monthly flows over a one-year period, aggregated to selected annual volumes. Note that the four plots are representative of the bilateral relationship patterns in the 1980's, 1990's, 2000's and 2010's.

For the three groups (EU, NAFTA, and APEC), most of the trade flows occur within the same group. This phenomenon is most prominent within the EU group where the imports and exports are all in red shade that denotes EU countries in Figure 6. The trade flows of NAFTA countries are least confined within the group, mainly because the U.S. alone trades a lot with both EU and APEC countries.

For individual countries, most noticeable are changes in the share and direction of trade of U.S., China Mainland, Mexico and Japan. Over the years, U.S. maintains the most distinctive one among all countries because of its large trading volumes and wide range of trading counterparties. The destinations of U.S. exports gradually shift from Japan and European countries to China Mainland and Mexico. In the 1980's Japan accounted for the largest importing and exporting flow among APEC countries. As shown clearly in Figure 6, China Mainland's slice of pie in global trades grew steadily in size and becomes the largest in the 2010's. Mexico experienced

Figure 5: Time series plots of the value of good traded among 13 countries over 1982 – 2015. The plots only show the patterns of the time series while the amplitudes are not comparable between plots because the range of the y-axis are not the same.

a similar steady growth in global trades although less prominent than that of China Mainland. The trading patterns are most stable of the EU countries. The EU countries almost keep the same portions in the size of imports and exports over years.

The explanatory statistical analysis and visualization tools provide very clear and powerful but only descriptive observations. It is clear that there exists a possibly lower dimensional latent network, underlying the large scale dynamic network on the surface. However, there are few statistical tool available to quantify this latent structure. In the next section, we present a new methodology that is able to quantify the latent dynamic networks that underpins the observed surface dynamic networks as well as the relationship that connect the latent networks and the surface networks.

