1-13-2023

A Forestry Trade Data Discrepancy Analysis

Morgan Alyce Daigle
Louisiana State University and Agricultural and Mechanical College

A FORESTRY TRADE DATA DISCREPANCY ANALYSIS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Agricultural Economics and Agribusiness

by
Morgan Alyce Daigle
B.S., University of Lafayette Louisiana, 2019
May 2023

Acknowledgments

I would like to express my deepest appreciation to my major professor, Dr. Jinggang Guo. I could not have undertaken this journey without his optimistic kindness, motivation, and support. This thesis would not have been possible if not for the advice, instruction, and patience that Dr. Hector Zapata offered to me during both the completion of my thesis, and my graduate studies. I would also like to extend my sincere thanks to my committee member, Dr. Maria Bampasidou for being a strong and calming presence which empowered me to be confident in my endeavors.

I would also like to extend my gratitude to my current and former graduate classmates: Whitney McKinzie, Junior Betanco, Nicolas Alvarez, Adriana Alfaro, and Stephen Adebayo. Our discussions and comradery have encouraged me more than words can describe.

I would be remiss in not mentioning the administrative program specialist and administrative coordinator for the Department of Agricultural Economics and Agribusiness at LSU: Marla Jones and Jody Bisset for their thoughtfulness, and rigorous attention to detail. Without these ladies, I would not have had the resources to reach my goals.

Finally, I would like to extend a special thanks to my family. I am eternally grateful to my parents (Kim and Jonathan Daigle), my sisters, brother-in-law, and nephew (Devyn Daigle, Kameron, Seth, and Rhett Thibodeaux), along with my grandparents (Lela Dickinson, and the late John and Helen Daigle), and my fiancée (Everett Franks) for their perpetual encouragement, love, and understanding. Without their support, I could not have made it to where I am today.

Table of Contents

Abstract

Trade data across multiple databases experience unavailability across some countries/forestry products, inconsistency, and unreliability; and these qualities manifest as discrepancies in the data. Literature provides evidence of discrepancies and inconsistencies within international trade statistics, including documented cases in which they are present within agriculture sector trade.

While researchers have worked to pinpoint factors to explain discrepancies, studies on the forestry trade databases are not as prevalent. Therefore, more research needs to be conducted to identify discrepancies within forest sector products trade data to understand the nature of discrepancies found between different bilateral trading partners.

The goal of this thesis is to identify and analyze discrepancies in forestry trade data found in bilateral trade series sets for the United States and its top trade partners of forestry products. Discrepancies will be identified using simple mathematical formulations and compared across trade flows and forestry products. Then a unique time-series trade data discrepancy analysis approach is conducted to estimate the nature of trade data discrepancies. The thesis aims to provide a framework for proceeding researchers to utilize in order to apply time-series analysis techniques to trade data discrepancies across any product and country. It also aspires to fill in gaps in the literature of trade data discrepancy analyses that examine forestry product trade data.

Results indicate that discrepancies are present between the bilateral import and export quantity statistics for industrial roundwood, sawnwood, plywood and wood chips and particles for the trade flows from Canada to the U.S, Brazil to the U.S, China to the U.S, the U.S. to China,

the U.S. to Japan, and the U.S. to the U.K. Augmented Dickey Fuller (ADF) and cointegration tests reveal that each unique time series bilateral trade quantity pair exhibits unique data generating processes. Therefore, estimating the discrepant relationship between bilateral import and export quantities of forestry products does not follow one clear cut method; but most discrepancies can be estimated through simple or multiple linear regressions, vector error correction models (VECM), or auto-regressive distributed lag error correction models (ARDL ECM).

Chapter 1. Introduction

1.1.Motivation

In 2020, trade accounted for 51.57% of the world's total GDP (World Bank, 2020). The data collected from world trade across all traded product sectors is in demand by researchers and policymakers in most countries worldwide. They use the data to perform empirical analyses, utilize product sector models, and develop food security indicators, among other applications. Consumers of trade data rely on its availability, reliability, timeliness, and consistency. If one of these qualities is missing from the data, it can cause a ripple effect of inaccuracies in the results of analyses, the output of models, the development of indicators, and potentially lead to erroneous policy implementation.

In 2019, forestry products accounted for 2.9% and $164 billion of all traded commodities, and the United States (U.S.) was the top exporter and the third-largest importer of forestry products by weight in metric tons (MT) (Chatham House, 2018). "Forests absorb greenhouse gases that affect climate change, provide clean water, protect watersheds, reduce erosion, provide food and medicine, serve as a defense during natural disasters, and provide habitat to more than half of the world's land-based species" (WWF, 2019). Demand for land, food, energy, timber products, and growth increases society's dependence on forests (WWF, 2020). In 2019, the U.S. alone experienced a high (5%-25%) increase in dependence on agriculture and natural resources (Development, 2021)[1]. The ability to accurately determine forestry product prices, supply, and demand quantities, and forecast developments in the forestry product markets is crucial for the future well-being of our society.

[1] Found in Index 2: Commodity Export Dependence, pg. 17

Researchers rely on mathematical models called Forest Sector Models (FSMs) to accomplish this goal (Riviere and Caurlo, 2020; Latta et al., 2013). FSMs can also be used to assess policies that deal with climate change, renewable energy production, and environmental protection (Latta et al., 2013). FSMs are valuable tools that can be used to understand society's demand and dependence on the forest sector. As most sector models, FSMs rely on trade data to represent the industry's natural, technological, and economic factors along with their interactions (Riviere and Caurlo, 2020; Latta et al., 2013). If forest sector trade data lacks availability, reliability, timeliness, or consistency, any output by the models will reflect those deficiencies, leading to inaccurate representations of the natural world and misperceptions of the role that the forest sector plays in society.

One commonly utilized source of forest sector trade data is the Forestry database compiled by the Food and Agriculture Organization of the United Nations Statistics (FAOSTAT) (Buongiorno, 2018). FAOSTAT Forestry Data consists of an annual time series of data on production, imports, and exports of Forest Sector products in each country starting in 1960 and on the trade flows of partner countries beginning in 1997 (FAO, 2020). This database is downloaded about every two minutes (Steele et al., 2021).

Trade data databases are commonly referred to as mirrored statistics. Import data from a destination country should reflect export data from the origin country. Trade data across multiple databases, including the FAOSTAT Forestry database, experience unavailability across some countries/products, inconsistency, and unreliability; these qualities manifest as discrepancies in the data. Discrepancies are identified in mirrored statistics when the reflections differ from one another. Discrepancies are commonly documented in trade data statistics,

including inconsistencies within FAOSTAT databases (Cafiero et al., 2014; Skjerstad et al., 2021; Kallio and Solberg, 2018). Discrepancies can be due to unavoidable errors, lack of resources needed for reporting, reporting errors, illegal activities, and other errors in the chain of data collection (Hamanaka, 2012). The consequences of using discrepant data include inconsistent research results, forecasts, and policy implementation, all of which should not be ignored.

While this thesis is focused on the FAOSTAT's Forestry Database, it is important to note that FAOSTAT databases are also used in other agricultural applications such as facilitating rural development, poverty reduction, assessing of food (in)security, and elasticities of demand of the forest and other agricultural industries, (Carfagna et al., 2013; Cafiero et al., 2014; Skejerstad et al., 2020). These databases are also subject to discrepancies. For example, when Jayasinghe et al. (2010) aimed to find which factors determine the seed trade and the relative importance of those factors, they noted that inconsistencies exist in FAOSTAT production statistics of corn production (Jayasinghe et al., 2010). In general, caution is advised when utilizing FAOSTAT databases.

Literature provides evidence of discrepancies and inconsistencies within international trade statistics, including documented cases in which they are present within agriculture sector trade. Researchers have also worked to pinpoint factors to explain discrepancies. However, specific studies on the FAOSTAT Forestry Database are not as prevalent, and to my knowledge there have been no previous trade data discrepancy analyses that apply time series analysis techniques. Therefore, more research needs to be conducted in search of the most accurate way to identify discrepancies within forest sector products trade data, understand the relationships

between discrepancies found in different trading partners, and, more precisely, identify their underlying nature.

1.2. Research Questions and Objectives

The goal of this thesis is to explore whether discrepancies exist within forest sector trade data, as it pertains to the U.S. and its top trading partners of forest products; and if they exist, which factors influence their presence. The following four objectives will be used to achieve this goal. First, three simple mathematical formulations will be applied to identify discrepancies in forestry trade data for the United States and its top forestry trading partners over the top traded forest sector products. The three mathematical formulations will be defined as Excess Trade (EX1), Excess Net Trade (EX2), used in respect to unilateral statistics, and GAP which is used in respect to bilateral trade flow statistics. The results from discrepancy identifying formulas will be compared across different countries, different forestry products, and trading partner pairs. Second, further investigation of bilateral trade flow is conducted by using an analysis of correlation to identify potentially discrepantly reported trade statistics . Third, a times series econometric analysis on the bilateral forestry trade data quantities will be employed to determine the estimated nature of discrepancies. Lastly, a discussion will ensue to connect results from the discrepancy analysis as they relate to factors that potentially cause discrepancies including corruption and tariffs.

1.3.Research Methods

This thesis first aims to identify discrepancies in the FAOSTAT forestry database while focusing on the United States and its trading partners of Forestry Sector products. First, each country's unilateral forestry trade data quantity statistics will be examined using the simple

mathematical formulations of Excess Trade and Excess Net Trade. If discrepancies exist within the reported unilateral statistics, then they will be compared across each country. The bilateral trade quantity statistics will be analyzed by applying a third mathematical formulation, the GAP equation, to identify discrepancies between pairs consisting of the United States and each of the top trading partners of its top traded forestry sector products. To further explore partner country pairs bilateral data, correlation between reported exports from an origin country and reported imports from a destination country from a destination will be determined. Then a time series analysis on bilateral forestry product trade quantities is conducted including Augmented Dickey Fuller (ADF) tests, Johansen Cointegration tests, and Autoregressive Distributed Lag (ARDL) Cointegrating Bounds procedure to determine the most appropriate way to estimate discrepancies and their potential causes. Finally, Granger non-causality tests will be utilized to make inferences on the bilateral import and export statistics for bilateral series sets in which discrepancy estimations do not seem feasible by the aforementioned methods.

The United States' trading partners will be examined in two ways; the top three trading partners when the U.S. is the importing country and then the top three when the U.S. is the exporting country. The top trading partners are in terms of the weight of goods traded rather than the value because this proposed thesis will focus on the quantity of forestry products traded rather than the value.[2] Six countries will be examined: the U.S., China, the U.K., Japan, Canada, and Brazil. Where the top three countries that import forestry products from the U.S. are China, the U.K., and Japan; the top three countries that export forestry products to the U.S. are Canada,

[2] Quantity of traded forestry products is used rather than value in order to avoid Cost Insurance and Freight (CIF) and Free on Board (FOB) reporting value differences.

Brazil, and China (Chatham House, 2018). The flow of trading partner country pairs will also be examined for a total of 6 pairs: (1) U.S. to China, (2) U.S. to the U.K., (3) the U.S. to Japan, (4) the U.S. from Canada, (5) the U.S. from Brazil, and (6) the U.S. from China. Figure 1.1 illustrates the trading partner flows to be explored by this research.

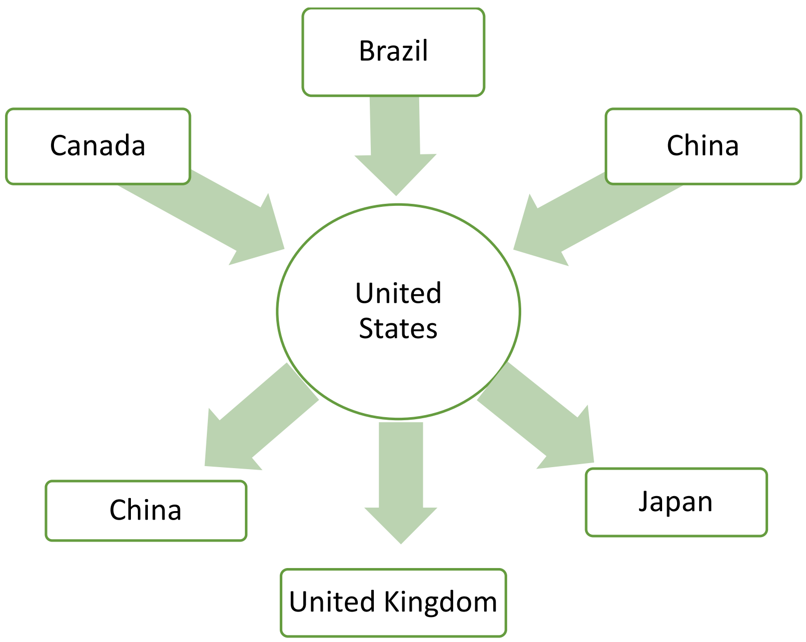

Figure 1.1. Flow of Trading Partner Pairs.

The top three forestry commodities that the United States imports by weight include (1) lumber and sawn wood, (2) board and plywood, and (3) wood pulp, chips, and waste products (Chatham House, 2018). The top three forestry commodities that the United States exports by weight are (1) wood pulp, chips, and waste products, (2) lumber and sawnwood, and (3) board and plywood (Chatham House, 2018). The commodities can all be considered forestry product outputs, of which their primary input is roundwood. The listed commodities can also be regarded as inputs to varying degrees. A simplified flow chart of the inputs and outputs of the forestry

products is illustrated in Figure 1.2, of which the products to be examined are industrial roundwood, sawnwood, wood chips, particles and residues, wood pulp, and plywood.

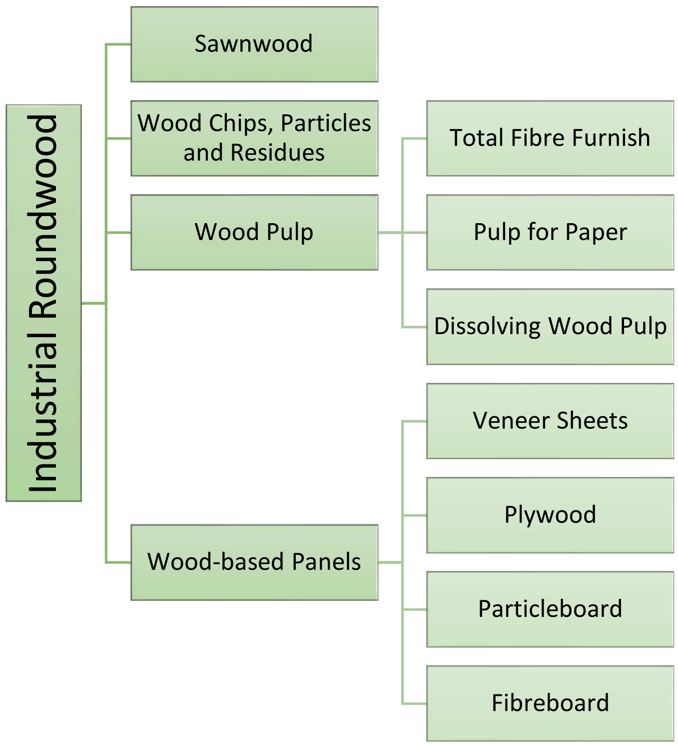

Figure 1.*2.* Forestry Product Inputs and Outputs Flow.

1.4. Key Contributions

There is extensive evidence of discrepancies in trade statistics, but more research needs to be done on the forest sector, and to my knowledge this is the first paper that applies time series analysis techniques to a trade data discrepancy analysis in the U.S. The key contributions of this thesis are to provide more evidence of the presence, or lack thereof, of discrepancies in forestry trade databases. If discrepancies are identified, this work will also contribute to the body of research that explains potential drivers of forestry sector product trade discrepancies between

trading partner country pairs. In addition, the contributions will help further the knowledge of where and why discrepancies are present in forestry trade data. Ultimately, developing a framework for subsequent researchers to utilize when examining trade data of any product sector over any trading partner country pairs.

