Disaggregating the United States GTAP region into 51 US-state subregions

Amanda M. Countryman, Joseph Francois and Hugo Rojas-Romagosa

April 2017

1 Introduction

The opening of the North Western Route (NWR) that crosses the Canadian Arctic or the expansion of the Panama Canal will reduce the trade costs faced by the United States (US) economy. However, these trade cost reductions will be asymmetric between different US subregions, in particular between the East and West Coasts of the US.

For instance, both the NWR and the Panama Canal expansion will diminish the trade costs between Asia and the South and East coasts of the United States, while the trade costs between Asia and the West Coast will not be affected. On the other hand, the trade costs between Europe and the West Coast will change, while those between Europe and the East Coast will not.

To properly analyze the impact of these asymmetric trade cost changes in a global economic modeling framework, the US must be disaggregated into subregions. Accordingly, the objective of this research is to assess the difference between trade cost reductions employed in a model where the US is included as one region and in a framework that disaggregates the US into key port regions to illustrate the importance of regional disaggregation in some cases. To accomplish this, we first carefully document the disaggregation of the US into state level regions in the GTAP database followed by two relevant policy experiments that highlight the impact of disaggregation on simulation results.

In sections **??** and **??**, we explain how using the GTAP9.1 global multi-region input-output database, we divide the US region into 51 subregions: the 50 US States plus the District of Columbia (DC). The 51 subregions can then be aggregated into five "port regions" that correspond to five main US ports: Seattle, Washington in the Northwest; Long Beach, California in the Southwest; New Orleans, Louisiana in the South; Charleston, South Carolina in the Southeast and New York/New Jersey in the Northeast. The decision to first disaggregate to the 51 States and then up to the five port regions, was based on the availability of data at the US state level and for the flexibility to use the 51 state-level disaggregation for future studies that require a different subregional aggregation of the US. Section **??** presents applications employing the new database providing relevant comparisons to results that model the US as an aggregate region. Section **??** provides a summary and discusses the relevance of this work for economic modeling.

2 GTAP database and SplitReg software

There is no state-level input-output table for the United States.[1] Without these primary data, a piece-meal procedure is needed to estimate state-level macroeconomic and trade flows, which are needed to disaggregate the US into 51 sub-regions.

To fully disaggregate the US, the first step is to use the GTAP9.1 database with base year 2011, which has 141 regions.[2] These regions are divided between 121 individual countries and 20 "residual" regions (countries without primary IO data that are grouped into regional aggregates) for a total of 122 countries/regions. The United States is one region within the database, yet our work disaggregates the US to state-level regions by using state-level data from the Bureau of Economic Analysis (BEA) and the Department of Commerce in tandem with a software program that facilitates regional splitting within the GTAP framework.

When we expand the United States from 1 region into 51 regions (50 States plus the District of Columbia), the database is expanded to 172 regions: 120 individual countries, plus RoW (aggregated residual GTAP regions), plus 51 US subregions. The SplitReg program (Horridge, 2011b) allows for the disaggregation of any GTAP region into multiple regions within the GTAP database using as much detail as the modeler determines. At a minimum, data for value added by sector for each new subregion are used to split industry and final demand by sector for each new subregion, imposing the same technology and budget shares as the original region. A second program, GTAPAdjust (Horridge, 2011a) ensures that the resulting new database is balanced, and also allows for different trade, tax or technology patterns to be imposed exogenously by subregion. Accordingly, we use state-level sectoral value added from BEA, followed by state-level international trade data from the Department of Commerce, as well as interstate trade data from the Commodity Flow Survey to create a balanced database with 51 new distinct state-level regions that replace the US in the GTAP database.

[1]The Bureau of Economic Analysis (BEA) of the US Department of Commerce has a Regional Input-Output Modeling System (RIMS II), which is used for regional analysis, but it is not a State-level IO table. To create RIMS II, the national IO tables are used, which are adjusted for regional leakages with location quotients. However, there are no sub-national data available for industry inputs, which are needed to create a State-level IO table.

[2]See Appendix **??**.

3 Procedure to split the US region into 51 state-level regions

In the following sections we detail the procedure to split the US into 51 regions and adjust the database.

3.1 US state-level macroeconomic data

First, we use the SplitReg software to disaggregate the main macroeconomic US variables (e.g. consumption, production, taxes, and trade) into 51 US states using the 2011 BEA US-State level value-added (VA) data as weights for 31 sectors. [3] The number of sectors (31) is an upper-limit imposed by the BEA State-level VA data, which cannot be fully matched to the 57 sectors in the GTAP database. This initial split procedure creates the GTAP9_USdisag database.