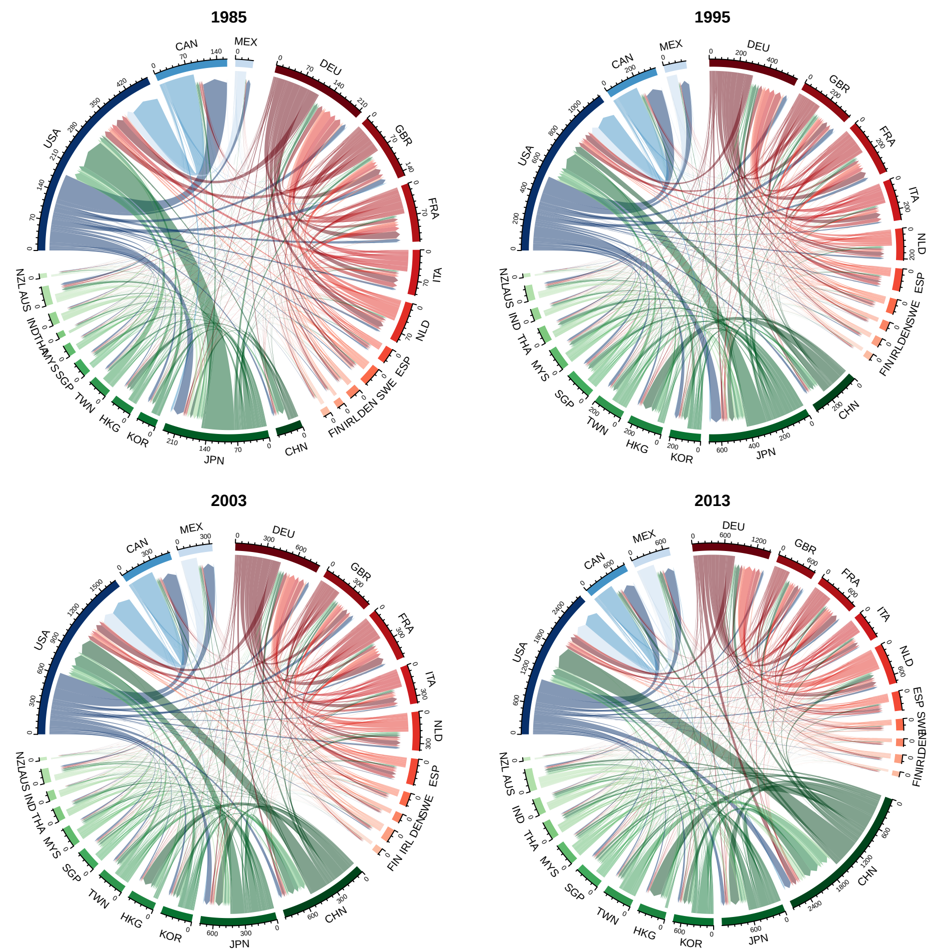

Figure 6: Circular trading plots that are representative of the bilateral relationship patterns in the 1980's, 1990's, 2000's and 2010's. The arrowhead indicates the direction of exports. The width of the arrow at its base represents the size of trade flow. Numbers on the outer section axis correspond to the size of trading flows in billion dollars.

B Asymmetric Export and Import Loadings

Now we apply Model (2.2) to the international trade volume data. We use the ratio-based method in (3.11) as well as scree plot to estimate the number of latent dimensions. The comparison between these two methods of estimating importing and exporting dimensions in different time periods is shown in Table 2. Note that Model (2.2) assumes different exporting and import loadings $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$. Similar to Figure 1 , the scree plot method selects the minimal number of dimension that explain at least 85 percent of the variance in the original data. The percentage of total variance explained by the 4×4 factor model is shown in the last line. With the additional flexibility of allowing different row and column loading matrix, the estimated dimension is slightly smaller than that in Table 1, though the ratio estimate becomes less stable.

As shown in Table 2, most dimension estimators are smaller than 4 and the factor model with dimension 4×4 explains at least 85% of the total variance. Thus, latent dimension 4×4

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Ratio	(1, 1)	(1, 1)	(8, 1)	(1, 1)	(11, 1)	(6, 1)	(6, 3)	(2, 1)	(2, 1)	(2, 2)
Scree	(2, 2)	(2, 2)	(3, 3)	(3, 3)	(4, 4)	(5, 4)	(4, 4)	(4, 4)	(3, 3)	(3, 3)
(4,4)	(98, 98)	(95, 96)	(92, 94)	(91, 92)	(85, 91)	(85, 90)	(88, 89)	(91, 90)	(94, 93)	(95, 94)
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Ratio	(5, 2)	(2, 2)	(2, 2)	(2, 2)	(2, 2)	(1, 1)	(1, 1)	(1, 1)	(1, 1)	(1, 1)
Scree	(3, 3)	(3, 3)	(2, 2)	(2, 2)	(2, 2)	(3, 3)	(3, 3)	(3, 3)	(3, 3)	(3, 3)
(4,4)	(93, 92)	(95, 94)	(96, 95)	(97, 97)	(96, 96)	(94, 94)	(92, 92)	(93, 94)	(95, 95)	(93, 93)
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Ratio	(1, 1)	(6, 6)	(1, 1)	(6, 6)	(1, 6)	(1, 6)	(1, 5)	(5, 5)	(7, 1)	(1, 1)
Scree	(3, 3)	(3, 3)	(3, 3)	(4, 4)	(4, 4)	(4, 3)	(3, 3)	(3, 3)	(3, 3)	(3, 3)
(4,4)	(94, 93)	(93, 93)	(91, 91)	(88, 89)	(88, 91)	(89, 91)	(92, 93)	(94, 94)	(95, 95)	(90, 91)

Table 2: Comparison of estimated latent dimension of $\boldsymbol{F}_t$ in Model (2.2) between ratio-based and scree plot methods. The last line presents the percentages of variance explained by the 4×4 factor model in (export, import), respectively.

will be used for illustration and comparison between different period. In the following analysis, we employ the same visualization tools as those in Section 4. However, there are separate plots for loading matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ since Model (2.2) differentiates the importing and exporting dimensions.