1.5. Outline

The remainder of this thesis can be outlined as follows. Chapter 2 will review literature including potential drivers of trade discrepancies, methods used to detect and analyze discrepancies, and possible approaches to mend discrepancies. Chapter 3 will be composed of the methodological framework on which this proposal is based. Chapter 4 consists of a description of the relevant data that will be used. Chapter 5 will consist of the results of the unilateral and bilateral trade data discrepancy analysis along with a discussion. Lastly, Chapter 6 encompasses concluding remarks.

Chapter 2. Review of Literature

Mirror statistics are commonly used for trade data discrepancy analyses. Mirror statistics refer to partner trading countries' import, export, and production statistics. More specifically, the import statistics of a destination country are compared to the export statistics from the origin country. They can be examined as aggregates of all trade or disaggregated into different industries and products. Discrepancies occur when the reflections of mirrored trade partner countries do not align. Most discrepancies are identified as underreporting or overreporting, which can be manifested by positive or negative trade discrepancies (Ferrantino et al., 2012). Some differences in the mirrored trade statistics are unavoidable, while others might indicate illegal activities, shortfalls of a country's customs reporting practices, or the data collection process itself.

There is extensive literature analyzing the discrepancies among international trade data as well as some specifically targeting forestry products. Some authors identify the discrepancies using several variations of a simple formulation, while others identify the discrepancies using optimization linear programming practices. Further, some authors perform analyses to find the determinants of discrepancies, or trade in general, using methods that include regressions and tests for correlation. They work to learn information about the causes of discrepancies and relationships between trading partners. Authors find that discrepancies could be due to misclassification, misattributions induced by transshipping and re-exporting, or the presence of tariffs; however, not all of the studies were conclusive.

This literature review is organized as follows. First, the potential drivers of trade data discrepancies will be reviewed. After which, methods used by preceding researchers to identify

and analyze discrepancies will be discussed. Lastly, literature involving how to fix discrepancies in trade data is overviewed.

2.1. Potential Drivers of Trade Data Discrepancies

Many researchers accept that an unavoidable discrepancy between the values of mirrored trade data will always be present due to transportation and shipping cost differences. Three standard international guidelines for trade valuation include Free Alongside Ship (FAS), Free on Board (FOB), and Cost, Insurance, and Freight (CIF); most imports are registered using CIF while exports use either FAS or FOB (Ali et al., 2019; Hamanaka, 2012; Yurik et al., 2020; Sen, 2000; Ferrantino and Wang, 2008; Markowicz and Baran, 2020)[3]. Researchers generally accept that a six to ten percent difference between import and export values can explain the discrepancy due to transportation cost (Ferrantino et al., 2012; Hamanaka, 2012; Yurik et al., 2020; Sen, 2000).

Another unavoidable discrepancy within trade data is due to time and distance between partners; when a country exports product towards the end of the year, it might not arrive at the destination until the following year (Ferrier, 2021; Vincent, 2004; Kazunobu, 2020; Ali et al., 2019; Hamanaka, 2008; Sen, 2000). This time and distance lag is especially prevalent when using monthly data but usually dissipates when annual data is used. The lag becomes more extreme the farther the trading partners are from each other.

[3] Free On Board (FOB) includes registering the cost of the traded product(s) as well as the cost of loading the product(s) onto its transportation mode. Free Alongside Ship (FAS) includes FOB costs excluding the cost of loading product(s) onto its transportation mode. Cost Insurance and Freight (CIF) includes FOB costs as well as the costs of insurance and freight.

It is often the case that discrepancies cannot only be explained by differences in transportation cost or time and distance lags. They can be caused unintentionally through human recording and measurement errors, or intentionally through illegal activities that aim to avoid trade controls, tariffs, or taxes. The remainder of Section 2.1 identifies events that might cause trade data discrepancies due to reporting errors, measurement errors, or illegal activities.

One such discrepancy might arise through transshipping or re-exporting, which can be caused intentionally or not. This practice can be described as moving traded products through an intermediary country before reaching the destination (Ferrier, 2021). It is widely accepted that importers will know more about the country of origin than exporters' knowledge of the final destination; however, this is not always the case (Ali et al., 2019; Ferrantino and Wang, 2008). If the destination country has more information, reported imports to the destination country will be larger than reported exports from the origin country. If the origin country has more information, reported exports from the origin country will be larger than the reported imports to the destination country. Several possibilities could arise from transshipment: the origin country reports the intermediary country as the final destination, the actual destination country reports the intermediary as the origin, both the intermediary and origin countries report exports to the destination, goods are stored in the intermediary country for a while and re-exported to a different country, or the intermediary adds a mark-up to the shipment value before re-exporting to the final destination (Ali et al., 2019). Transshipping is a legal practice commonly used; for example, Hong Kong is widely used as an intermediary (Ferrantino et al., 2012). However, it is illegal if it is intentionally used to hide the country of origin (Ferrier, 2021). Re-exporting can also be unlawful when used to legally re-export illegally acquired goods or attempt to bypass

intellectual property rights, as is the case with Czech beer re-exports from Russia (Yurik et al., 2020). The word transshipment should be interpreted carefully within the literature. Although it is primarily used as a synonym for re-exports, it is sometimes used to describe goods in transit (Ferrantino and Wang, 2008).

Goods in transit are used to describe shipments of traded goods that pass through intermediary countries without ever being unloaded except for switching the means of transportation (Ferrantino and Wang, 2008; Ali et al., 2019). It is possible that true re-exports are accidentally or intentionally confused with goods in transit (Ferrantino and Wang, 2008).

The two central systems of reporting international trade data are the General Trade System and the Special Trade System. The General Trade System is meant to record all goods physically crossing the country's national borders, excluding goods in transit (Ferrantino and Wang, 2008; Sen, 2000). The Special Trade System captures exports produced or transformed within a country and consumed imports (Ferrantino and Wang, 2008; Sen, 2000). For example, re-exports that were not transformed within the intermediary country would be included in the general system, but not the special (Ferrantino and Wang, 2008). Countries can use either system, which means that not every country reports the same way; so, discrepancies can occur when trading partners do not record traded goods using the same system.

Mis-invoicing takes place when the volume, quantity, or quality of a trade shipment is manipulated when reporting to customs, resulting in over-invoicing (under-invoicing) if the value is invalidly increased (decreased) (Pardo-Herrera, 2021). This practice masks the actual value of the traded goods, which might be used for tax and tariff evasion schemes or to evade capital controls (Ali et al., 2019; Ferrantino et al., 2012). Transfer pricing is a form of mis-invoicing

committed by related parties in different countries, for example, different branches of a multinational corporation. It involves under- or over-invoicing so that firms can shift their profits from high-tax countries to low-tax countries to avoid tariffs, especially when the firms are shipping to themselves (Ferrantino and Wang, 2008; Ferrantino et al., 2012; Hamanaka, 2012). Smuggling can be considered the most extreme case of under-invoicing, in which no value for the transaction is recorded (Ferrantino and Wang, 2008; Ferrantino et al., 2012; Vincent, 2004; Ali et al., 2019). Smuggling can occur when the traded good is illegal in one country but not the other, or illicit in both countries. Smuggling can lead to a trade discrepancy in the former case and no visible discrepancy in the latter (Ferrantino and Wang, 2008).

Misattribution deals with the origin or destination of a good. This occurs when traders falsify either the actual origin or destination to customs (Ali et al., 2019; Ferrantino and Wang, 2008). Traders might engage in misattribution to take advantage of special programs or in attempts to evade taxes and tariffs (Ali et al., 2019; Ferrantino and Wang, 2008). Misattribution will lead to trade data discrepancies and might be coupled with mis-invoicing. Misattribution is sometimes referred to as direction misclassification (Hamanaka, 2012).

Misclassification is also a cause for trade discrepancies. It can be done intentionally to abet in illegal activities or might be due to different classification practices performed by each country. Misclassification by customs offices might occur when one good is recorded under different commodity codes by either the origin or destination country due to somewhat similar commodity descriptions and names (Hamanaka, 2012). If one trading partner reports import/export data of a product while the other partner does not, it is reasonable to believe the product has been misclassified by one of the partners (Hamanaka, 2012). Misclassification will

result in trade data discrepancies even if the value and quantity of the product are reported correctly; this is most prevalent among emerging technologies that have not been assigned classification yet or for traded products with vague classification standards (Ferrantino et al., 2012). Misclassification might also be referred to as mismanifesting (Ferrier, 2021). Mismanifesting can occur when the good is mislabeled to be disguised as something else, as is the case with the illegal honey trade (Ferrier, 2021).

Discrepancies might also arise from differing conversion factors used within trading partner countries (Kallio and Solberg, 2018). Conversion factors vary across countries, leading to mistakes in import/export trade data. Differing exchange rates between customs offices and exchange rate fluctuations can also account for discrepancies between trading partner countries (Ferrier, 2021; Pardo-Herrera 2021; Vincent, 2004; Ferrantino et al., 2012; Hamanaka, 2012; Sen, 2000). It is often the case that customs officials are responsible for converting the value of imports from the origin country's currency to the destination country’s money at the time of payment (Sen, 2000). Some countries are under a floating exchange rate regime of which customs officials may not be informed of the current exchange rate daily; therefore, if high fluctuations of exchange rates are prevalent, it will lead to discrepancies among the trade data (Sen, 2000).

2.2. Detecting and Analyzing Discrepancies

One way to detect discrepancies is to use optimization methods. Kallio and Solberg (2018) use linear programming targeting the forestry industry to assess if the volume of wood biomass, according to FAOSTAT forest database statistics, was sufficient to supply the raw material needed for reported production quantities. They found inconsistencies in the data for several countries.

Buongiorno (2018) also aimed to reveal potential discrepancies in the FAOSTAT forest statistic database through optimization with goal programming. They estimate the consumption of forestry product inputs based on the reported production of various outputs, conditional on the amount of input needed per unit of output. This author compares estimated and reported consumption to make inferences on the quality of data, and also found discrepancies in the data for several countries.

Another method used to identify discrepancies is through simple mathematical equations. Ferrier (2021) uses three different methods to capture the discrepancy among trade data of honey using mirror statistics. The author's excess trade methods define excess exports as total exports divided by production; and excess net exports as exports minus imports, divided by production. The author indicates that solutions of either equation that are greater or equal to one indicate discrepancies, which could be explained by mismanifesting imported honey and re-exportation to the United States. Ferrier (2021) also establishes a GAP equation as the percentage difference between quantities of exports reported in an origin country and those imports reported by the destination country. If GAP is positive, a larger quantity is recorded by the origin country than the destination country.

Pardo-Herrera (2021) uses forest industry COMTRADE data and simple equations to conduct a trade discrepancy analysis. They find instances of under- and over-valuation in the Peruvian timber trade and infer possible causes of the discrepancies.

Torres-Rojo (2021) identifies discrepancies in FAOSTAT and SEMARNAT data by defining their equation, Apparent Demand (A.D.) to equal production plus imports minus exports. The annual growth rate for the A.D. for wood products is used to obtain the expected A.D.; the illegal

logging is defined as expected A.D. minus observed A.D. They also use a second approach to estimate illegal timber by comparing wood inputs and outputs through a wood balance analysis.

Vincent (2004) uses European Forest Institute (EFI) data to develop a relative measure of discrepancies in the trade statistics defined as imports minus exports, divided by imports plus exports. The measurement ranges from -1 to +1, so positive values indicate that the importer reports more trade, negative values indicate that the exporter reports more trade, and zero means no discrepancy is present.

Kazunobu's (2020) measure for discrepancies is formulated by taking the natural log of exports minus the natural log of imports. In contrast, Ferrantino et al. (2012) define their measurement as the natural log of imports minus the natural log of exports.

Ferrantino and Wang (2008) developed three measurements for the mirrored statistic discrepancies. The first measures differences at the commodity level as imports minus exports, divided by imports, and multiplied by 100. Their second measure is extended to account for the distance between the importing and exporting countries. The third measure is extended further to develop an absolute average aggregation index by partners or commodities.

Ali et al. (2019) uses a ratio between reported exports from the origin country and imports reported by the destination country to develop a unit-free measurement. Hamanaka (2012) also uses a ratio to identify discrepancies, but theirs is defined as import-side data divided by export-side data. They determine that an absolute value greater than 1.10 percent indicates a discrepancy. The 1.10 percent is representative of the conceptual CIF-FOB cost of transportation. By using a multiple mirror technique, they look at all trading partners, including intermediaries, to develop a complete picture of where discrepancies might originate. Sen (2000) also utilizes

the 10 percent rule for CIF-FOB costs when analyzing the International Monetary Fund's Direction of Trade Statistics (DOTS). They define Trade Difference (T.D.) as imports minus exports, divided by imports, and multiplied by 100. When the T.D. statistic is positive; the destination country reports more imports than the origin country exports.

Markowicz and Baran (2020) aim to improve the methods commonly used to identify discrepancies in international trade data mirrored statistics by proposing a new value-weighted aggregated index for quantity as an indicator of trade data quality. They compare their index to two others: an aggregated quantity-based index and an aggregated value-based index. All three of the indices take on a value from zero to two, where zero indicates the best quality; they find that quantity-based methods are more accurate.

After identifying discrepancies in trade data, several studies have worked to analyze their underlying causes through a variety of methods. Ferrier (2021) explores whether discrepancies in honey trade data can be explained by the presence of antidumping tariffs and corruption levels by running regressions of their discrepancy measure, GAP, over a dummy variable for antidumping tariffs among trading partners, corruption level indicators, and fixed effect dummy variables for year and country. They also explore whether the discrepancies can be explained by transshipping by running a regression over the logged imports to a destination from an intermediary country over the logged exports from an origin country to the intermediary. Likewise, Ferrantino et al. (2012) explore whether discrepancies in U.S./China trade data can be explained by value added tariffs (VATs), level of corruption indicators, and taxes through regression analysis; they run Ordinary Least Squares (OLS) and panel fixed effect regressions. Unlike the previously mentioned article, Ferrantino et al. (2012) report that they are unable to

reject the null of no serial correlation in the error term, so they run a first difference regression as an alternative to fixed effects. Similarly, Vincent (2004) examined whether discrepancies in forest product trade statistics can be explained by tariffs, and corruption indicators[4] as well as the distance between trading partners, and the prices of imports and exports. They too ran an OLS regression and found issues of heteroskedasticity in the error term, so they estimated a generalized least squares (GLS) model with fixed effects. Kee and Nicita (2016) explore the effect of non-tariff measures (NTM) including Sanitary and Phytosanitary Measures[5] and Technical Barriers to Trade[6] and inferred that discrepancies are increased by the presence of NTMs. The articles mentioned above found evidence that discrepancies could be explained by tariffs, corruption levels, transshipping, distance between trading partners, prices of imported and exported forestry products, and NTMs. Figures 2.1, 2.2, 2.3 and 2.4 summarize the tariffs, corruption indicators, unit prices of bilaterally traded forestry products, and NTMs for the six trade flows explored by this thesis respectively.[7]

Another approach to analyzing discrepancies takes a look at the effect of time and distance shipping lags on partner country trade discrepancies, Kazunobu (2020) examines Japan's export data and its partner countries' import data. They hypothesize that exports shipped at the end of a calendar year will not arrive or appear in reported imports until the following year. They use monthly data and a probit model with importer, product, and year fixed effects, to estimate

[4] Corruption indicators include the World Governance Indicators (WGI): Control of Corruption, Government Effectiveness and Rule of Law.
[5] Sanitary and Phytosanitary Measures includes labeling, hygiene, maximum pesticide residue limits, and testing requirements.
[6] Technical Barriers to Trade includes labeling, product quality, packaging, and certification requirements.
[7] Figures 2.1, 2.2, 2.3, and 2.4 are referenced in the discussion relative to the results from this paper's analyses.

the probability that a discrepancy occurs when Japan exports its traded products at the end of the year.