With this first step of disaggregation, we have a balanced GTAP9.1 database created with 172 regions. The new database includes the original 120 individual GTAP countries, except the US, plus a RoW composite of the 20 original "residual" regions, and 51 US-States. This database can be aggregated to 5 US subregions, each of which is assigned a main port of entry: Northeast (NJ/NY), Southeast (Charleston), South (Louisiana), Southwest (Long Beach) and Northwest (Seattle).

To use the BEA data, we created a concordance between BEA sectors (which use the North American Industry Classification System (NAICS)) and GTAP sectors to link the 31 BEA sectors to a GTAP aggregation comprised of 31 sectors. The output is "GDP_byUS_states.xlsx" that gives GDP (value added) for 51 states (50 US States plus the District of Columbia) and 31 and sectors.

3.2 US inter-State trade

The second step in the disaggregation process is to include US interstate trade data. The values for interstate trade are equal to zero after the initial regional split in step 1, because the original GTAP database value for trade between the US and itself is equal to zero (the variables VXWD(US,US) and VIWS(US,US)). In other words, the import and export data to and from the same region is zero for the US, which is the case for GTAP countries except for "residual" or composite regions.[4]

In order to create a final database that contains interstate trade, we must first manually alter and re-balance the original, starting database to account for the total value of interstate economic activity between US regions as the value for US-US trade in the initial database. To populate the US interstate trade data we use the Commodity Flow Survey (CFS) from the US Census Bureau. The CFS provides information for trade in goods (agriculture and manufacturing sectors); however,

[3] We use 2011 data to correspond with the GTAP 9.1 baseyear of 2011.

[4] In particular, the BEA sector-state value-added value weights are multiplied by the value of US-US trade in the initial database, which is equal to zero in the first step of the SplitReg procedure, and the resulting US-State to US-State data for economic activity is subsequently equal to zero in the split database.

there are no interstate services trade data available for the US, according to the BEA. Therefore, we use proportionality assumptions to estimate interstate trade in services, based on the US-wide proportion of services trade in total trade.

3.2.1 Internal US trade in goods

In this section, we describe the use of the Commodity Flow Survey (CFS) to obtain the interstate trade flows for agricultural and manufacturing goods. The CFS provides information on commodities shipped, monetary value, weight, mode of transportation, as well as the origin and destination of shipments of commodities. Thus, CFS provides detailed data on the movement of goods in the United States.

The CFS data uses both the North American Industry Classification System (NAICS) and the SCTG (Standard Classification of Transported Goods) codes. We use a combination of both codes, to obtain a conversion to the 31 GTAP sectors that we use in our overall disaggregation.[5] This allows us to determine which product (by sector) is traded between two US states –i.e. interstate trade.

We use the CFS database for 2012, which has more than 4.5 million observations.[6] We aggregate the CFS over US states and 31 sectors, and subtract within-state trade. This leaves approximately 37,500 bilateral interstate trade observations. Total shipments amount to 13.9 US$ trillion, of which 6.5 trillion (46.7% of the total) represents interstate trade. Compared to GDP at current US$ for 2012 (approximately 16.2 US$ trillion, according to BEA), we conclude that internal US trade in goods will represent approximately 40% of GDP. To make the CFS data comparable with the GTAP data, we scale down the 2012 interstate trade to 2011 values using GDP nominal growth (from the BEA) between 2011 and 2012, which gives a factor of 0.9605.

3.3 Internal US trade in services

For the US there are no interstate services trade data available, but different assumptions can be used to identify inter-regional service imports in the literature (Guci and Mead, 2014). The CFS database, for instance, only gives limited information on retail trade and other business services, but there is no information for other services sectors.

Given the lack of data, we use an approximation to estimate interstate trade in services data: we use the assumption that interstate trade is proportional to international US trade in services values. In particular, we use the trade data from the GTAP 9.1 database where US primary sector exports are 6.2% of total US

[5]Inspection of the data reveals that the NAICS is used to identify the destination sector and the SCGT the sector of origin. Since we need information on the origin/nature of the product being traded (e.g. an agricultural good used in another industry), we chose to use the SCTG as the main sectoral identifier. However, when SCTG does not identify the sector (SCTG=0), then the NAICS sector identifier is used.