Figures 7 and 8 present the heat maps for exporting loading $\boldsymbol{A}_1$ and importing loading $\boldsymbol{A}_2$, respectively. They are designed in the same way as those in Figure 2. The patterns are strikingly similar in the heat maps of $\boldsymbol{A}$, $\boldsymbol{A}_1$, and $\boldsymbol{A}_2$. Plots (a) in all three Figures 2, 7 and 8 represent the latent hub of United States. Plots (b), (c), and (d) in all figures represent the latent hub of European countries, Japan/China Mainland, and NAFTA countries (except US), respectively. The loadings of countries on these top four latent hubs evolve in the same way among these three figures.

There are a few noticeable differences in the import and export behavior though. For example, US's import activities dominate the import hub #1 throughout the period, but its export activities weaken in the export hub #1 during the 2000's, facing competition from the Asian countries. China's export activities start in the early 1990's but it's import activities only show dominance in the 2000's.

Figure 9 plots the trading network among four latent hubs as well as the relationship between countries and latent hubs for four selected years. Since we use different export (left) and import (right) loading matrix, the relationships between countries and latent hubs are different for import and export activities. The meanings of row and column dimensions of the latent factor matrix $\boldsymbol{F}_t$ are different too. Specifically, the rows of $\boldsymbol{F}_t$ represents the exporting hubs while the columns correspond to the importing hubs. Thus we distinguish the row and column hubs and have eight circle nodes for the latent hubs in Figure 9. The nodes annotated with "Ex" and "Im" correspond to the export (row) hubs and the import (column) hubs, respectively. We notice symmetry between the exporting and importing nodes or hubs, indicating empirically the validity of Model (2.1), for certain years. For example in 1995, the exporting node "Ex1" and importing node "Im1" both represent the United States hub; the exporting node "Ex2" and importing node "Im2" both represent the Europe hub; and the exporting node "Ex3" and importing node "Im3" both represent the Asia hub;the exporting node "Ex4" and the importing node "Im4" both represent the Canada & Mexico hub. However, in this paper we do not devise a formal statistical

Figure 7: Latent export loadings for trading volume on $r = 4$ hubs for a series of 30 rolling five-year periods indexed from 1984 to 2013.

Figure 8: Latent import loadings for trading volume on $r = 4$ hubs for a series of 30 rolling five-year periods indexed from 1984 to 2013.

method for testing Model (2.1) and (2.2), which is an important problem for future research.