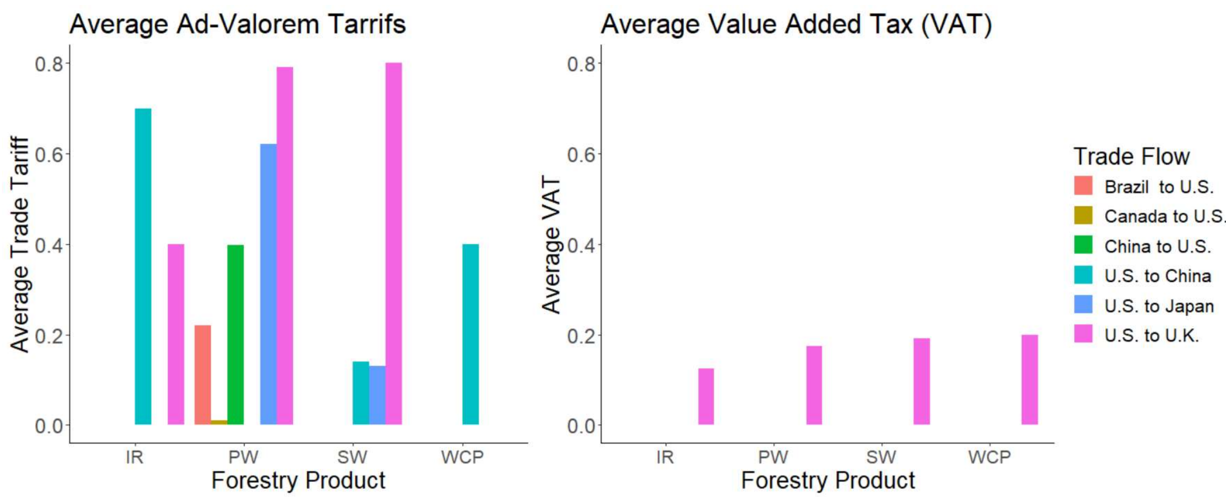

Figure 2.1. Average Tariffs and Duties for Forestry Products.
Source: WTO, Tariff Download Facility; GOV.U.K, UK Integrated Online Tariff.
Note: Higher tariff rates are expected to increase discrepancies.

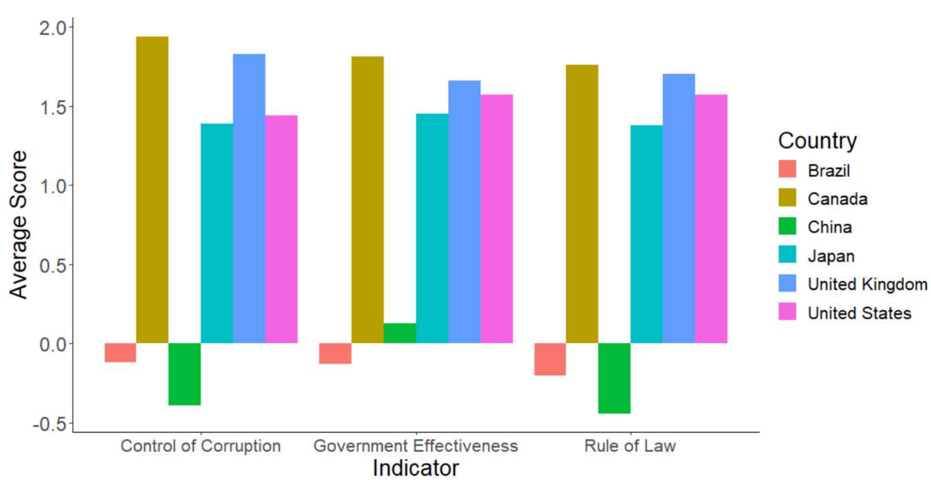

Figure 2.2. Average World Governance Indicators.
Source: Worldwide Governance Indicators, 2018.
Note: Scores range from -2.5 (weak) to 2.5 (strong). Weak scores are expected to lead to more discrepancies compared to strong scores.

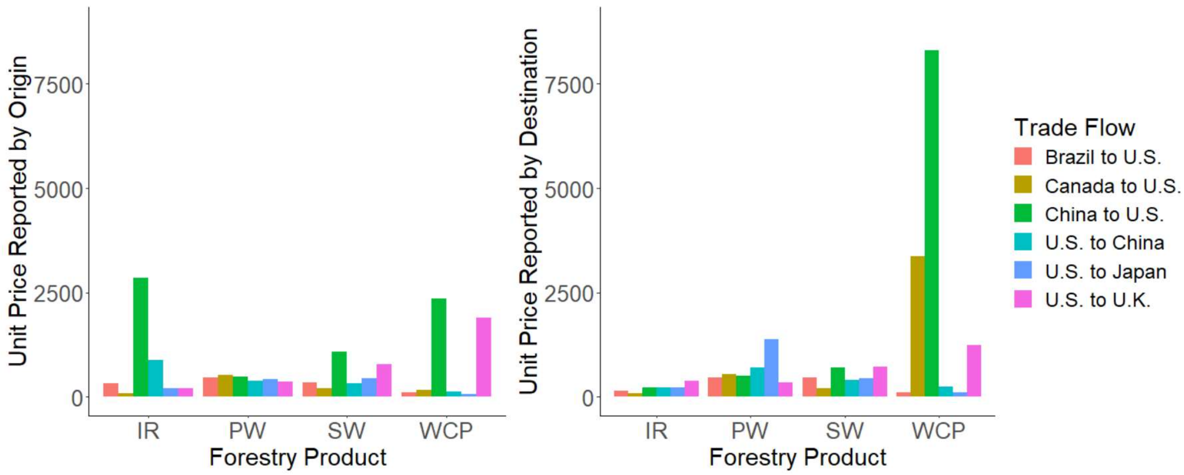

Figure 2.3. Average Unit Price per Forestry Product in Each Bilateral Flow.

Source: FAO Forestry Trade Flows; International Trade Centre (ITC) Trade Map.

Note: Averages are attained from bilaterally traded forestry products from 1997-2021 and reported in USD. Exporter refers to the origin of the traded forestry product while importer refers to the destination country. For example, for Brazil to U.S. exporter is Brazil and importer is the U.S.

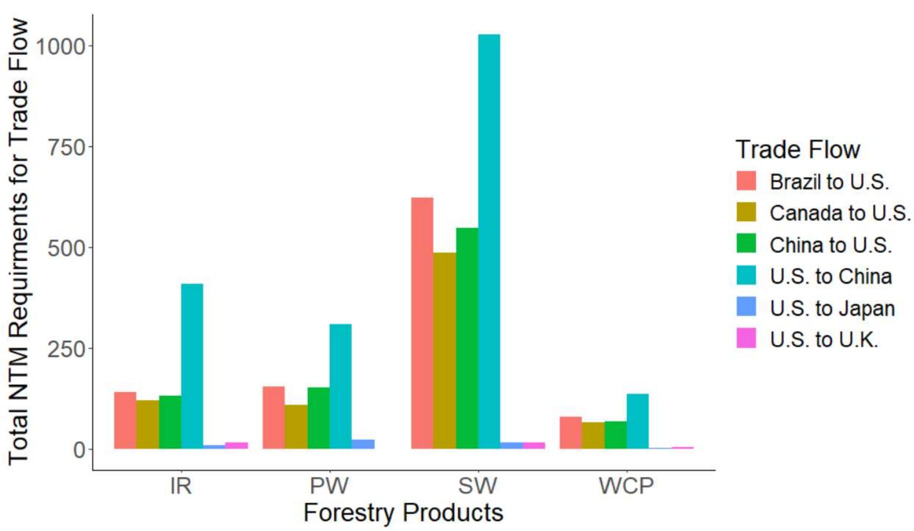

Figure 2.4. Total Non-Tariff Measures (NTM) Regulations for Trade.

Source: International Trade Centre (ITC) Market Access Map; WTO Integrated Trade Intelligence Portal.

Note: Total NTMs are aggregated across all HS codes for each forestry product and includes requirements for importing and exporting country.

Instead of exploring the potential causes of discrepancies, Ali et al. (2019) asks whether sector trade data discrepancies found in Malaysian trading partner pair data vary among different partner country pairs, vary among sectors, and change over time. They use Pearson's product-moment correlation coefficient to measure the ability of one country's reported trade to explain the values of its partner's.

2.3. Fixing Discrepancies

Buongiorno and Johnston (2018) advise that data uncertainty can pose as an issue on long-term projections in Forest Sector Models (FSMs) because any error in the historical data will be retained in parameter estimates, initial conditions of the model, and carried into projections and scenario analysis. They suggest that researchers should practice caution when using trade data provided by databases like the FAOSTAT, and the consequences of the uncertainty should be fully understood. When it comes to mending the discrepancies found in trade statistics, it seems that there is no option other than improving the data collection and/or estimation methods.

Carfagna et al. (2013) discuss that accurate and timely data is necessary for knowledge-based planning, facilitating rural development, and the reduction of poverty and food insecurity, especially in developing countries. Despite its importance, the quality and quantity of the statistics are low because traditional estimation methods that use probabilistic sampling can be too expensive and infeasible for some developing countries. To improve the quality and quantity of the data they suggest adopting new technologies like Geographic Information Systems (GIS), Global Positioning Systems (GPS), integrating various administrative databases to create a pre-

census list of farmers, and adding an agricultural module questionnaire to the existing population census. In their analysis of international forest statistics, Kallio and Solberg (2018) state that a lack of reliable and official data is obvious and improving the data should be of primary importance, but do not provide suggestions for improvement. Buongiorno and Johnston (2018) also suggest that improvements in data collection are needed, as well as further econometric research and scenario-building methods.

When examining forest sector statistics in Russia, Pzyhev et al. (2021) similarly call for an improvement of the volume and quantity of forestry statistics. They recommend improvements will need to be made via a collaboration between academia and business communities representatives. Ferrantino and Wang (2008) suggest that many discrepancies most likely arise from reporting methods and errors; they recommend cross-checking statistical results with persons involved in customs agency operations to get an insider's view of reporting practices. Pardo-Herrera (2021) recommends improving and increasing audits performed by authorities, continuous training of customs staff for reporting purposes, and an increase in country tax exchange partnerships. They also propose that a collaborative approach led by the National Statistics Office in each country to combine data from tax agencies, financial intelligence units, customs officials, and forest inspection agencies might help to identify potential illegal activities.

Chapter 3. Methodology

This chapter presents the methodology for this thesis. Section 3.1 will discuss the methodological framework that will be executed to analyze the forestry product trade data for each country's unilateral trade and production statistics. Section 3.2 will review the methodological framework used to analyze the bilateral forestry product trade statistics for each trade flow pair; Canada to the U.S, Brazil to the U.S, China to the U.S, the U.S. to China, the U.S. to Japan, and the U.S. to the U.K. Section 3.3 discusses factors that have been found to affect discrepancies in bilateral trade flow and their expected affects.

3.1. Unilateral Trade Data Analysis Methodological Framework

Section 3.1 will discuss the mathematical equations that will be used to identify discrepancies that are present in each individual country of interest's import, export, and production statistics. The methods presented in Section 3.1 will be performed on each of the six countries proposed to be examined in this thesis: the United States, China, United Kingdom, Japan, Canada, and Brazil.

The first mathematical equation to be defined is Excess Trade. Excess Trade is present when a country's export of a particular forestry product exceeds its production of that forestry product (Ferrier, 2021). Excess Trade will be applied to each of the six countries of interest for each of the forestry products of interest. Excess Trade will be defined as EX1. The mathematical formulation is expressed in Equation (1):

$$EX1 = \frac{X_{ipt}}{y_{ipt}}, \tag{1}$$

where X_{ipt} is the exports reported by country, i, of product, p, in the year, t. The production in the country, i, of product, p, in the year, t, is denoted by y_{ipt}. When EX1 is greater than one, exports of product, p, in the year, t, exceed the production.

Extending the Excess Trade formula to include imports leads to the Net Excess Trade equation. Net Excess Trade will be defined as EX2. Similar to EX1, EX2 will also be applied to each of the six countries and products of interest and then to aggregated country regions. The mathematical equation used to represent EX2 is defined in Equation (2):

$$EX2_{ipt} = \frac{X_{ipt} - M_{ipt}}{y_{ipt}}, \qquad (2)$$

where X_{ipt} is the exports of from country, i, of product, p, in the year, t, as reported by country i. Imports to country, i, of product, p, in the year, t, as reported by country i will be defined as M_{ipt}. Production in the country, i, of product, p, in the year, t, reported by country i is denoted as y_{ipt}. When EX2 is greater than 1 indicates that a country is exporting more than it imports and produces (Ferrier, 2021). When EX2 is negative, there were more imports than reported exports in that year, and if EX2 is less than negative one then imports were greater than exports and production. A discrepancy exists if EX2 is greater than one. If EX1 is greater than one, but EX2 is not then no discrepancy exists.

3.2. Bilateral Trade Analysis Methodological Framework

First, a discussion of how Pearson's Correlation Coefficient will be used to provide an initial analysis of bilateral trade flow. The GAP equation will then be defined as the measurement used for discrepancy. Followed by a discussion of the economic model that will be used to analyze the relationship between imports and exports in each bilateral trade country pair. Afterwards, the general econometric model formulated to correspond with the economic model will be

discussed. Lastly, an overview of the specific times series econometric models that will be utilized to explain each bilateral trade relationship will follow brief explanation of the methodological framework for econometric model selection.

3.2.1. Correlation

To further explore the trading partner flows, the Pearson product-moment correlation coefficient formula will be applied. Correlation can be used to measure bivariate association (Rodgers and Nicewander, 1998). The correlation between exports reported by an origin country and imports reported by the destination country can be used to determine whether exports of a product recorded by the origin country explain the imports of that product recorded by the destination country (Ali et al., 2019). The value ranges from -1 to 1, where 1 indicates a strong positive linear association and -1 indicates the presence of a strong inverse relationship, 0 indicates no linear relationship (Hill et al., 2018). If exports can explain imports, a strong positive relationship is expected. If no relationship exists, then it is reasonable to believe that the trading partner relationship is discrepant. The correlation formula is shown in Equation (3) below:

$$r_{ijp} = \frac{\sum_t^T \left(\left(X_{ijpt} - \overline{X_{ijp}} \right) \left(M_{jipt} - \overline{M_{jip}} \right) \right)}{\sqrt{\sum_t^T \left(\left(X_{ijpt} - \overline{X_{ijp}} \right)^2 \left(M_{jipt} - \overline{M_{jip}} \right)^2 \right)}}, \tag{3}$$

where r_{ijp} is the correlation between exports of product p from the country, i, to country, j, and imports of the product, p, into the country, j, from the country, i. X_{ijpt} are exports of product, p, from the country, i, to country, j, in the year, t, as reported by country, i. M_{ijpt} are imports of the product, p, into the country, j, from the country, i, in the year, t, as reported by country, j. $\overline{X_{ijp}}$ and $\overline{M_{ijp}}$ are the mean values of X_{ijpt} and M_{ijp}, respectively.