[6]The CFS has been conducted every five years as part of the Economic Census, in 1993, 1997, 2002, 2007, and the most recent is in 2012.

exports, and the corresponding value for manufacturing is 70.3%, and 23.4$ for services exports. Using the services sector-specific shares in conjunction with the interstate exports in non-services sectors from the CFS database, we obtain the interstate trade in services values.[7]

Table 1 shows the final US trade values when we merge the CFS interstate non-services trade data and our estimated interstate services trade, with the GTAP9.1 trade data. From the GTAP9.1 database, total US exports abroad in 2011 are equal to 1.7 billion US$, of which 6.5% is from primary sectors, 69.1% from manufacturing sectors and 24.4% from services. Internal (inter-US) trade is approximately four times higher than external trade (extra-US), and this holds for all three main sectors analyzed. On the other hand, the sectoral composition of inter-US exports is very similar to extra-US, since the same external trade shares are employed to construct the internal services trade approximations.

3.4 Rebalancing internal US national output

Since the initial interstate trade data for the US as a whole is zero in the GTAP database, when we assign non-zero values to the state-level trade matrix, this will unbalance the input-output tables. For instance, total production (VOM) must equal internal production (VDM) plus exports (VXMD). When we disaggregate the US into separate states, then interstate exports will be added to VXMD and thus, must be subtracted from internal production (VDM).

We use the following procedure to rebalance the GTAP database:

- We assign the US (aggregate) internal trade value to all trade variables in the GTAP base year data source. For the internal US trade data, we assume that world prices are equal to domestic prices (i.e. there are no internal US export taxes), and also that US to US imports are equal to US to US exports. Under these conditions, all the GTAP trade variables are equal (VIMS=VIWS=VXMD=VXWD) for internal US trade flows.

- We subtract these US_dom values from the internal production values (VDM), such that the equality of total output (VOM) is preserved: VOM =VDM + VXMD. Furthermore, since VDM is the sum of private household consumption (VDPM), government consumption (VDGM) and intermediate input consumption by firms in specific sectors (VDFM), we subtract exports from the components of VDM in a way that is proportional to the weight of each component.

[7] Even though the CFS provides some data on retail trade (TRD) and other business services (OBS), we do not use this data as it does not wholly represent all services. Therefore, we estimate TRD and OBS trade values in the same way as other services sectors.

- In a few sectors, we find that the CFS interstate (internal) exports are higher than US domestic production.[8] To solve this problem, we limit internal US trade in these sectors to be at most 95% of total domestic production.[9]

To assess how reasonable the estimations are for interstate US trade, we compare the US interstate trade data with the EU25 region.[10] We observe that the sectoral shares of EU25 exports differ from the shares we obtain for the U.S. when comparing extra- and inter-EU exports to corresponding values for the US. These results lend credibility to our proportionality assumption methodology. Moreover, when assessing the horizontal shares by sector, the EU has much smaller inter-EU trade shares (around 60%) than the corresponding US shares (around 75%). The contrast is particularly noticeable in the services sector where inter-EU trade is half of total trade in services, while inter-US trade in services is 76% of the total. This observations confirms the prior that the US is more integrated that the EU. In other words, there is still some scope for further EU integration, especially in services.

3.5 External US trade at the state-level

SplitReg assigns the values for total US exports by State using sectoral value-added as weights in the initial split, but this does not map correctly to observed external (international) trade by state-sector pairs, since external trade in goods is mainly linked to international ports, which are not present in most States. Thus, we improve the initial US-split by using detailed US district-level international bilateral trade data from the US International Trade Commission that was obtained from the Department of Commerce USA Trade Online Database (2014).[11] In particular, we use 2013 US international trade for 42 districts that can be assigned to US subregions by GTAP sector and international trade partner outside the US. As expected, some interior states do not have districts assigned. For import data (into the US), we used the data pre-processed by USITC that assigned source-specific imports to GTAP sectors and US subregions. The district-level US export data was not pre-processed by USITC, so we followed the same procedure as the import data processing, and converted the bilateral export data from HS6 codes to GTAP

[8] These sectors are: agriculture (AGR), Chemicals and plastic products (CRP), Metals (MET), Motor vehicles and parts (MVH), Electronics (ELE) and Other manufactures (OMF). There could be several reasons for this discrepancy: the GTAP global re-balancing adjustments that in some cases change reported country data, the use of two different years (2011 and 2012) can overlook year to year structural changes in the export composition, and measurement errors with the CFS database.