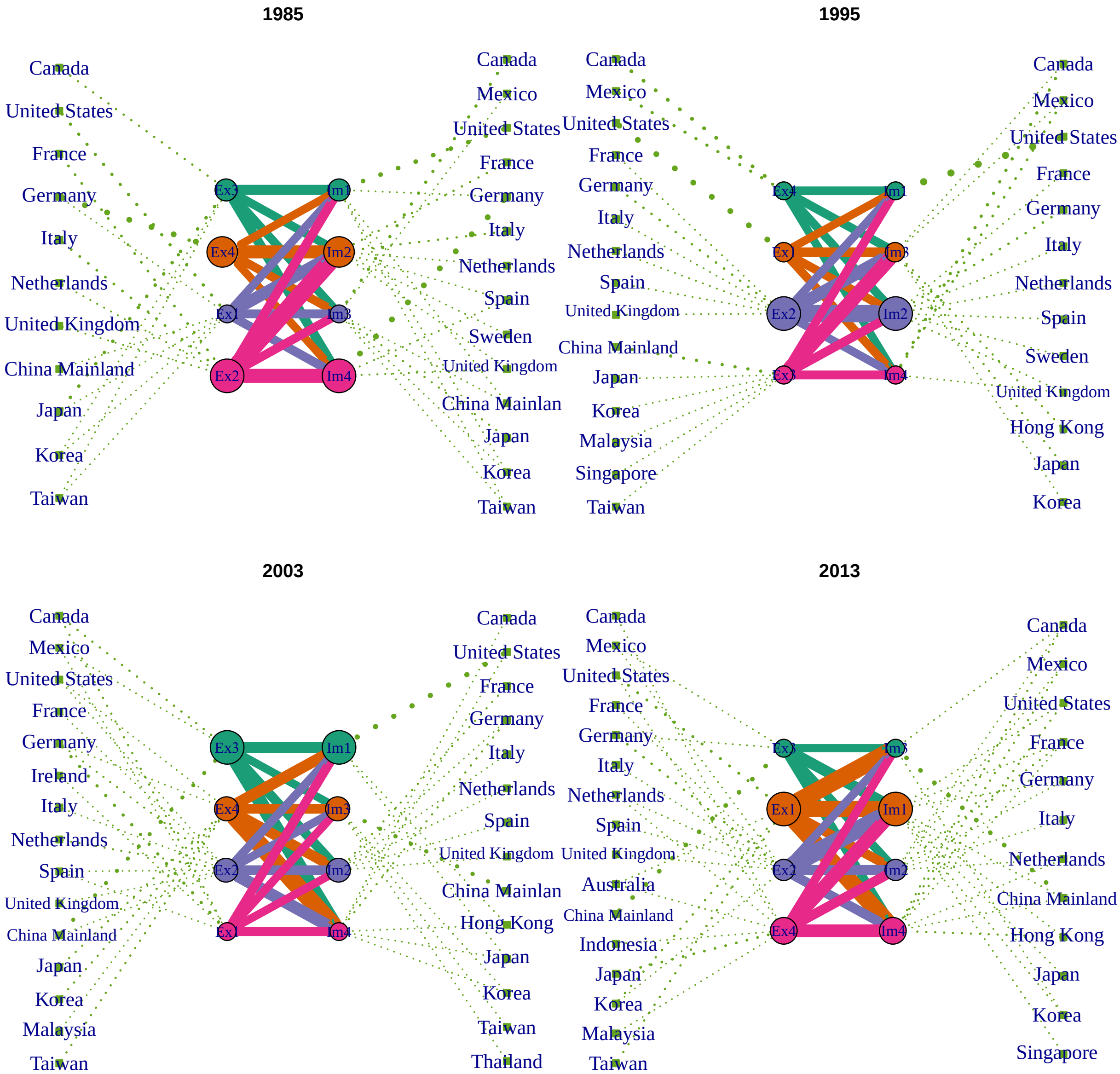

Figure 9: Trading volume network plot of latent hubs and relationship between countries and the latent hubs. Thickness of the solid line represents the volume of trades among latent hubs. Thickness of the dotted lines represents the level of connection between latent hubs and countries. Note that a country can be related to multiple latent hubs. To provide a clear view, $\widehat{\boldsymbol{A}}_1$, and $\widehat{\boldsymbol{A}}_2$ are truncated by rounding $10\widehat{\boldsymbol{A}}$ and then normalizing the non-zero entries to have column sum one.

Finally, we apply a hierarchical clustering algorithm (Xu and Wunsch, 2005; Murtagh and Legendre, 2014) to cluster countries based on their contribution patterns over years under Euclidean distance and the *ward.D* criterion. The dendrograms in Figure 10 are obtained by applying on $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively, the same method that is used to get Figure 4. Similar to the symmetric case, we observe that geographically or culturally proximate countries are usually in the same group and behave similarly, and countries with similar trading behaviors also tend to be clustered in the same group. For example, one can easily identify the European group and the Asia-pacific group from the dendrograms and that, in the 2000's, Canada and Germany are in the same group – both trade in large volumes to United States and China. The overall structure of international trading seems steady over years: fours groups in all years can be labeled as 'United States', 'European active', 'Asia-pacific active', 'European-Asia-pacific less active'. However, the relationship between individual countries are changing over the period. The patterns of evolution resonate to some of the observations from Figures 7, 8 and 9.