3.2.2. GAP Analysis

The GAP equation will be used to identify discrepancies that might occur within trading partner flows. GAP will be used on three partner flows in which the U.S. is the importing country and three of which the U.S. is the exporting country. U.S. importing flows will be (1) Canada to the U.S, (2) Brazil to the U.S, (3) China to the U.S. The U.S. exportation flows will be examined between (1) U.S. to China (2) U.S. to U.K. (3) U.S. to Japan. GAP represents the difference between exports reported in the origin country, i, and imports reported by the destination country, j (Ferrier, 2021). The mathematical equation for GAP is defined below:

$$GAP_{ijpt} = M_{ijpt} - X_{jipt}. \tag{4}$$

Exports reported by country, i, to country, j, of product, p, in the year, t, is defined as X_{ijpt}. M_{jipt} represents the imports reported by country, j, from the country, i, of product, p, in the year, t. If GAP_{ijpt} is positive, a larger volume of traded good, p, is recorded by the importing country than the exporting country (Kee and Nicita, 2016)[8]. This result could indicate that the importer is overreporting or the exporter is underreporting. If no discrepancy occurs, GAP should be equal to 0.

3.2.3. Econometric Model

The econometric formulation that will be used to study the relationship between reported imports and exports of bilateral trade pairs is presented in Equation (5) (Kee and Nicita, 2016). The model can be interpreted as imports reported by country j of forestry product p in year t can be explained by the exports reported by country i of product p in time t, an intercept

[8] Kee and Nicita (2016) define the econometric model for GAP in Section 2 as a theoretical model in their research in relation to tariff and non-tariff measures. They use GAP to minimize heteroskedasticity and outliers while estimating the Ad Valorem Equivalent of Non-Tariff Measures.

parameter, β_0, and a random error at time t. If a bilateral trade pair flow's relationship is non-discrepant, it is expected that $\beta_0 = 0$ and $\beta_1 = 1$ and any errors from the non-discrepant relationship are random with a mean of zero.

$$M_{jpt} = \beta_0 + \beta_1 X_{ipt} + e_{Mt} \quad (5)$$

$$e_{Mt} = M_{jpt} - \beta_0 - \beta_1 X_{jip}. \quad (6)$$

In Equation (6), by rearranging the simple regression formula, the formula for the error term can also be considered an estimation for the GAP statistics defined in the previous section: if $\beta_0 = 0$ and $\beta_1 = 1$ then Equations (4) and (6) are the same. If $\beta_0 \neq 0$ there is some constant factor that is contributing to the discrepancies within the reported bilateral statistics, if $\beta_1 \neq 1$ then import and export statistics are not proportional to one another, and the estimated discrepancies are reflective of the non-proportional relationship between the imports and export statistics.

Ordinary least squares (OLS) will be used for parameter estimation and import, and exports statistics used in estimation are assumed to be stationary. If import and/or export quantity statistics are trend stationary, then a trend term will be added to Equation (5). Model diagnostics tests will be performed on estimated regression models for (trend) stationary bilateral trade pairs. The variance inflation factor (VIF) and condition index (CI) of each estimated regression will be observed to ensure no multicollinearity is present; if the VIF is ≤ 5 and the CI is ≤ 20 then there are no signs of multicollinearity in the model. Ramsey's reset test is used to check the null hypothesis that model is in correct functional form. The White's test and Breusch-Pagan tests will be used to test the null hypothesis that errors are homoscedastic; if the two tests are rejected, then heteroskedasticity is present and it must either be corrected for, or

heteroskedasticity consistent (HC) t-statistics are reported in the final estimation results. The null hypothesis of no autocorrelation among the errors is tested using the Durbin Watson (DW) test or Durbin Watson-h probability value if the model includes lagged dependent variables; if the DW test is rejected then the regression results in autocorrelated residuals. To test the null hypothesis of normally distributed errors, the Jarque-Bera test is used (Hill et al., 2018).

To properly model the regression relationships in the GAP regression, the variables must be checked for non-stationarity. If a time series variable's data set is weakly stationary, its mean and variance do not change over time (Enders, 2009)[9]. If the variables are nonstationary, but treated as such, the regression will lead to spurious results (Granger and Newbold, 1973). Spurious regression results are results that appear statistically significant when the dependent and explanatory variables are actually unrelated (Hill et al., 2018). Because each bilateral import and export pair is unique, they might not share the same unique time-series properties. For example, some series might be stationary in levels while some might be nonstationary in levels but integrated of the same order, while others may be integrated of different orders. If the import and export statistics are not stationary in levels, then alternate time-series models will be considered. Alternative time-series models are explained in more detail in Sections 3.2.4 – 3.2.5.

Augmented Dickey-Fuller (ADF) tests are used to determine whether one or both series in a bilateral series set are nonstationary. ADF tests are a set of statistical procedures that can be used to determine the order of integration for a time-series sequence by utilizing a series of steps

[9] From Enders (2019, Chapter 2, pg. 54) a time-series is weakly stationary if its mean and autocovariances do not change over time. For the remainder of this thesis, stationarity refers to weak stationarity.

to test whether there is a unit root present within the series (Enders, 2009)[10]. To employ the ADF tests, Equations 7-10 are estimated beginning with the least restrictive model:

$$\Delta y = \mathrm{a_0} + \gamma y_{\mathrm{t-1}} + \mathrm{a_2 t} + \sum_{\mathrm{i=2}}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t-i+1}} + \mathrm{e_t} \tag{7}$$

$$\Delta y = \mathrm{a_0} + \gamma y_{\mathrm{t-1}} + \sum_{\mathrm{i=2}}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t-i+1}} + \mathrm{e_t} \tag{8}$$

$$\Delta y = \gamma y_{\mathrm{t-1}} + \sum_{\mathrm{i=2}}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t-i+1}} + \mathrm{e_t} \tag{9}$$

$$\phi_i = \frac{SSR_{Restricted\ Model} - SSR_{Unrestricted\ Model})/restrictions}{SSR_{Unrestricted\ Model}/(T - K\ parameters)}. \tag{10}$$

The t-statistics, or tau, from the parameter estimations are compared to Dickey Fuller critical values to test the null hypothesis $\gamma = 0$, or unit root. The rejection region can be defined as reject the null hypothesis if the t-statistic is greater than critical tau value. The ADF testing procedure begins with the estimating the least restrictive model Equation (7). If $H_0: \gamma = 0$ in Equation (7) cannot be rejected, then the phi statistic (ϕ_3) defined in Equation (10) is used to test joint hypotheses $H_0: \gamma = a_2 = 0$[11]. For the joint hypothesis test, critical phi statistics given by Dickey Fuller, and the rejection region can be defined as reject the null hypothesis if $\phi_i > \phi_C$. If $H_0: \gamma = a_2 = 0$ is not rejected, then the next restricted model, Equation (8), is estimated[12]. From Equation (8) $H_0: \gamma = 0$ is tested where a rejection implies no unit root, and failure to reject leads

---

[10] Lag-orders for ADF tests are determined by minimizing the AIC.

[11] If $H_0: \gamma = 0$ is rejected then conclude no unit root.

[12] If $H_0: \gamma = a_2 = 0$ is rejected, then $H_0: a_2 = 0$ is tested using the t-distribution where a rejection leads to a test on $H_0: \gamma = 0$ using the t-distribution. If $\gamma = 0$ using the t-distribution is not rejected then the series is trend stationary.

to the next joint hypothesis (ϕ_1), $H_0: \gamma = a_0 = 0$. If the ϕ_1 test cannot be rejected then the most restrictive model, Equation (9) is estimated where $H_0: \gamma = 0$ is tested again. The tests are conducted until either $\gamma = 0$ cannot be rejected for the most restrictive estimated model Equation (9), implying the series has a unit root (non-stationary); or until $\gamma = 0$ is rejected meaning that the series does not have a unit root or is stationary.

Several outcomes could result from the ADF tests which will each require specific time-series econometric modeling considerations. If the both the import and export statistics are found to be stationary in levels, they are integrated of order zero, I(0); and they can be estimated as described by the econometric model from Equation (4). If either the import or export or both are trend stationary, then they can be estimated with the possible inclusion of a trend variable. Import and export statistics within a bilateral set that are found to be integrated of the same order can be checked for Johansen cointegration. Before cointegration is tested for, the vector autoregressive (VAR) model order used for the cointegration test is determined using Likelihood Ratio (LR) Tests.

The likelihood ratio statistic is used to determine the Vector Autoregressive (VAR) model order to estimate when conducting Johansen Cointegration Tests. The test statistic follows a χ^2 distribution and degrees of freedom is the number of restrictions in the hypothesis (Enders, 2009). It should be noted that other appropriate methods for model selection for VAR include Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and other similar statistics. To utilize the likelihood ratio statistic for VAR order selection the testing scheme described by Lütkepohl (1993) is used where 4 is the largest VAR lag-length to be tested[13]. If

[13] The largest VAR lag-length to be tested is chosen to be four based on Enders, 2009, Chapter 6, page 397.

imports and exports are integrated of different orders but cannot be specified using a VAR representation, then ARDL Bounds cointegration can be conducted. If the imports and exports are integrated of different orders, determined by the ADF tests, and can be specified as VAR based on LR tests, then Granger non-causality will be tested utilizing augmented VAR estimation discussed Section 3.2.6.

Once the appropriate VAR order is selected for each I(1) bilateral import and export pair, Johansen Cointegration Tests can be performed. If the import and exports from a bilateral trade flow are cointegrated then they share a common stochastic trend. In other words, if imports and exports are both I(1), it is expected that regressing imports over exports would result in an error that is I(1), but if that error term is stationary, then the two series are cointegrated and share some sort of long-run equilibrium. In terms of discrepancies, where the error term is the estimated discrepancy[14], if imports and exports of a series set are cointegrated then in the long-run, their resulting estimated discrepancy series are deviations from their long-run equilibrium. Therefore, cointegration is important to test whether a long-run relationship does exist between imports and exports for a particular forestry product for a bilateral trade partner pair. If the statistics are not cointegrated, then a long-run relationship does not exist between the two. The Johansen Cointegration test relies on characteristic roots and rank of a matrix to determine whether the import and export variables are cointegrated (Johansen and Juselius, 1990). The Johansen tests will be performed using SAS VARMAX procedure. To illustrate the cointegration test, consider Equations (11) and (12):

$$y_t = A_1 y_{t-1} + \varepsilon_t, \tag{11}$$

[14] From Equation (5).

subtracting both sides of the equation by y_{t-1} leads into Equation (12),

$$\begin{aligned} \Delta y_t &= A_1 y_{t-1} - y_{t-1} + \varepsilon_t \\ &= (A_1 - I)x_{t-1} + \varepsilon_t \\ &= \pi x_{t-1} + \varepsilon_t, \end{aligned} \tag{12}$$

where y_t is an $(n \cdot 1)$ vector of reported import and export statistics; A_1 is an $(n \cdot n)$ matrix of parameters and I is an identity matrix. For this research, n = 2 for the cointegration test for the relevant products for each bilateral trade flow pair. In Equation (12), $\boldsymbol{\pi} = \boldsymbol{\alpha\beta'}$ where π is the $(n \cdot n)$ long-run impact matrix, β is the $(n \cdot r)$ cointegrating vector, α is $(n \cdot r)$ vector of weights, and r is rank (Johansen and Juselius, 1990). The purpose of the Johansen cointegration process is to determine if the π matrix contains information about the long-run relationship between reported import and export statistics by examining its rank. If the rank of π is full rank, or equal to the number of variables (in this case two), then the variable vector is stationary, and imports and exports are not cointegrated. If the rank of π is less than full rank, but greater than zero then the two series are cointegrated. If the rank of π is zero, then the two series are not cointegrated and not stationary (Enders, 2009)[15]. To determine the rank and the significance of the rank of π, two statistics are used for hypothesis testing, the Trace and Maximum Eigenvalue statistics, defined in Equations (13) and (14) respectively:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \tag{13}$$

$$\lambda_{max}(r, r+1) = -T ln\left(1 - \widehat{\lambda_{r+1}}\right), \tag{14}$$

[15] Chapter 6 pages 390-391.

where r refers to the rank of the π and $\widehat{\lambda_i}$ are its estimated characteristic roots or eigenvalues. The trace statistic is used to test the null hypothesis that the rank is equal to r against the alternative that rank is greater than r. The maximum statistic tests the null hypothesis that the rank is equal to r against the alternative that the rank is equal to r+1. If the import and export statistics for a product from a particular bilateral trade flow set are cointegrated then a vector error correction model (VECM) will be used to estimate the relationship, and long-run equilibrium between imports reported by a destination country and exports reported by the origin country.

The Auto Regressive Distributed Lag (ARDL) cointegration, or bound cointegration testing technique will be applied to forestry product trade quantity sets integrated of different orders. The ARDL cointegration technique is applicable to determine the long-run relationship between such series of differing integration orders as long as the series are not I(2) (Nkoro and Uko, 2016; Pesaran et al, 2001). The ARDL model can be estimated using OLS as long as the error term is a white noise process. It is assumed for the purpose of this testing procedure that import quantity is dependent on export quantity so that the model can be specified as in Equation (15) which offers a computationally convenient way to estimate the ARDL error correction model (Kripfganz and Schneider, 2022):

$$\Delta M_t = \delta_0 + \sum_{i=1}^{p-1} \lambda_i \Delta M_{t-i} + \sum_{i=0}^{q-1} \delta_i \Delta X_{t-i} + \varphi_M M_{t-1} + \varphi_X X_{t-1} + u_t. \tag{15}$$

The orders for p and q are chosen by minimizing the AIC of all potential lag orders. Within the ARDL model specification, φ_1 and φ_2 correspond to the long-run relationship between variables while λ_i and δ_i are the short-run dynamics. The first step of the ARDL cointegration test process is the F-bounds hypothesis test cointegration between the import and export statistic

tests through the null hypothesis that $\varphi_M = \varphi_X = 0$ or that a long-run relationship does not exist against the alternative hypothesis of cointegration or $\varphi_M \neq \varphi_X \neq 0$ using a F-statistic or Wald test. The distribution of the test statistic is non-standard and critical bounds for the F-statistic are given by Pesaran et al (2001) and calculated based on the number of variables in the system and whether the model includes an intercept and/or trend.[16] If the calculated F-statistic is lower than the bounds, then it can be concluded that the variables are I(0) and not cointegrated; if the F-statistic is greater than the bounds, then the null hypothesis is rejected, and it can be concluded that cointegration exists; and if the F-statistic falls between the bounds, then the test is inconclusive. If the null hypothesis of the F-bounds test is rejected, then the next step is to test the null hypothesis that $\varphi_M = 0$ against the alternative that $\varphi_M \neq 0$ using the conventional t-statistic with a non-standard distribution where critical values are again provided by Pesaran et al (2001). If both the F-bounds and T-bounds tests are rejected then the third step is to test the hypothesis that $\theta = 0$ against the alternative that $\theta \neq 0$ using a Wald test (Kripfganz and Schneider, 2022). This test is performed on θ instead of φ_X because the OLS estimator of θ is asymptotically normally distributed; the formula for reparameterization of φ_X into θ is in Equation (16) where α is the speed-of-adjustment parameter (Kripfganz and Schneider, 2022; Nkoro and Uko, 2016):

$$\alpha = -\varphi_M \qquad (16)$$
$$\theta = \frac{\varphi_X}{\alpha}.$$

[16] it is important to note that these critical values are not always appropriate for small sample sizes, but critical values for small sample sizes are available from Narayan (2005).