[9] We also then apply the same proportionality assumption estimations from the previous section to be consistent. This means that for these adjusted sectors, we use the internal US trade shares from the CFS data and not the original values.

[10] EU25 excludes Croatia, Bulgaria and Romania, which were the last countries to join the EU. Croatia joined in 2013 and thus, was not an EU country in the 2011 baseline year. Bulgaria and Romania joined in 2007 and were just starting the trade integration process with the rest of the EU in 2011.

[11] Data are available here:
https://www.census.gov/foreign-trade/reference/products/catalog/usatradeonline.html.

Table 1: United States and the EU25, external and internal exports by main sector, million US$

	United States			*vertical shares*		
	total	extra-US	intra-US	total	extra-US	intra-US
Total exports	7,029.7	1,771.6	5,258.1			
horizontal shares		25%	75%			
Primary sectors	439.4	114.9	324.6	6.3%	6.5%	6.2%
horizontal shares		26%	74%			
Manufacturing sectors	4,783.3	1,224.5	3,558.8	68.0%	69.1%	67.7%
horizontal shares		26%	74%			
Services sectors	1,807.0	432.3	1,374.7	25.7%	24.4%	26.1%
horizontal shares		24%	76%			
	EU-25			*vertical shares*		
	exports	extra-EU	intra-EU	total	extra-EU	intra-EU
Total exports	6,399.1	2,699.4	3,699.7			
horizontal shares		42%	58%			
Primary sectors	189.0	56.5	132.4	3.0%	2.1%	3.6%
horizontal shares		30%	70%			
Manufacturing sectors	4,768.8	1,919.5	2,849.3	74.5%	71.1%	77.0%
horizontal shares		40%	60%			
Services sectors	1,441.3	723.3	718.0	22.5%	26.8%	19.4%
horizontal shares		50%	50%			

Notes: Horizontal shares are the proportion of intra and extra exports to total exports in each sector. Vertical shares are the proportion of each sectors' exports in the total extra or intra exports.

Source: Own estimations using the CFS and GTAP9 databases.

sectors using a HS6 to GTAP concordance. Finally, the state-level shares of total US imports and exports were used to disaggregate US international trade into each subregion. We also created a concordance to map the Department of Commerce database countries to the GTAP regions.

To add the external trade data to the GTAP9_USdisag database, we took the following steps:

1. We assign the trade data shares from districts to US-States.[12]

2. We obtained the matrix with the shares of imports into US states (by state of entry) and the exports from US states (by State of exit).

3. These trade shares are then imposed exogenously into the SplitReg procedure so the final trade values reflect the external trade structure from the Department of Commerce trade data.

There are two caveats regarding the external trade adjustment:

[12] We use shares because the source data is for 2013 and the GTAP9.1 database is for 2011.

1. There is no external trade data for services, only for agricultural and manufacturing goods. Therefore, we are implicitly using the original SplitReg –based on state-sector value-added– to assign US external trade in services.

2. Puerto Rico is an independent region in the original GTAP9.1 database, but there are no values for trade between Puerto Rico and the US.[13] Thus, since there is no official trade data between both initial regions,[14] we aggregate Puerto Rico into the Rest of the World residual region in the initial database.

3.6 Using SplitReg with exogenous targets

Once the we have the adjusted external trade US data (from Section 3.5), the aggregated interstate trade data (from Section 3.2), plus the detailed interstate data from CFS (Section 3.2.1), we proceed in the following way:

1. We first use the US-level interstate trade data and the corresponding adjustments in internal production values (VDM) and its components (see Section 3.4, the file 01USint_agg.har), as well as values for the US interstate trade level for each sector (i.e. values in the US to US trade vector). This results in an adjusted, balanced, GTAP9.1 database that has the US as a composite region. We use this as the initial (adjusted) baseline database to split consumption, production, trade, and remaining relevant variables for the US into 51 US State regions, using the sector-state value added as weights and supplemental data to target state-level trade.

2. While weights for state-level value added are the primary drivers for the SplitReg procedure, we also specify target values for state-level trade (interstate trade and international trade by state). To accomplish this, we create a US state-level trade matrix that includes the CFS data for interstate trade in goods (no services), and the adjusted US state-level international trade data from the Department of Commerce to set as targets. SplitReg employs the file that contains all values for the state-level trade matrix (interstate and international trade by state.) As previously described, services data are split according to proportionality assumptions. The SplitReg program provides diagnostic files to ensure that no errors persisted during the splitting procedure for the newly created database.