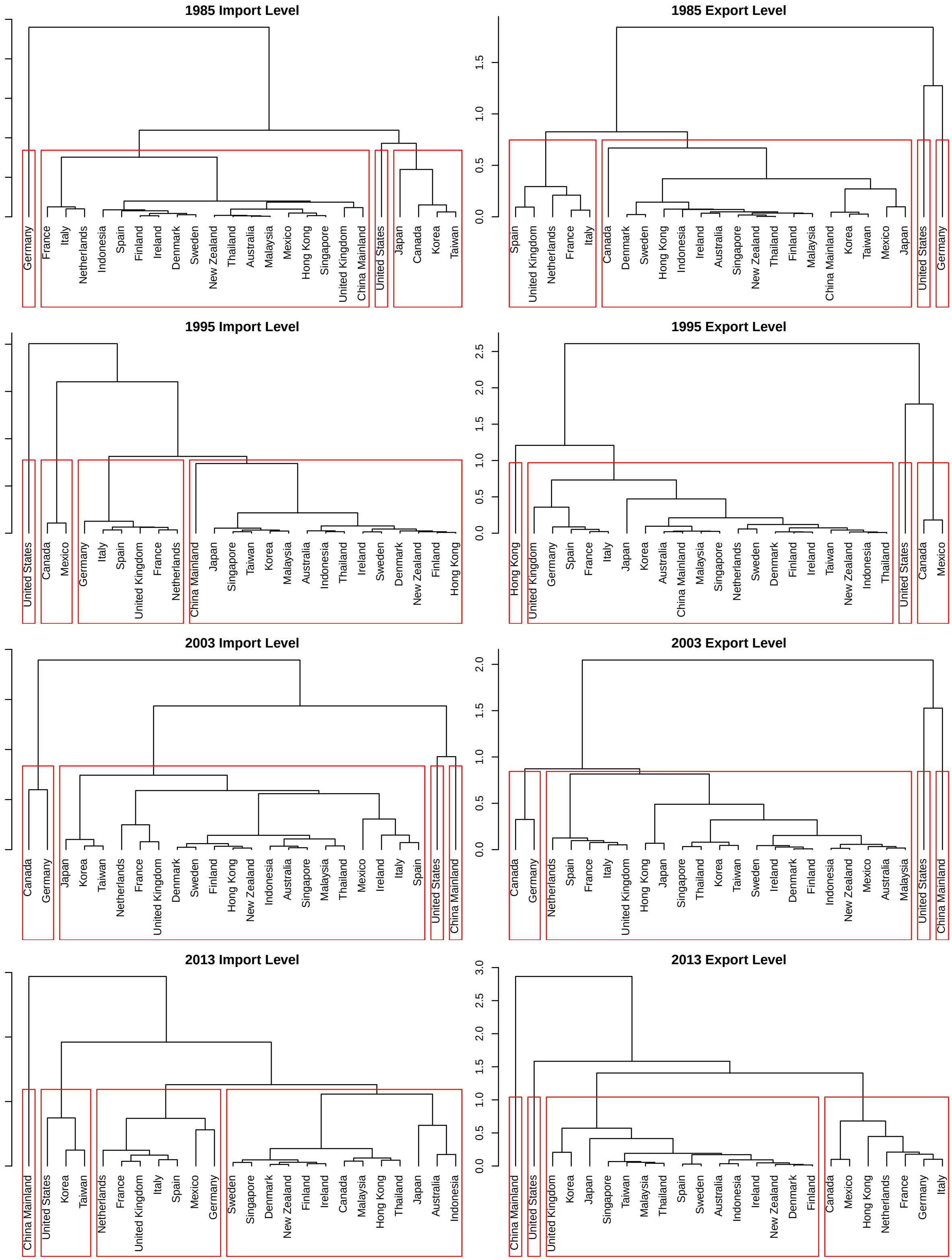

Figure 10: Clustering of countries based on their trading volume latent hub representations.