If the null hypothesis of the F-bounds, T-bounds, and $\theta = 0$ tests are rejected, then the ARDL ECM can be estimated using OLS. The ARDL ECM specification is discussed in Section 3.2.5.[17]

3.2.4. Vector Error Correction Model

By utilizing the VECM approach, the long-run GAP or discrepancy can be estimated via the error correction term, and each country in a bilateral pair's response to discrepancies can be evaluated. The vector error correction model (VECM) is formulated in Equation (17) (Johansen and Juselius, 1990)[18]:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \varepsilon_t$$
$$\alpha\beta' y_{t-1} = \alpha_i(\beta_M M_{t-1} - \beta_X X_{t-1}). \quad (17)$$

The parameters of interest here are β which is the cointegrating vector and α which is the speed of adjustment parameter. For the case of imports and exports reported by bilateral trade flow partners, the generalized error correction term is defined where M is imports and X is exports. The error correction term is the same whether the dependent variable is imports or exports, and only the alpha parameter will differ. The error correction term demonstrates the long-run equilibrium shared by imports and exports, and the alpha parameter is each respective reporter's estimated response when the error correction term is positive (M>X). For this research, imports are normalized where $\beta_M = 1$; so, from this model, it is clear that the error correction term has the same formulation as GAP, or discrepancy. The VECM allows us to

[17] Determining the relevant case in which to model the ARDL ECM can be deduced by visual inspection of the time series data; testing for no cointegration and no serially correlated errors of each suspected case; and performing likelihood ratio tests with limiting chi-square distribution with n-r degrees of freedom where r = 0,…,n-1 (Pesaran et al, 2001; Pesaran et al, 1998; Kripfganz and Schneider, 2022; Nkoro and Uko, 2016).

[18] Although VECM have not been utilized for bilateral trade data statistics thus far, there are instances in literature in which researchers utilize the estimation method for unilateral trade statistics in relation to economic growth (Lawal and Ezeuchenne, 2017; Md Reza et al., 2019).

estimate the discrepancy error, and each reporter's response to discrepancies occurring. There are five potential cases of the VECM that might occur: the first case is that there is no separate drift or constant term in the VECM (Equation (8)), the second case is that a constant enters only through the error correction term, the third case is that there is a separate constant or drift in the VECM that does not enter the error correction term, the fourth case is that there is a linear trend in the error correction term, and the fifth case is where a separate linear trend is present in the VECM. For this research, for each eligible forestry product trade statistic pair, the results from the ADF test will guide which VECM case to estimate, and a likelihood ratio test is preformed using the statistic in Equation (9) to test between the restricted and unrestricted applicable case:

$$-T\left(\ln\left(1-\widehat{\lambda_R}\right)-\ln\left(1-\widehat{\lambda_{UR}}\right)\right), \tag{18}$$

where $\widehat{\lambda_\iota}$ is the estimated eigenvalue from the restricted (R) and unrestricted (UR) models. For example, this hypothesis can be used to test the null that the restricted model, or case two, performs as well as the unrestricted or case three. [19] If the null hypothesis is rejected, then the VECM with a drift in the model and no separate trend in the error correction term is used. If the null hypothesis is not rejected, then a VECM with a drift inside the error correction term is used.

Following the aforementioned likelihood ratio tests, once the case in which to model each bilateral trade flow statistic pair for the relevant products and country flow pairs has been determined, the VECMs will be estimated using Maximum Likelihood Estimation (MLE) with SAS's VARMAX procedure. Hypothesis tests are conducted on the estimated alphas (speed of

[19] For example, if the ADF results show that a variable is I(1) with a trend then cases four and five can be estimated. Then the likelihood ratio test can be used to test the restricted case four against the unrestricted case five.

adjustment), and beta (cointegrating relations) parameters. If a case two VECM is estimated, then the constant within the error correction term will be checked for significance using the likelihood ratio test statistic in Equation (19). The null hypothesis to test the significance of the constant within the error correction term illustrated in Equation (20), is $\beta_0 = 0$, against the alternative $\beta_0 \neq 0$. If the null hypothesis is rejected then the error correction term does include a constant and it can be observed that the estimated discrepancy does not adhere to the GAP formulation; so, in that case the relationship between import and export statistics might be inherently discrepant. If the null hypothesis is not rejected then the error correction term does not contain a constant which can imply that in the long run, the import and exports return to a long run equilibrium in which $M - X = 0$. In other words, if the null hypothesis that $\beta_0 = 0$ is not rejected then the estimated discrepancy adheres to the GAP formulation. If the null hypothesis cannot be rejected, then the VECM is re-estimated without a constant within the error correction term.

$$T\left(\ln\left(1 - \widehat{\lambda_R}\right) - \ln\left(1 - \widehat{\lambda_{UR}}\right)\right) \tag{19}$$

$$\alpha\beta' y_{t-1} = \alpha_n(\beta_M M_{t-1} - \beta_0 - \beta_X X_{t-1}). \tag{20}$$

Within the error correction term, a hypothesis test is also conducted on the β_X parameter. If the relationship between reported import and export quantities are equal to one another in the long run, then it is expected that $\beta_X = -1$. If $\beta_X \neq -1$ then import and export quantities are not equal to one another in the long run; and one side is either overreporting or underreporting. The null hypothesis $\beta_X = -1$ against the alternative that $\beta_X \neq -1$ is also tested using the likelihood ratio statistic in Equation (20).

The alpha parameter, or speed of adjustment for each relationship is tested using Chi-squared test. For each VECM the two null hypotheses, $\alpha_M = 0$ and $\alpha_X = 0$ are tested against the

alternatives $\alpha_M \neq 0$ and $\alpha_X \neq 0$, respectively. If the null hypothesis $\alpha_M = 0$ ($\alpha_M = 0$) is not rejected, then importer (exporter) side statistics do not respond to deviations from the long-run estimated relationship. In other words, if an alpha parameter is not statistically significant then it can be concluded that the respective reporter does not experience any adjustment in statistics when estimated discrepancies occur. Model diagnostic tests will be conducted on the VECMs.

3.2.5. Auto-Regressive Distributed Lag Error Correction Model

The ARDL ECM is specified in Equation (21) and is similar to the error correction component defined in the VECM section where the error correction term (ECT) replaces the long-run terms as the estimated residuals from the long-run model. By utilizing the ARDL ECM estimation approach, the estimated long-run GAP or discrepancy can be identified as the ECT, and any adjustments that the importer's statistic might undergo in response to discrepancies can also be identified through the speed of adjustment parameter[20].

Once the ARDL ECM is estimated, inferences on both the long-run and short-run relationships between import and export statistics can be conducted. In Equation (21) α is the speed of adjustment parameter. The ECT is specified in Equation (22) where θ was defined in Section 3.2.3. Note that the specification in Equation (21) and (22) is for an ARDL ECM with an unrestricted intercept and no trend:

$$\Delta M_t = \delta_0 + \sum_{i=1}^{p} \lambda_i \Delta M_{t-1} + \sum_{i=0}^{q} \delta_i \Delta X_{t-i} + \alpha ECT_{t-1} + u_t \tag{21}$$

[20] To my knowledge, ARDL error correction models have not been used on bilateral trade statistics but have been utilized in exploring unilateral trade statistics in relation to economic growth and financial development (Iheanacho, 2017; Ahad, 2017).

$$ECT_{t-1} = \varepsilon_{t-1} = M_{t-1} - \theta X_{t-1}. \quad (22)$$

Once the ARDL ECM is estimated, model diagnostics can be performed to ensure that no autocorrelation, heteroskedasticity, or misspecification problems are present. Models are checked using the Ljung-Box and Breusch-Godfrey test for no autocorrelation, Breusch-Pagan test for no heteroskedasticity, Shapiro-Wilk test for normality, and Ramsey Reset test for correct functional form. Cumulative sum of recursive residuals (CUMSUM) and CUMSUM of square (CUSUMSQ) plots are examined to test the model's stability (Iheanacho, 2017; Brown et al. 1974; Pesaran et al, 2001). R statistical software is used for the ARDL bounds cointegration procedure and ECM specification. Packages ARDL, dLAGM, and manual calculations are utilized to cross check and confirm results. Hypothesis tests on the long-run coefficient reparameterizations are conducted in SAS. Granger-causality tests can be conducted on the ARDL ECMs. Granger causality tests can be conducted on the short-run parameter by testing the null hypothesis that $\delta_i = 0$, tests can be conducted on the long-run parameter by testing $\alpha = 0$, and strong granger causality can be tested through the hypothesis $\delta_i = \alpha = 0$ (Hung-Pin, 2014).

3.2.6. Non-Causality Tests Using Augmented VAR

Granger non-causality can be tested on a bilateral trade statistic set through the estimation of an augmented VAR regardless of the order of integration. The test assumes that X causes Y, or in this scenario, exports cause imports. An augmented or over-fitted VAR is estimated then tests on restrictions of the parameters are conducted using a modified Wald (MWALD) test where the Wald statistic has an asymptotic chi-squared distribution with k degrees of freedom (Guru-Gharana, 2012). The Augmented VAR is estimated as a VAR[k+dmax] where dmax is the maximum order of integration from the bilateral series set and k is the optimal number of lags

for the VAR as determined by the likelihood ratio test. The first step for the procedure is to determine the maximum order of integration for the series set by means of ADF tests discussed in Section 3.2.3. Next by using the LR tests discussed in Section 3.2.3 the true lag-length, k, can be determined (Guru-Gharana, 2012; Zapata and Rambaldi, 1997). The next step is to estimate the unrestricted VAR(k + dmax) in levels using a suitable estimation technique. For the purposes of this analysis, OLS is used. Lastly, the MWALD test can be conducted which involves testing the significance of the k lags while ignoring the last dmax lag (Guru-Gharana, 2012; Zapata and Rambaldi, 1997). The Augmented VAR is formulated in Equation (23) and the hypothesis tests are specified in Equation (24):

$$\begin{bmatrix} M_t \\ X_t \end{bmatrix} = \begin{bmatrix} \alpha_M \\ \alpha_X \end{bmatrix} + \sum_{i=1}^{k} \begin{bmatrix} \beta_{MM,i}\beta_{MX,i} \\ \beta_{XM,i}\,\beta_{XX,i} \end{bmatrix} \begin{bmatrix} M_{t-i} \\ X_{t-i} \end{bmatrix} + \sum_{i=k+1}^{k+d} \begin{bmatrix} \gamma_{MM,i}\gamma_{MX,i} \\ \gamma_{XM,i}\,\gamma_{XX,i} \end{bmatrix} \begin{bmatrix} M_{t-i} \\ X_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Mt} \\ \varepsilon_{Xt} \end{bmatrix} \qquad (23)$$

$$\begin{gathered} H_0\colon \beta_{MX,i} = 0\ for\ all\ i\ \leq k\ \rightarrow Exports\ do\ not\ Granger\ cause\ Imports \\ H_0\colon \beta_{XM,i} = 0\ for\ all\ i\ \leq k \rightarrow Imports\ do\ not\ Granger\ cause\ Exports, \end{gathered} \qquad (24)$$

where k[21] is the optimal lag order and (k+d) is the extra lag for dmax, the error terms ε_{Mt} and ε_{Xt} are assumed to be normally distributed with no serial correlation based on Portmanteau tests. If the relationship between imports and exports is not discrepant then it is expected that exports will granger cause imports. A stronger indication that the relationship is non-discrepant is that exports Granger cause imports and imports granger cause exports. If neither hypothesis test can be rejected, then the series set must be discrepant.

[21] Optimal lag length order (k) for VAR is determined through LR testing scheme described by Lütkepohl (1993).

Chapter 4. Data

An overview of the data needed to follow the methodological framework described in Section 3.1 is discussed in the proceeding subsections. The relevant data includes forestry sector trade data, information on the distance between trading partners, indicators that represent the corruption within a country, and tariff data.

4.1. FAOSTAT Forestry Data

The primary data source for this proposed thesis will be the FAOSTAT Forestry Database which consists of an annual time series of data on production, imports, and exports of Forest Sector products in each country starting in 1961 and on the trade flows of partner countries beginning in 1997 (FAO, 2020). The data used for Excess Trade (EX1) and Net Excess Trade (EX2) will be from the database's annual series of data on production, imports, and exports. The data used for the GAP and Correlation Methods will come from the trade flows of partner countries' time series set. The data is available for download as an Excel file.

FAO collaborates with the United Nations Economic Commission for Europe (UNECE), the Statistical Office of the European Union (Eurostat), and the International Tropical Timber Organization (ITTO) to collect survey data on the import, export, and production statistics of forestry industry products (FAO, 2020). The survey is referred to the Joint Forest Sector Questionnaire (JFSQ), and each agency is responsible for administering the questionnaire to a respective region in order to alleviate the responsibility from falling on one single agency. The data is collected and disseminated through the FAOSTAT database and other publications (Steele

et al., 2021). Practitioners of the JFSQ check the responses, throw out gross misestimations[22], and estimate missing data as they see fit (FAO, 2020). The JFSQ does ask for the source of respondents' data; however, practitioners do not have time to check each country's estimation procedures (Steele et al., 2021). Therefore, discrepancies in the data could occur at any point during the data collection process.

Utilizing trade data for analyses has several shortfalls that challenge researchers. First, for some countries, when values are unknown for a certain year, the value from the previous year is used (Skejerstad et al., 2020; FAO, 2020). This practice is applied at the discretion of the FAOSTAT data aggregating practitioner (FAO, 2020). Repeating the values of previous years removes the variance in the unknown data point that researchers might wish to observe (Skejerstad et al., 2021). Another challenge when using trade data is that it contains zero values for some imports, exports, and production: dropping the zero-valued observations is one way to handle the issue (Jayasinghe et al., 2010; Dal Bianco et al., 2016; Kazunobu, 2020). However, this might not be the best approach because the zero values represent the absence of trade rather than missing values so, they are important and require examination (Jayasinghe et al., 2010). Researchers have suggested other options to combat this issue include replacing the zero with a small arbitrary number or the use of the pseudo-Poisson maximum likelihood estimator, which can admit zero observations (Jayasinghe et al., 2010; Dal Bianco et al., 2016).

[22] Gross misestimations are determined at the discretion of the survey practitioner.

4.2. International Trade Centre (ITC)

The International Trade Centre (ITC) disseminates trade quantities and prices for bilateral trade flows. Trade data from 2018-2021 is collected from the ITC Trade Map to supplement the FAOSTAT trade data collected for each of the bilateral trade flows. ITC Trade Map derives its data from several sources. Trade data for the U.S. is derived from the U.S. Census Bureau and UN COMTRADE databases; trade data for Canada is derived from Statistics Canada and UN COMTRADE databases; trade data for Brazil is derived from Ministério do Desenvolvimento, Indústria e Comércio Exterior and UN COMTRADE databases; trade data for China is derived from the General Customs Administration of China and UN COMTRADE databases; trade data for Japan is derived from the Japanese Ministry of Finance and UN COMTRADE databases; and trade data for the U.K. is derived from the UN COMTRADE database.

ITC Trade Map allows users to look up trade data quantities and prices by Harmonized System (HS) codes. The HS is a standardized numerical method to classify traded goods used by customs authorities around the world (Pardo-Herrera, 2021). HS codes are updated every five years, so several HS versions will be employed for this research include 1996, 2002, 2007, 2012, and 2017. ITC Trade Map provides a HS correspondence flow chart for the ease of tracking HS code changes. The HS product codes that correspond to the forestry products in this study for the 2017 revision are listed in Table 4.1 below. The total number of HS codes in each revision year is listed in Table 4.2.

Table 4.1. Harmonized System (HS) Product Code Classifications for Forestry Products.

Forestry Product	HS 2017 Classification
Industrial Roundwood, Coniferous	4403.11/21-26
Industrial Roundwood, Non-Coniferous, non-Tropical	4403.91/93-99
Industrial Roundwood, Non-Coniferous, Tropical	4403.12/41/49
Sawnwood, Coniferous	4406.11/91 4407.11/12/19
Sawnwood, Non-Coniferous	4406.12/92 4407.21/22/25-29/91-97/99
Wood Chips, Particles and Residues	4401.21/22/40
Plywood	4412.31/33/34/39/94/99

Source: (FAO, 2020) pg. 32-33.

Table 4.2. Harmonized System (HS) Codes for Forestry Products: Number per Revision Year.

	Revision Year:					
Forestry Product:	2017	2012	2007	2002	1996	Average
Industrial Roundwood	18	7	7	7	7	9
Sawnwood	22	14	14	8	8	12
Plywood	6	5	5	9	9	6
Wood Chips and Particles	2	2	2	2	2	2

Source: ITC Trade Map.