3. The result is a modified, balanced GTAP9.1 database that includes 172 regions and 31 sectors.

[13] The reason given by the Center for Global Trade Analysis is that there is no official international trade data reported separately for the US territories of Puerto Rico and the US Virgin Islands, and both are included in the official US trade statistics.

[14] The CFS database does not have data for Puerto Rico either, since it does not cover shipments originating from business establishments located in Puerto Rico and other U.S. possessions and territories.

4 Applications using the new US database

To illustrate the contribution of our modified database that splits the US into subsectors, we now compare results from two experiments to show the importance of disaggregation for large countries in certain cases. As previously discussed, the newly created database now allows for shocks in a CGE modeling framework that vary by subregion. Accordingly, we consider two scenarios and compare results that employ the modified database relative to the standard GTAP 9.1 database that has the US as a composite region. First, we assess the impacts of trade cost reductions associated with the potential future use of Arctic transportation routes through the Northern Sea Route (NSR) and the Northwest Route (NWR). Second, we investigate the economic effects of a hypothetical drought scenario that only affects the Southwest region of the US. For certain events or policies that may affect parts of the country differently, we are able to provide richer and more informed results by employing a more highly disaggregated database. Such is the case for changes in trade costs for the Arctic Routes Scenario that affect each coast of the US differently, as well as drought that has persisted in the Southwestern region of the US.

For the scenarios considered in this analysis, we chose to aggregate the individual states in the modified database into 5 representative regions of the US including the Northeast (US_1), Southeast (US_2), South (US_3), Southwest (US_4), and Northwest (US_5). We chose to include 5 port regions to illustrate the changes in trade costs that vary along the US coastline. The state-level mapping to the 5 US regions is described in the appendix. Furthermore, the fully disaggregated database should be used with caution, and in its current form is best use as an input for a higher level of aggregation rather than employing the full 51 state-level database. The reason for this includes primarily the limitations imposed by the US District level data from the Department of Commerce. The availability of state-specific, rather than District level international trade data will allow for further information and greater accuracy when splitting the US into state-level subregions.

4.1 North Western Route experiment

There is a rich literature on the steady reduction of Arctic sea ice (**???**) that may lead to the potential future use of Arctic sea routes as viable shipping lanes for trade. In this experiment, we follow (**??**) and examine the economic effects of commercial use of the North West Route (NWR) that would reduce the shipping distances and time between Northeast Asia (i.e. China, Korea, Japan) and the East Coast of the United States and Canada. Our work augments the literature by comparing the results from (**?**), which implements weighted average trade cost reductions for the US as a composite regions, to results from a modeling framework that employs the newly created GTAP database that represents the US as 5 subregions. To accomplish this, we use the actual distance reductions, if any, that would apply to each US subregions, given the use of the NWR, and estimate the US subregion-specific trade cost reductions, following (**?**). The resulting trade cost reductions are employed in a CGE modeling framework and compared.

Following the literature, we assume that both the NWR and NSR are fully operational year round by the year 2030, and that logistics issues related to navigating the Arctic have been resolved[15] This serves as an "upper bound" scenario that assumes that the NWR becomes a perfect substitute for the Panama Canal, and as such, all commercial shipping between Northeast Asia and the East coast of North America will use the shorter, faster, and cheaper NWR instead of the current route through the Panama Canal route. The focus of this experiment is on the economic effects of Arctic shipping routes on the US. While the NWR is the only Arctic route that will affect shipping distances for US trade, we also must account for trade cost reductions that will simultaneously affect East Asia and Europe. The climate literature confirms that the NWR will only be ice free if the NSR is ice free as well. Therefore, we employ the trade cost reductions from use for the NSR which remain the same for both cases for this scenario.

The economic analysis follows a multi-step process. In the first step, we estimate changes in physical distances between East Asia and the US subregions along major and prospective shipping routes using both the NWR (new routes) and the Panama Canal (current routes). The second step employs a regression-based gravity model of trade to map the distance reductions into estimates of the bilateral trade cost reductions between trading partners at the industry level. While physical shipping costs account for part of the total reduction, there are other trade costs including time and distance barriers that are estimated in the gravity modeling framework. In the third step, we integrate our trade cost reduction estimates into a computable general equilibrium (CGE) model of the global economy to simulate the effect of the commercial opening of both the NSR and NWR on bilateral trade flows, macroeconomic outcomes and the total amount of CO_2 emissions.[16] Where our work differs is that we estimate the trade cost reductions for the US subregions for a unique set of shocks to employ in the CGE modeling framework to compare to the results from **?**. Therefore, each step described is completed twice: once for the US as a composite region and a second time by disaggregating the US into the 5 subregions.