Chapter 5. Results and Discussion

In this Chapter, the results from the unilateral and bilateral trade data discrepancy analysis will be reported and discussed. First the results from the unilateral trade data analysis, EX1 and EX2, will be reviewed. Then the bilateral trade data analysis will commence starting with an overview of the descriptive statistics for the import and export reported quantities for each bilateral trade flow. Followed by the correlation analysis between imports and exports of each bilateral trade country pair. Then the GAP analysis will be discussed. Next the results from each estimated model for bilateral trade quantities of each forestry product will be presented. The last section in this chapter is a discussion of the bilateral trade partner flows and forestry products analysis results, and how results might be related to discrepancy inducing factors discussed in reviewed literature.[23]

5.1. Unilateral Trade Analysis

EX1 and EX2 were used to analyze each country's unilateral forestry product export quantities and net export quantities and their relation to the country's production of that forestry product each year. Table 5.1 reports the EX1 and EX2 results for the U.S, Canada, Brazil, China, Japan, and the U.K. A brief discussion of the results follows Table 5.1.

[23] From Figures 2.1, 2.2, 2.3, and 2.4.

Table 5.1. Excess Trade (EX1) and Excess Net Trade (EX2) Results.

Country:	Forestry Product:	Percentage of Years greater than 1 EX1	EX2	Average Percentage EX1	EX2
United States	IR	0%	0%	3.54%	3.02%
	SW	0%	0%	5.94%	-22.73%
	PW	0%	0%	4.13%	-14.30%
	WCP	0%	0%	10.25%	9.65%
Canada	IR	0%	0%	2.47%	-0.21%
	SW	0%	0%	66.68%	63.34%
	PW	0%	0%	25.97%	-2.12%
	WCP	0%	0%	1.77%	-5.23%
Brazil	IR	0%	0%	0.31%	0.26%
	SW	0%	0%	11.43%	10.47%
	PW	0%	0%	34.57%	34.49%
	WCP	6%	6%	39.44%	39.41%
China	IR	0%	0%	0.19%	-15.46%
	SW	0%	0%	2.92%	-17.09%
	PW	0%	0%	38.69%	8.19%
	WCP	0%	0%	16.30%	-27.05%
Japan	IR	0%	0%	0.58%	-69.67%
	SW	0%	0%	0.41%	-34.54%
	PW	0%	0%	3.98%	-59.62%
	WCP	0%	0%	0.05%	-380.55%
United Kingdom	IR	0%	0%	3.77%	-10.33%
	SW	0%	0%	4.44%	-444.22%
	PW	71.67%	0%	>500%	<-500%
	WCP	0%	0%	7.22%	0.16%

Note: Abbreviations are for industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP). IR, SW, and PW unilateral data ranges from 1961-2020 (59 yrs.) while WCP data ranges from 1998-2020 (23 yrs.)

The EX1 and EX2 analysis for unilateral trade flow for the U.S., Canada, Brazil, China, Japan, and the U.K in Table 5.1 revealed that the data sets for trade and production statistics for industrial roundwood, sawnwood, plywood and wood chips and particles all are discrepancy free with one exception. Wood chips and particles trade and production statistics reported by Brazil did contain a discrepancy in one (6%) of the 23 reported years where net exports exceeded

production (EX2 > 1) . Because Brazil also reports excess exports of wood chips and particles exceeded production (EX1 > 1) in the same year as EX2 > 1, it is possible that Brazil imported wood chips and particles under the wrong HS code (misclassification), and then re-exported the product under the correct code. The results in Table 5.1.1 also indicate that the U.K. export quantities of plywood exceed production (EX1 > 1) in 43 (71.67%) of the 59 reported years; but because net exports never exceed production (EX2 < 1) the inference can be made that the U.K. is transshipping plywood. While transshipping can lead to discrepancies, it is a routine and legal practice and does not necessarily mean that misclassification, or other discrepant reporting practices occurred. Therefore, the Brazil's unilateral reported trade quantity and production statistics for wood chips and particles is the only country/product that contained a discrepant reporting year. Interestingly, Brazil is also the country with one of the lowest average Control of Corruption, Rule of Law, and Government Effectiveness based on the World Governance Indicators. [24]

5.2 Bilateral Trade Analysis

The purpose of the bilateral trade flow analysis is to identify discrepancies in the reported imports and export quantities for the trade flow from Canada to the U.S, Brazil to the U.S, China to the U.S, the U.S. to China, the U.S. to Japan, and the U.S. to the U.K.

5.2.1 Descriptive Statistics

The following section reports the descriptive statistics for the import and export series for each forestry product; industrial roundwood, sawnwood, plywood, and wood chips and particles, for each bilateral trade flow. Table 5.2 reports the import and export quantities sum,

[24] Figure 2.2.

mean, standard deviation, variance, minimum and maximum of each forestry product for each bilateral trade flow pair in thousands. The descriptive statistics can reveal information about the import and export statistics of each bilateral trade flow; for example, if the summary statistics for imports and exports do not match, discrepancies can be expected.

United States as Importer

Table 5.2 implies that the trade flow from Canada to the U.S. had the greatest number of traded forestry products whenever the U.S. is the importing country. Sawnwood is the top traded product for the trade flows from Canada to the U.S. and Brazil to the U.S. while plywood is the top traded product for the trade flow from China to the U.S. The forestry product with the lowest total traded quantity is wood chips and particles for all three trade flows when the U.S. is the importing country. If import statistics are equivalent to the mirrored[25] export statistics it may be expected that the data is not discrepant, but if they are not equivalent then it can be inferred that the reported quantities are different from one another, or discrepant. For the trade flow from Canada to the U.S. sawnwood and plywood reported trade quantities, Canada's reported minimum export quantity is equivalent to the U.S's reported minimum import quantity, and no other bilateral trade quantity descriptive statistics are equal. Table 5.2 indicates that the forestry products traded from Canada to the U.S. will exhibit discrepant reporting years, but sawnwood and plywood may be less discrepant than industrial roundwood and wood chips and particles. For the trade flow from Brazil to the U.S. plywood's reported bilateral trade quantities share the same minimum value, wood chips and particles' bilateral statistics share the same minimum and

[25] "Mirrored statistics" is a common term for trade data, where import statistics reported by a destination country should be an exact reflection of export statistics reported by the origin.

maximum value, and no other bilateral quantity statistics are equivalent. Therefore, Table 5.2 reveals that forestry products traded from Brazil to the U.S. will exhibit discrepancies, but plywood might have less discrepancies, and wood chips and particles the least compared to other traded products. For the trade flow from China to the U.S. plywood's bilateral trade quantities share the same maximum quantity, wood chips and particles share the same minimum quantity, and no other descriptive bilateral statistics are equivalent. Thus Table 5.2 implies that the reported bilateral trade quantities of forestry products for the trade flow from China to the U.S. will contain discrepancies, but plywood and wood chips and particles will potentially exhibit fewer discrepancies than industrial roundwood and sawnwood.

Table 5.2. Descriptive Statistics for Bilateral Trade Flow Reported Quantities (1997-2021).

Trade Flow			Total Quantity		Standard Deviation Quantity		Minimum Quantity		Maximum Quantity	
ORIG.	DEST.	PROD.	M	X	M	X	M	X	M	X
Canada	U.S.	IR	529.29	468.36	16.40	10.41	4.23	3.97	68.72	36.15
		SW	7020.63	6733.40	126.77	109.20	96.83	96.83	487.91	434.95
		PW	144.15	148.17	2.48	2.39	2.19	2.19	10.45	10.03
		WCP	37.35	38.89	2.48	2.14	0.0002	0.0037	10.20	7.84
Brazil	U.S.	IR	10.18	3.18	1.27	0.25	0.0016	0.0010	6.44	1.10
		SW	134.42	180.64	3.25	6.27	0.19	1.54	11.66	29.07
		PW	114.36	113.49	3.99	4.32	0.31	0.31	12.36	12.63
		WCP	20.15	19.27	1.70	1.71	0	0	5.90	5.90
China	U.S.	IR	22.98	0.30	3.02	0.02	0.0042	0	15.34	0.07
		SW	12.95	5.62	0.43	0.20	0.02	0.01	1.35	0.62
		PW	336.67	331.17	9.89	10.01	0.38	0.39	29.00	29.00
		WCP	0.12	0.16	0.01	0.02	0	0	0.04	0.08
U.S.	China	IR	656.50	546.24	23.53	22.16	0.72	0.69	62.55	62.55
		SW	385.12	356.21	11.31	11.44	3.48	1.90	36.48	36.48
		PW	9.46	10.42	0.63	0.62	0.0002	0.01	2.18	2.18
		WCP	13.57	16.22	0.75	1.12	0.0020	0.0066	2.74	5.21
U.S.	Japan	IR	627.66	570.48	19.13	13.20	5.40	4.49	90.79	60.44
		SW	109.58	111.51	3.21	3.41	1.30	0.95	16.73	17.31
		PW	1.50	2.04	0.06	0.08	0.0029	0.02	0.30	0.40
		WCP	564.43	486.86	23.57	13.58	6.31	8.09	89.41	57.15
U.S.	U.K.	IR	17.10	55.96	0.69	2.59	0.01	0.06	2.22	8.54
		SW	35.96	32.87	0.81	0.76	0.53	0.20	4.36	4.36
		PW	4.99	5.20	0.50	0.59	0.0013	0	2.11	2.85
		WCP	0.54	1.16	0.08	0.08	0.0002	0.0001	0.42	0.42

Note: X refers to export quantities reported by the origin country and M refers to import quantities reported by the destination country. Abbreviations for forestry products are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP). Quantities are reported in hundreds of thousands.

United States as Exporter

Table 5.2 implies that the trade flow from the U.S. to the Japan had the greatest number of traded products out of the three trade flows explored when the U.S. is the exporting country. Table 5.2 reveals that industrial roundwood is the top traded products among the trade flows when the U.S. is the exporting country. For the trade flow from U.S. to China and U.S. to Japan plywood is the least traded quantity, and for U.S. to the U.K. wood chips and particles is the least traded quantity. If import statistics are equivalent to the mirrored export statistics it may be expected that the data is not discrepant, but if they are not equivalent then it can be inferred that the reported quantities are different from one another, or discrepant. For the trade flow from the U.S. to China industrial roundwood, sawnwood, and plywood bilateral quantity statistics share the same maximum values, and no other descriptive statistics are equivalent; so, discrepancies are expected in the reported mirrored statistics for this trade flow and wood chips and particles may be the most discrepant. For the trade flow from the U.S. to Japan, no bilateral trade descriptive statistics are the same which may indicate that of the trade flows when the U.S. is the exporter, U.S. to Japan might exhibit the greatest number of discrepant reporting years across the forestry products under evaluation. For the trade flow from the U.S. to the U.K. only wood chips and particles had bilateral descriptive statistics that were equivalent to one another: the maximum reported quantities and standard deviations are the same. From Table 5.2 it may be expected that the forestry products traded from the U.S. to the U.K. do experience discrepancies with wood chips and particles consisting of the least number of discrepant reporting years.

5.2.2. Correlation Analysis

The correlation analysis will examine the correlation coefficients from each import and export pair of trade quantities for industrial roundwood, sawnwood, plywood, and wood chips and particles for each bilateral trade flow pair. Table 5.3 reports the results from the correlation analysis rounded to the nearest 2-decimal place and is followed by a discussion.

Table 5.3. Correlation Coefficients for Imports and Exports of Bilateral Trade Country Pairs.

	Trading Partner Flows					
Forestry Product:	Canada to United States	Brazil to United States	China to United States	United States to China	United States to Japan	United States to the United Kingdom
IR	0.77	0.86	-0.05	0.75	0.65	0.16
SW	0.96	0.48	0.38	0.97	0.99	0.78
PW	0.98	0.99	1.00	0.97	0.42	0.95
WCP	0.95	1.00	0.46	0.93	0.90	0.92

Note: Trading partner flows are listed as the unidirectional flow from the exporting country to the importing country. Abbreviations are the same following Table 5.2. Correlation is measured for years 1997-2021. Abbreviations for forestry products are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP). Quantities are reported in hundreds of thousands.

United States as Importer

Table 5.3 indicates that correlation is overall the highest for traded plywood quantities when the U.S is the importing country and the lowest for industrial roundwood. This is interesting because plywood is the only forestry product across all three trade flows when the U.S. is the importing country that are subject to tariff restrictions[26]. The trade flow from China to the U.S. seems to have the least correlated bilateral import and export reported quantities, and Canada to the U.S. has the highest. The table implies that the correlation coefficient rounded to the

[26] From literature review, it is expected that higher tariffs lead to more discrepant trade data; so, plywood having the highest value for correlation contradicts this expectation.

nearest second digit for Brazil to the U.S. plywood (0.99) and wood chips and particles (1.00), and China to the U.S. plywood (1.00) have almost a perfectly positive relationship with one another which could lead to the expectation that less discrepancies are present for these flows/forestry products in relation to the other reported flows/forestry products when the U.S. is the importing country. The correlation coefficient for China to the U.S. industrial roundwood (-0.05) indicates a weak negative linear relationship between bilateral import and export quantities implying that their reported trade relationship is highly discrepant. [27]

United States as Exporter

When the U.S. is the exporting country, Table 5.3 indicates that correlation is overall the highest for traded quantities of wood chips and particles. Industrial roundwood again seems to have the least correlated bilateral trade quantities compared to sawnwood, plywood, and wood chips and particles. The trade flow from the U.S. to China seems to have the most correlated bilateral import and export reported quantities overall which is interesting because China's WGIs are the lowest among the trade partners when the U.S. is the exporter. Table 5.3 implies that bilateral statistics for sawnwood traded from the U.S. to Japan is almost perfectly positively correlated (0.99) which may indicate this trade flow/forestry product is less discrepant compared to others when the U.S. is the exporter.

[27] It is expected that strong positive linear associations indicate non-discrepant bilateral trade quantities.

5.2.3. GAP Analysis

The GAP analysis analyzes the discrepancies between imports and exports of each bilateral trade flow pair. The GAP analysis identifies what percentage of years from 1997-2021, 25 years total, discrepancies are present between the mirrored trade statistics. Table 5.4 presents the results of the GAP analysis followed by an interpretation of the results.

Table 5.4. GAP Analysis for Forestry Products for Bilateral Trade Pair.

		Forestry Products:				
Trade Flow		IR	SW	WCP	PW	Average
Canada to U.S.	% Of Years (+) Gap	52	24	40	32	37
	% Of Years (-)	28	2	8	16	13.5
	% Of Years Gap	80	44	48	48	55
Brazil to U.S.	% Of Years (+) Gap	68	36	8	24	34
	% Of Years (-)	32	36	0	24	23
	% Of Years Gap	100	72	8	48	57
China to U.S.	% Of Years (+) Gap	96	84	28	32	60
	% Of Years (-)	4	12	36	16	17
	% Of Years Gap	100	96	64	48	77
U.S. to China	% Of Years (+) Gap	92	36	20	8	39
	% Of Years (-)	8	8	28	40	21
	% Of Years Gap	100	44	48	48	60
U.S. to U.K.	% Of Years (+) Gap	2	24	16	40	11.5
	% Of Years (-)	68	20	60	16	41
	% Of Years Gap	88	44	76	56	66
U.S. to Japan	% Of Years (+) Gap	40	8	20	8	19
	% Of Years (-)	48	36	28	44	39
	% Of Years Gap	88	44	48	52	58

Note: Trade flow refers to the origin country (exporting) to the destination country (importing). Positive GAP means that the importer reported a larger quantity and negative GAP means the exporter reported a larger quantity. Abbreviations for forestry products are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP). Quantities are reported in hundreds of thousands.

United States as Importer

When the U.S. is the importing country, the results from Table 5.4 imply that on average the trade flow with the most discrepancies found between reported bilateral forestry product trade quantities is China to the U.S. The trade flow from Canada to U.S. had the least average discrepancies across the four forestry products in the analysis. Table 5.4 indicates that across all three trade flows when the U.S. is the importer, the U.S. is found to overreport or its partner underreport more often than the U.S. underreports or its partner overreports. For all three trade flows when the U.S. is the importing country, industrial roundwood seems to be the product that experiences the greatest number of discrepant reporting years. Figure 5.1 below is a time series plot for import and export quantities of each forestry product for the trade flow from Canada to the U.S and illustrates the GAPs for this bilateral pair.

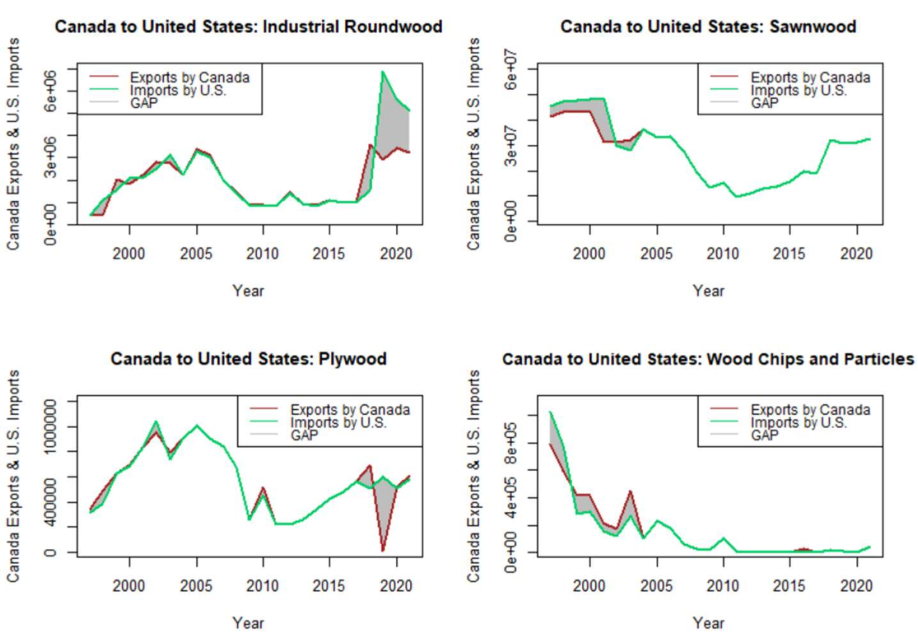

Figure 5.1. Time Series Plots for Canada to the U.S.

United States as Exporter

When the U.S. is the exporting country, Table 5.4 indicates that on average, the trade pair that experienced the greatest number of discrepant reporting years across all four forestry products is the flow from the U.S. to the U.K. The trade flow from the U.S. to Japan had the least number of discrepancies on average across the four traded forestry products. Industrial roundwood is again the traded forestry product with the greatest number of discrepancies for all three trade flows when the U.S. is the exporting country. Table 5.4 reveals that the U.S. is overreporting or its partner is underreporting for the trade flows from U.S. to U.K. and U.S. to Japan. From the U.S. to China is the only trade flow across all trade flows reported in Table 5.4 in which the U.S. underreports or its partner overreports. Figure 5.2 below is a time series plot for import and export quantities of each forestry product for the trade flow from the U.S. to China and illustrates the GAPs for this bilateral pair.

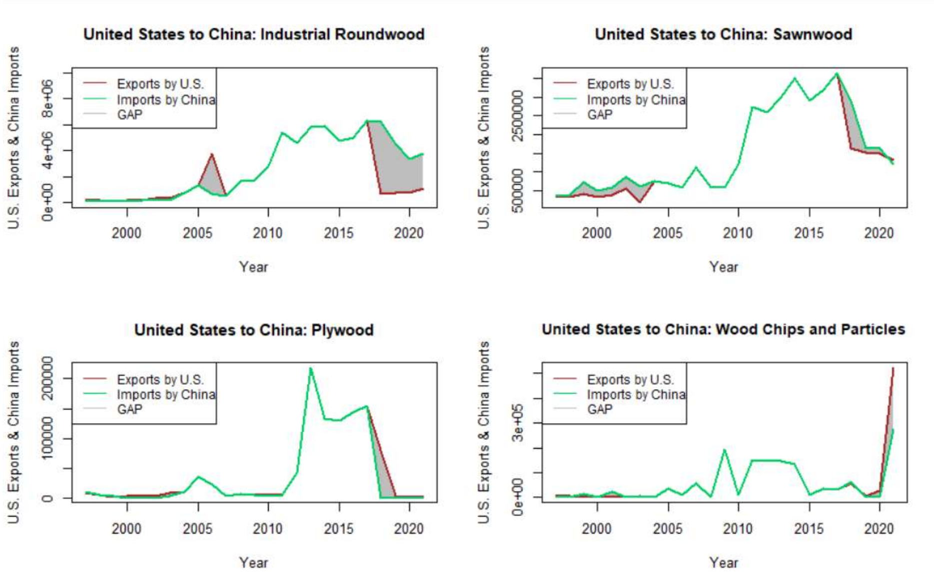

Figure 5.2. Time Series Plot for U.S. to China.

5.2.4. Bilateral Trade Pair Econometric Results

Visual inspections of the bilateral import and export quantities time series plots for the trade flows from Canada to the U.S, Brazil to the U.S, China to the U.S, the U.S. to China, the U.S. to Japan, and the U.S. to the U.K in the Appendix Figures A.2 – A.5 respectively, illustrate potential outliers. The presence of outliers can lead to spurious stationarity and/or cointegration (Franses and Haldrup, 1994). Formal outlier detection was conducted using the SAS ARIMA procedure's OUTLIER statement and confirmed the presence of additive and level shift outliers. In order to cope with the presence of outliers for ADF tests, dummy variables were included into Equations 7-19 demonstrated by Equations 25-27, and ADF were performed utilizing Dickey-Fuller critical values (Franses and Haldrup, 1994):

$$\Delta y = \mathrm{a}_0 + \gamma y_{\mathrm{t}-1} + \mathrm{a}_2\mathrm{t} + \sum_{\mathrm{i}=2}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t}-\mathrm{i}+1} + \sum_{i=0}^{p+1}\sum_{j=1}^{m} w_{ij}^{O} D_{t-1}^{Oj}\, \mathrm{e}_{\mathrm{t}} \tag{25}$$

$$\Delta y = \mathrm{a}_0 + \gamma y_{\mathrm{t}-1} + \sum_{\mathrm{i}=2}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t}-\mathrm{i}+1} + \sum_{i=0}^{p+1}\sum_{j=1}^{m} w_{ij}^{O} D_{t-1}^{Oj} + \mathrm{e}_{\mathrm{t}} \tag{26}$$

$$\Delta y = \gamma y_{\mathrm{t}-1} + \sum_{\mathrm{i}=2}^{\mathrm{p}} \beta_{\mathrm{i}} \Delta y_{\mathrm{t}-\mathrm{i}+1} + \sum_{i=0}^{p+1}\sum_{j=1}^{m} w_{ij}^{O} D_{t-1}^{Oj} + \mathrm{e}_{\mathrm{t}}, \tag{27}$$

where D^{Oj} indicates dummy variables that are equal to 1 at the time an outlier occurs and 0 otherwise. Table 5.5 summarizes the results from outlier testing and ADF tests. Table 5.5 indicates that all bilateral series sets were subject to outliers with the exception of U.S. imports of sawnwood from China. Results from the ADF tests subject to outlier dummy variables is found in Table A.1 in the Appendix.

Table 5.5. Summary of Bilateral Series Sets Data Generating Processes.

				Levels Outlier Detection		First Difference Outlier Detection				
ORIG.	DEST.	PROD.	Reporter	Outlier Type	Year	Outlier Type	Year	ADF Results	Proceeding Consideration	Evidence of Cointegration
Canada	U.S.	IR	ORIG.	Shift	2019	Additive	2019	I(1)	Bounds Cointegration	Yes
			DEST.	Shift	2018			TS		
Canada	U.S.	SW	ORIG.	Shift	2002	Additive	2002	I(1)	Johansen Cointegration	Yes
			DEST.	Shift	2008	Additive	2018	I(1)		
Canada	U.S.	PW	ORIG.	Shift	2009	Additive	2009	I(1)	Johansen Cointegration	Yes
			DEST.	Additive	2013	Additive	2019	I(1)		
Canada	U.S.	WCP	ORIG.	Shift	1999			I(0)	Regression in Levels	
			DEST.	Additive	2003			I(0)		
Brazil	U.S.	IR	ORIG.	Additive	2004			I(0)	Bounds Cointegration	Yes
			DEST.	Additive	2004	Additive	2004	I(1)		
Brazil	U.S.	SW	ORIG.	Additive	2003	Additive	2004	I(1)	Johansen Cointegration	No
			DEST.	Additive	2020	Additive	2021	I(1)		
Brazil	U.S.	PW	ORIG.	Additive	2005	Additive	2018	I(1)	Johansen Cointegration	No
			DEST.	Shift	2018	Additive	2007	I(1)		
Brazil	U.S.	WCP	ORIG.	Additive	2010			I(0)	Regression in Levels	
			DEST.	Additive	2010			I(0)		
China	U.S.	IR	ORIG.	Additive	2004			I(0)	Bounds Cointegration	Yes
			DEST.	Additive	2011	Additive	2012	I(1)		
China	U.S.	SW	ORIG.	Shift	2019			I(0)	Regression in Levels	
			DEST.	NA				I(0)		
China	U.S.	PW	ORIG.	Additive	2010	Additive	2018	I(1)	Johansen Cointegration	Yes
			DEST.	Additive	2010	Additive	2018	I(1)		
China	U.S.	WCP	ORIG.	Additive	1999			I(0)	Regression in Levels	
			DEST.	Additive	1999			I(0)		
U.S.	China	IR	ORIG.	Shift	2011	Additive	2011	I(1)	Johansen Cointegration	No
			DEST.	Additive	2017	Additive	2018	I(1)		
U.S.	China	SW	ORIG.	Shift	2011	Additive	2011	I(1)	Johansen Cointegration	Yes
			DEST.	Shift	2011	Additive	2018	I(1)		

(Table cont'd.)

				Levels Outlier Detection		First Difference Outlier Detection				
ORIG.	DEST.	PROD.	Reporter	Outlier Type	Year	Outlier Type	Year	ADF Results	Proceeding Consideration	Evidence of Cointegration
U.S.	China	PW	ORIG.	Additive	2013			I(0)	Regression in Levels	
			DEST.	Additive	2018			I(0)		
U.S.	China	WCP	ORIG.	Additive	2021	Additive	2021	I(1)	Johansen Cointegration	Yes
			DEST.	Additive	2021	Additive	2021	I(1)		
U.S.	Japan	IR	ORIG.	Additive	2003			I(0)	Bounds Cointegration	Yes
			DEST.	Additive	2009	NA		I(1)		
U.S.	Japan	SW	ORIG.	Additive	1997			I(0)	Regression in Levels	
			DEST.	Additive	1997			I(0)		
U.S.	Japan	PW	ORIG.	Additive	1998	Additive	2013	I(1)	Johansen Cointegration	Yes
			DEST.	Additive	1997	Additive	1998	I(1)		
U.S.	Japan	WCP	ORIG.	Shift	2002			I(0)	Regression in Levels	
			DEST.	Shift	2000			I(0)		
U.S.	U.K.	IR	ORIG.	Shift	2011			I(0)	Regression in Levels	
			DEST.	Additive	2008			I(0)		
U.S.	U.K.	SW	ORIG.	Additive	2007	Additive	22	I(1)	Bounds Cointegration	Yes
			DEST.	Additive	2007			I(0)		
U.S.	U.K.	PW	ORIG.	Additive	1997			I(0)	Regression in Levels	
			DEST.	Additive	1997			I(0)		
U.S.	U.K.	WCP	ORIG.	Additive	2005			I(0)	Regression in Levels	
			DEST.	Additive	2005			I(0)		

Note: Abbreviations for forestry products are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

Table 5.6 below presents the final results from regressions estimated for bilateral import and export quantities that were found to be stationary in levels. Table 5.7 consists of model diagnostic testing and Table 5.8 displays hypothesis testing results performed on the estimated regressions from Table 5.6. Each final estimated model is interpreted in the following subsections for trade flows when the United States is the importing country (Canada to the U.S, Brazil to the U.S, and China to the U.S) when it is the exporting country (U.S. to China, U.S. to Japan, and U.S. to U.K.) respectively. ADF tests revealed that the bilateral import and export quantity statistics of forestry products for the six trade flows analyzed experience different combinations of integration orders: 1) both import and export statistics form a bilateral series pair are stationary in levels, 2) both sides are integrated of the first order, or 3) one side of the bilateral quantity statistic set is integrated of the first order while the other is stationary or trend stationary. As discussed in Chapter 3, in the first case the bilateral series set will be analyzed through estimated regressions in levels; the second case leads to Johansen cointegration testing; and the third case leads to Bounds cointegration testing. Table 5.5 also summarizes the results from Johansen and Bounds cointegration tests.[28]From Chapter 3, discrepancies for Johansen cointegrated bilateral series sets are estimated through VECMs; discrepancies for Bounds cointegrated bilateral series sets are estimated through ARDL ECMs; and non-causality testing is conducted on bilateral series sets that are not cointegrated.

[28] Johansen cointegration results are presented in Table A.2 in the Appendix; and Bounds cointegration results are presented in Table A.4 in the Appendix.

Table 5.6. Estimation Results from Regressions on Bilateral Trade Statistics Estimated in Levels.

ORIG (i):	Canada	Brazil	China	China	U.S.	U.S.	U.S.	U.S.	U.S.	U.S.
DEST (j):	U.S.	U.S.	U.S.	U.S.	China	U.K.	U.K.	U.K.	Japan	Japan
DEP.Variable	$M_{jWCP,t}$*	$M_{jWCP,t}$	$M_{jSW,t}$	$M_{jWCP,t}$	$M_{jPW,t}$	$M_{jIR,t}$	$M_{jPW,t}$	$M_{jWCP,t}$	$M_{jSW,t}$	$M_{jWCP,t}$*
FUNCT.Form	Log-Log	Linear	Linear	Linear	Linear	Log-Log	Log-Log	Linear	Linear	Log-Log
Dummy:	dum16	NA	dum19	NA	dum18	dum10; dum19	dum97	dum05	dum18	NA
$\widehat{\beta_0}$	0.16 (0.87)	4176 (1.20)	10670** (2.49)	316.92 (1.55)	-945.12** (-2.35)	4.85*** (2.80)	5.48*** (6.36)	653.76** (2.44)	9686 (0.70)	-3.95** (-2.22)
$\widehat{\beta_{Xip,t}}$	0.97*** (34.30)	0.99*** (52.34)	0.63 (1.54)	0.24** (2.50)	1.01*** (184.70)	0.50*** (3.36)	0.39*** (4.99)	-0.05 (-0.75)	0.95*** (38.70)	1.27*** (10.10)
$\widehat{\beta_{Mjp,t-1}}$	0.01 (0.27)									
$\widehat{\beta_{Xip, t-1}}$							-0.08 (-1.02)			
$\widehat{\beta_{Mjp, t-2}}$										
$\widehat{\beta_{Xip,t-2}}$							-0.01 (0.10)			
$\widehat{\beta_{Xip,t-3}}$										
$\widehat{\beta_{D1,t}}$	-6.82*** (-33.66)		-71952* (-1.77)		-78899*** (-46.35)	-3.64** (-2.24)	1.36** (2.24)	43418*** (16.75)	143297*** (3.43)	
$\widehat{\beta_{D2,t}}$						-2.67*** (-3.12)				
R^2	0.9955	0.9917	0.2489	0.2137	0.9994	0.5803	0.8639	0.9890	0.9858	0.9612

Note: T-values for coefficient estimations are reported in parentheses: (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance. An asterisk (*) on the dependent variable indicates that heteroskedastic consistent (HC) t-values are reported. $\beta_{Xip,t}$ can be read as Beta on export quantities reported by origin country (i) for forestry product (p). Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

Table 5.7. Model Diagnostic Tests for Regressions on Bilateral Trade Statistics Estimated in Levels.