For the first step of our analysis, we estimate the precise distance reductions for bilateral trade flows associated with the NWR for each US subregion, and as a weighted shipping distance for the US as a composite region. Accordingly, we use shipping industry data for the physical distance between ports in the estimation of the distance between two trading partners in the gravity equation **?**.[17] From the data collected in **?**.

[15] The use of 2030 as the benchmark year is mainly for illustration purposes and the use of another year does not affect the main economic results.

[16] In practical terms, we build on the results of **?**, mapping these into transport and trade cost reductions associated along the SSR through the Panama canal and the NWR; yet here, we focus on the impact on trade-related linkages between the US subregions and Asia.

[17] This paper provides a detailed explanation on how the shipping distances are constructed.

We observe that the NWR shipping distance is approximately 25% shorter than the Panama Canal for the routes between Northeast Asia and the Northeast (Newark), while the distance reduction is just about 5% for the route to the South (Louisiana). Finally, when we use the weighted US distance to China, Korea and Japan we observe that the NWR reduces the total shipping distance by around 10% for the US as a composite region. This highlights that the US composite results essentially overestimate the the trade cost reductions for the South, yet underestimate the trade cost reductions for the Northeast when created a trade-weighted average for the US rather than using distance reductions at the subregional level. Therefore, the NWR is is expected to have a more substantial effect on the Northeast relative to the South, and will have negligible effects in the Northwest and Southwest where the shipping distances between East Asia remain the same.

The second step in our analysis is to use the gravity model of trade to estimate the linkage between shorter shipping distances and trade cost reductions. We follow the gravity model estimated in **?**. Controlling for country-specific structural features of the gravity model, estimates of pairwise coefficients provide measures of the impact that distance between two trading partners has in terms of *trade costs* between the two partners. In the present context, when we substitute the current shipping distances using the Panama Canal and the SSR routes with the new Arctic route distances (NSR and NWR), we obtain a measure of how much current trade costs will be reduced by the shorter physical shipping distances associated with the Arctic routes. We complete this using the trade weighted averages for the US as a composite regions and at the subregional level for use with the new, modified database.

Using the data on shipping distance changes, as previously described, in tandem with distance and tariff elasticities in **?**, we can assess how much the decrease in shipping distance translates into effective trade cost reductions. The calculation is as follows:

$$\Delta \text{cost}_{jsd} = \frac{\beta_{j,\text{distance}}}{\beta_{j,\text{tariff}}} \Delta \ln(\text{distance}_{sd}) \tag{1}$$

where Δcost_{jsd} is the change in the total cost of goods sold as a share of the value of trade. They are defined for each sector j and for bilateral trade flowing from region s to region d.[18]

To link the gravity estimations with the CGE model, we allocate these total trade cost reductions from Equation (1) over actual international transport services costs ("atall" in the GTAP code) and the remainder as iceberg trade cost reductions ("ams" in the GTAP code).

[18] Note that these total trade costs are sector-specific and are not symmetric for country pairs. For instance, the trade costs from USA to Korea are slightly different than from Korea to the USA. As expected, since the distance changes associated with the NWR are one-third lower than the NSR changes, the related trade cost reductions are also smaller for the NWR.

We first estimate the shipping services costs reduction as the percentage distance reduction associated with the Arctic routes:

$$\text{atall}_{sd} = -\left(\frac{\text{NSR/NWRdistance}_{sd}}{\text{distance}_{sd}} - 1\right) \tag{2}$$

This reduction is applied directly to the international transport margin (ITM_{sd}), which is the wedge between the *fob* and *cif* trade values in the GTAP database. These are country specific margins: by country of origin (s) to country of destination (d). In this case, we are now able to provide subregional specific margins that correspond to the 5 US subregions.

Finally, iceberg trade costs account for several costs that hinder international trade, such as time, coordination, and other non-shipping service costs (cf. **?**). The iceberg trade costs are calculated as the difference between the total trade costs in Equation (1) and the shipping service cost reductions from Table **??**. Accordingly, the GTAP iceberg trade cost reductions ("ams") are calculated as:

$$\text{ams}_{jsd} = \begin{cases} \Delta\text{cost}_{jsd} - (\text{atall}_{sd} * \text{ITM}_{sd}) & \text{, if positive} \\ 0 & \text{, otherwise} \end{cases} \tag{3}$$

The trade cost estimates are employed in a computable general equilibrium (CGE) model of the global economy in our third and final step. The use of the Arctic routes will affect bilateral trade, sectoral production and consumption patterns, relative international and domestic prices, and the use of sector-specific factors of production, which naturally lends itself to analyses using a CGE modeling framework. Following **?**, we begin by creating a baseline scenario of the global economy in the year 2030 based on macroeconomic projections and comparing how the projected 2030 economy would respond to changes in trade costs. We then assess the changes in bilateral trade flows, relative prices, production and consumption throughout the global economy, given the opening of the Arctic routes. The particular CGE model we use is a modified version of a standard GTAP class CGE model. The macroeconomic projections and model characteristics are detailed in (**?**) and (**?**).[19]

We present the CGE results for each case (US composite and US subregions) as the differences between the baseline values in 2030 (i.e. the baseline scenario with no Arctic shipping lanes) compared to the counterfactual scenario where bilateral trade uses the Arctic routes, where applicable. We incorporate the estimates for both the transportation and trade cost reductions into the CGE model to assess the effects of the Arctic shipping lanes on bilateral trade flows, sectoral output, and other macroeconomic variables.[20]

[19]The main distinction between our model and the standard GTAP model is that we use a monopolistic competition framework with increasing returns to scale.

[20]As explained in Section **??**, we accomplish this by implementing technical efficiency gains in shipping and iceberg trade costs that are equivalent to the estimated reductions in total trade costs.

Table 2: North Western Route simulations, United States trade using model with one or five US sub-regions, percentage changes with respect to the baseline in 2030

	EU		**East Asia**		**RoW**	
	exports	imports	exports	imports	exports	imports
US: 1 region						
USA	-1.1	-0.3	9.8	3.4	-0.3	0.5
US: 5 regions						
USA_1	-1.7	-0.8	34.8	7.1	-1.4	1.4
USA_2	-1.2	-0.5	8.6	3.4	0.2	-0.4
USA_3	-1.2	-0.2	5.6	1.8	0.1	0.2
USA_4	-0.4	0.0	0.5	-0.9	0.2	0.4
USA_5	-0.9	-0.4	0.2	-1.2	0.3	0.3
USA total	-1.2	-0.4	8.6	2.3	-0.2	0.6

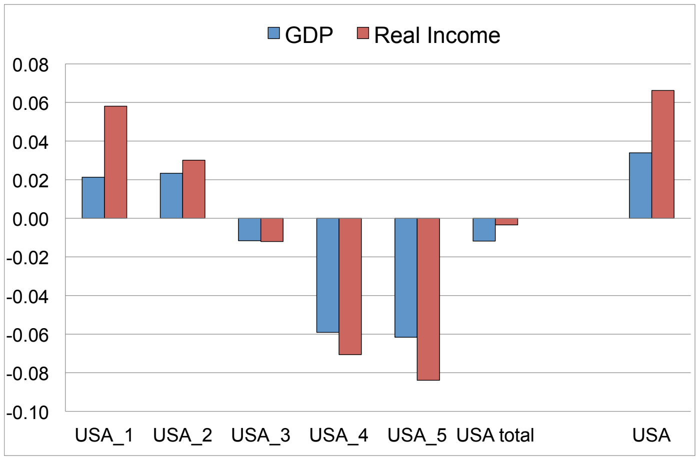
Notes: EU is the 28 European Union member states, East Asia is China/Hong Kong, Japan, Korea and Taiwan, RoW is the rest of World (i.e. the remaining regions. Source: Own estimations using CGE model and GTAP9.1 database.

Table 3: North Western Route simulations, United States and the EU trade with East Asia, percentage changes with respect to the baseline in 2030

	China		**Japan**		**Korea**	
	exports	imports	exports	imports	exports	imports
US: 1 region						
EU	7.0	10.9	11.6	15.1	8.9	11.2
USA	3.6	13.3	3.1	5.2	2.9	5.4
US: 5 regions						
EU	7.0	11.1	11.6	15.4	8.8	11.5
USA_1	7.8	51.6	5.8	9.2	4.6	11.3
USA_2	3.7	11.4	2.7	5.4	3.2	4.4
USA_3	1.9	7.1	2.3	3.8	1.7	3.9
USA_4	-1.0	0.7	-1.0	-0.1	-0.5	0.2
USA_5	-1.2	0.3	-1.3	-0.3	-1.0	0.1
USA total	2.5	12.5	2.2	3.2	1.8	3.6

Source: Own estimations using CGE model and GTAP9.1 database.

Figure 1: North Western Sea route, United States, GDP and real income changes with one USA region and five USA sub-regions, percentage changes with respect to baseline in 2030



4.2 Drought in the US Southwest experiment

To further illustrate the importance of country disaggregation in certain cases, we present a simple, hypothetical drought scenario. Here, we assume that drought affects only the Southwest subregion of the US (US_4). To simulate drought, we employ negative changes to agricultural productivity in the southwest by implementing a -20% shock to "aoall." This causes an inward shift in the supply of agricultural production for the Southwest in the model. To assess this same scenario, but for the US modeled as a composite region, we calculate the weighted average of the agricultural productivity shock, which is equal to -4.14%, and apply this shock to agricultural productivity in the US composite region.

The differences between the results for modeling the US as a composite region versus subregions are described in Tables 4 and 2. Noteworthy differences arise for agricultural production and real returns to land. We clearly observe the losses in the Southwest US region (USA_4) and that the one-region losses are higher than the aggregated losses with the five-region version.

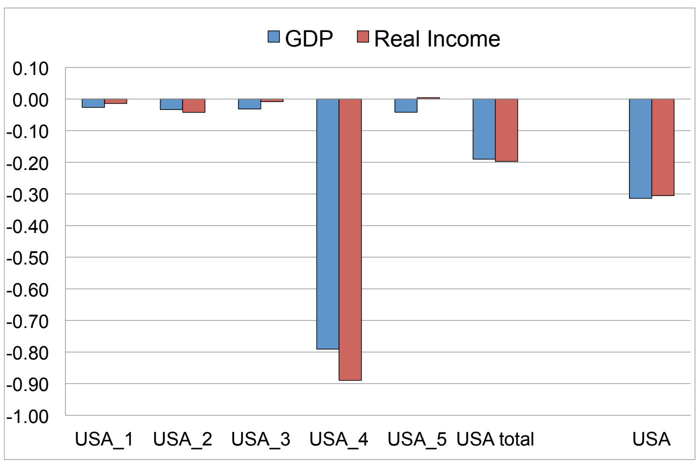
Regarding GDP and real income, differences persist between results for the one-region US and the US five-region case. We clearly identify the losses in the particular US sub-region affected (Southwest, USA_4). Modeling the US as a composite region may overestimate the total losses experienced for this particular case where the change in productivity is isolated in one subregion of the country rather than experienced uniformly across the entire US.

Table 4: Drought in the US Southwest simulations, United States production and return to factors using model with one or five US sub-regions, percentage changes with respect to the baseline in 2030

	US: 1 region			US: 5 region			
	USA	USA_1	USA_2	USA_3	USA_4	USA_5	USA total
Production:							
Agricultural products	-4.7	0.4	0.3	0.3	-23.7	0.2	-3.6
Processed foods	-1.4	-0.4	-0.4	-0.7	-0.7	-1.0	-0.6
Real returns to:							
Land	-4.1	3.8	2.9	2.6	-33.0	2.5	-3.8
Low-skill labour	-0.2	-0.1	-0.1	-0.1	-0.2	-0.1	-0.1
High-skill labour	-0.2	-0.1	-0.1	-0.1	-0.4	-0.1	-0.1
Capital	0.0	0.0	0.0	0.0	0.1	0.0	0.0
Natural resources	-0.6	-0.2	-0.3	-0.3	0.0	-0.3	-0.2

Source: Own estimations using CGE model and GTAP9.1 database.

Figure 2: Drought in the South West United States, GDP and real income changes with one USA region and five USA sub-regions, percentage changes with respect to baseline in 2030



References

Guci, L. and C. I. Mead (2014). “Domestic Trade in Services in Regional Input-Output Models,” BEA Working Papers 0106, Bureau of Economic Analysis.

Horridge, M. (2011a). “GTAPAdjust: A Program to Balance or Adjust a GTAP Database,” Item TPMH0104 at www.monash.edu.au/policy/archivep.htm, Centre of Policy Studies, Monash University, Australia.

Horridge, M. (2011b). “SplitReg: A Program to Create a New Region in a GTAP Database,” Item TPMH0105 at www.monash.edu.au/policy/archivep.htm, Centre of Policy Studies, Monash University, Australia.