			Diagnostic Tests:								
ORIG.	DEST.	PROD.	Ramsey Reset	VIF	Cond. Index	White's	Breusch-Pagan	Durbin Watson	Durbin-h	Ljung-Box	Jarque-Bera
Canada	U.S.	WCP	1.40	✓	✓	HCC SE	HCC SE	NA	1.83	1.56	0.34
Brazil	U.S.	WCP	0.03	✓	✓	0.32	0.25	2.12	NA	0.12	458.51***
China	U.S.	SW	0.33	✓	✓	1.97	1.46	1.44*	NA	0.28	3.91
China	U.S.	WCP	2.72*	✓	✓	1.74	1.74	1.67	NA	0.70	38.40***
U.S.	China	PW	0.37	✓	✓	3.00	2.66	1.47	NA	0.62	1.47
U.S.	U.K.	IR	3.08*	✓	✓	2.52	2.44	1.79	NA	0.01	0.44
U.S.	U.K.	PW	0.93	✓	✓	12.00	7.11	2.07	NA	0.73	12.70***
U.S.	U.K.	WCP	1.49	✓	✓	1.57	0.18	2.38	NA	1.57	212.01***
U.S.	Japan	SW	3.11*	✓	✓	0.76	0.21	2.28	NA	1.50	27.85***
U.S.	Japan	WCP	1.94	✓	✓	HCC SE	HCC SE	1.40*	NA	1.95	1.00

Note: ORIG. indicates the origin country (exporter), DEST. indicates the destination country (importer), and PROD. indicates the forestry product. VIF refers to the Variance Inflation Factor and Cond. Index is the condition index: a check mark indicates that the VIF ≤ 5 and Cond. Index ≤ 20 (multicollinearity is not present). Durbin Watson test is used if the estimated model does not include lagged dependent variables, and Durbin-h is used if lagged dependent variables are included in the estimated model. (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance.

United States as Importer[29]

When the U.S. is the importing country of forestry products, bilateral statistics for the following trade flows/products were stationary in levels: Canada to the U.S. wood chips and particles, Brazil to the U.S. wood chips and particles, China to the U.S. sawnwood, and China to the U.S. wood chips and particles. For each bilateral trade pair/forestry product, the simplest models were first estimated as $M_{jp,t} = \beta_0 + \beta_1 X_{ip,t} + e_t$. In Table 5.6 final estimated models for Brazil to the U.S. wood chips and particles and China to the U.S. wood chips and particles trade flow quantities conformed to the simplest model and passed model diagnostic tests reported in Table 5.7 except for the Jarque-Bera test for normally distributed residuals. From Table 5.6 the simple estimated model for Canada to the U.S. wood chips and particles trade flow did not pass model diagnostics. A visual inspection of the plot of imports over exports in Figure A.1 in the Appendix it seemed like a non-linear relationship exists between import quantities reported by the U.S. and export quantities reported by Canada of wood chips and particles, and a large outlier occurs in the year 2016.[30] Therefore, the final regression model for Canada to the U.S. wood chips and particles is estimated in log-log[31] functional form, and includes a dummy variable to accommodate the outlier in observation year 2016. It satisfies all model diagnostics except for failing the null hypothesis of homoscedastic errors and is reported with

[29] Residual plots are provided for one estimated levels regression when the U.S. is the importing country. Residual plots for Canada to the U.S. wood chips and particles is displayed in Figure A.8 in Appendix.
[30] By looking at the raw data for Canada to U.S's wood chips and particles in the year 2016 the U.S. reports an imported quantity of 24 m^3 and Canada reports an exported quantity of 24,281 m^3. This report seems like it could be a mistake of data entry, but it is difficult for practitioners to confirm whether or not it is a mistake of data entry or a reporting mistake. The log-log plot of Canada to U.S. wood chips and particles is in Figure A.7 in the Appendix.
[31] Note that final estimations from Table 5.6 do not all have the same functional form. Some were estimated in linear functional form and others log-log. Throughout this thesis the distinction will be made to clearly indicate to the reader which functional form was used in each specific case.

heteroskedasticity corrected t-values. The relationship between import quantities of sawnwood reported by the U.S. and export quantities reported by China contained a level shift outlier at observation year 2019 which was remedied by the inclusion of a dummy variable equal to 1 for years greater than or equal to 2019. From Table 5.7 China to the U.S. sawnwood passes all model diagnostic checks.

From Table 5.6 the estimated coefficients for $X_{tip,t}$ when the U.S. is the importing country reveal that the U.S. is underreporting, or its partner is overreporting contrary to the inference made from the GAP Analysis in Section 5.2.3. For the wood chips and particles trade from Canada to the U.S, $\widehat{\beta_{x\iota p}}$ indicates that a 1% increase in exports reported by Canada of wood chips and particles is estimated to increase imports reported by the U.S. by about 0.97%, ceteris paribus. In Table 5.6 the $\widehat{\beta_{x\iota p}}$ for wood chips and particles traded from Brazil to the U.S. reveals that when export quantities of wood chips and particles reported by Brazil increase by $1m^3$, import quantities of wood chips and particles reported by the U.S. increase by $0.99m^3$, ceteris paribus. Table 5.6 shows that $\widehat{\beta_{x\iota p}}$ indicates that a $1m^3$increase of sawnwood export quantities reported by China will lead to a $0.64m^3$ increase of sawnwood import quantities reported by the U.S; and that a $1m^3$ increase in wood chips and particles export quantity reported by China will lead to a $0.24m^3$ increase in wood chips and particles reported import quantities by the U.S.

Table 5.6 reports the largest R^2 value when the U.S. is the importing country for Canada to the U.S. wood chips and particles trade where about 99.6% of the variation in imports of wood chips and particles reported by the U.S. can be explained by exports of wood chips ant particles reported by Canada, a lag of imports reported by the U.S, and an intercept term. About 99.2% of the variation in import quantities reported by the U.S. from Brazil of wood chips and particles can

be explained by exports reported by Brazil to the U.S. and an intercept term. And about 21.4% of the variation in import quantities reported by the U.S. from China of wood chips and particles can be explained by exports from China to the U.S. of wood chips and particles and an intercept term. The estimated regression model for China to the U.S. wood chips and particles has the lowest reported R^2 in Table 5.6 as well as the largest number of discrepancies reported in the GAP Analysis in Table 5.4 for the levels bilateral series pairs when the U.S. is the importing country.

The estimated discrepancies or residuals when the U.S. is the importing country take on the forms: for Canada to the U.S. wood chips and particles: $\hat{e}_t = \log(M_{jp,t}) - 0.16 - 0.97\log(X_{ip,t}) - 0.01\log(M_{jp,t-1}) - 6.82D_{16,t}$. For Brazil to the U.S. wood chips and particles: $\hat{e}_t = M_{jp,t} - 4176 - 0.99X_{ip,t}$. China to the U.S. sawnwood: $\hat{e}_t = M_{jp,t} - 10670 - 0.64X_{ip,t} - 71952D_{19,t}$. And for China to the U.S. wood chips and particles: as $\hat{e}_t = M_{jp,t} - 316.92 - 0.24X_{ip,t}$. None of the estimated discrepancies (residuals) conform to the theoretical discrepancy from literature review, and the wood chips and particles trade from Canada to the U.S. insinuates that imports reported the U.S. might succumb to some type of time lag in the reporting process. The estimated coefficients on the export quantities of wood chips and particles for Canada to the U.S. (0.97) and Brazil to the U.S. (0.99) indicate an almost proportional relationship to their mirrored U.S. import quantity statistics, ceteris paribus. While the estimated coefficient on the export quantities of wood chips and particles from China to the U.S. reveals that the U.S. is largely underreporting or China overreporting their trade quantities. Table 5.8 reports hypothesis tests conducted on the relationships between import and export quantities and displays that the null hypothesis of $\beta_0 = 0$ could not be rejected for any of the regressions

estimated in levels when the U.S. is the importing country except for China to the U.S. sawnwood. Table 5.8 also reports that null hypothesis of $\beta_{XIp} = 1$ cannot be rejected for the relationships between Canada to U.S. import and export quantities of wood chips and particles, Brazil to U.S wood chips and particles, and China to the U.S. sawnwood; but is rejected for the relationship between import and export quantities of wood chips and for trade from China to the U.S. Therefore, it can be inferred that when the U.S. is the importing country among the bilateral series sets estimated in levels, a constant term might not statistically significantly affect the relationship between U.S' import quantities and its partner's export quantities except for China to the U.S. sawnwood. As well as, reported import and export quantities are significantly proportional for Canada to the U.S. wood chips and particles, Brazil to the U.S. wood chips and particles trade flows and China to the U.S. sawnwood, but not for China to the U.S. traded quantities of wood chips and particles. Therefore, based on the estimated regressions in levels when the U.S. is the importing country, Canada to the U.S and Brazil to the U.S trade of wood chips and particles is subject to a less discrepant relationship than the reported trade quantities from China to the U.S. of wood chips and particles.

The estimated beta coefficients for export quantities within a bilateral set reveal the estimated relationship between import and export quantities traded of a forestry product between a country trade pair. For Canada to the U.S. wood chips and particles and Brazil to the U.S. wood chips and particles the estimated beta coefficients, 0.97 and 0.99 respectively imply that there are statistically significant positive relationships between the logged imports and logged exports of wood chips and particles traded from Canada to the U.S. and between the import and export quantities of wood chips and particles traded from Brazil to the U.S. The

estimated beta coefficients also imply that there are strong statistical associations (correlations) between imports and exports of wood chips and particles for the two trade flows Canada to the U.S. and Brazil to the U.S. because the estimated beta coefficients are the correlation coefficients of imports and exports multiplied by the standard deviations of imports over the standard deviation of exports. The correlation coefficient of logged import quantities of wood chips and particles reported by the U.S. and logged export quantities of wood chips and particles reported by Canada is about 0.85 which indicates a strong, but not perfect linear association between the two. The correlation coefficient of import and export quantities of wood chips and particles traded from Brazil to the U.S. is about 1 which indicates a perfect positive linear association between the two. From both correlation coefficients it was expected that the aforementioned trade flow pairs were not discrepant or suffered from few discrepancies, and the hypothesis test in Table 5.8 showed that the relationships between imports and exports for wood chips and particles traded from Canada to the U.S. and Brazil to the U.S. are proportional at the 5% significance level which is also expected of discrepancy free trade relationships. For China to the U.S. trade of both sawnwood and wood chips and particles, the estimated beta coefficients on export quantities are lower than those estimated on the export quantities for the Canada to U.S. and Brazil to U.S. wood chips and particles trade flows. The estimated beta coefficients on export quantities for China to the U.S. sawnwood and wood chips and particles trade flows are 0.63 and 0.24 respectively of which only the estimate on wood chips and particles is statistically significant. Both estimated coefficients imply positive relationships between the import and export quantities, but weaker positive linear associations (correlations) than those for Canada to the U.S. and Brazil to the U.S. wood chips and particles trade pairs. The correlation coefficients for

China to the U.S. sawnwood and wood chips and particles trade are 0.38 and 0.46 respectively which does align with the estimated beta coefficients and may lead to the expectation that more discrepancies are present between these two trade flows compared to Canada to the U.S. and Brazil to the U.S. wood chips and particles trade flows.

Neither correlation nor the estimated beta coefficient serve as evidence that export quantities cause import quantities, but due to the nature of trade (i.e. import quantities in a destination country cannot exist without first being exported by the origin country), logically it is expected that export quantities must cause import quantities. If export quantities reported by an origin country do not cause the import quantities reported by a destination country then the reported trade data must be highly discrepant. It can also be expected that exports cause imports on a 1:1 basis, meaning that 1-unit of quantity exported will cause 1-unit quantity being imported and any deviation from this expectation is considered a discrepancy. As discussed previously in this subsection for the trade flows Canada to the U.S. wood chips and particles, Brazil to the U.S. wood chips and particles, and China to the U.S. sawnwood, import and export quantities are proportional at the 5% significance level. Only China to the U.S. wood chips and particles import and export quantities were not proportional at the 5% significance level.

Table 5.8 Hypothesis Tests for Regressions on Bilateral Trade Statistics Estimated in Levels.

			Null Hypothesis:			
ORIG. (X)	DEST. (M)	PROD.	$\beta_0 = 0$	$\beta_x = 1$	Other	Test Statistic
Canada	U.S.	WCP	0.87	1.07	$\beta_0 = \beta_{M,t-1} = 0$	0.66
Brazil	U.S.	WCP	1.20	0.46		

(Table cont'd.)

ORIG. (X)	DEST. (M)	PROD.	Null Hypothesis: $\beta_0 = 0$	$\beta_x = 1$	Other	Test Statistic
China	U.S.	SW	2.49**	0.906		
China	U.S.	WCP	1.55	7.86***		
U.S.	China	PW	2.35**	1.18		
U.S.	Japan	SW	0.69	2.12**		
U.S.	Japan	WCP	2.27**	2.24**		
U.S.	U.K.	IR	2.80***	3.30***		
U.S.	U.K.	PW	6.36***	7.81***	$\beta_{X,t-1} = \beta_{X,t-2} = 0$	0.72
U.S.	U.K.	WCP	16.89***	2.44**		

Note: The null hypotheses $\beta_0 = 0$ and $\beta_X = 1$ are tested using t-tests. Other null hypotheses are conducted using F-tests. (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

United States as Exporter[32]

When the U.S. is the exporting country of forestry products, bilateral statistics for the following trade flows/products were stationary in levels include U.S. to China plywood, U.S. to U.K. industrial roundwood, plywood and wood chips and particles, and U.S. to Japan sawnwood and wood chips and particles. For each of the aforementioned bilateral trade pair/forestry products, the simplest models were first estimated as $M_{jp,t} = \beta_0 + \beta_1 X_{ip,t} + + e_t$. From Table 5.6 none of the final estimated models for the trade pair/forestry products when the U.S. is the exporting country conformed to the simplest model. Table 5.6 reports that U.S. to China plywood trade quantities, U.S. to U.K. wood chips and particles trade quantities, and U.S. to Japan sawnwood trade quantities experienced a linear relationship among their respective bilateral

[32] Residual plots are provided for one estimated levels regression when the U.S. is the exporting country. Residual plots for U.S. to the China plywood is displayed in Figure A.9 in Appendix.

statistics; however, all were subject to outliers. A plot of U.S. exports over China imports of plywood revealed an outlier at observation year 2018; a plot of U.S. exports over U.K. imports of wood chips and particles indicated an outlier at observation year 2005; and an outlier at observation year 2018 was found to affect the relationship between U.S. exports and Japan imports of sawnwood. Linear models did not seem to be appropriate to estimate the relationship between U.S. exports and U.K. imports of industrial roundwood, U.S. exports and U.K. imports of plywood, nor for U.S. exports and Japan imports of wood chips and particles. After visually inspecting plots of the above trade flows/forestry products log of import quantities over log of export quantities, a log-log functional form seemed to be a more appropriate representation of the relationships between their import and export quantities. The relationship between logged export quantities from the U.S. and logged import quantities to the U.K. of industrial roundwood was subject to outliers at observation years 2010 and 2019. That between logged export quantities form the U.S. and logged imports quantities to the U.K. of plywood was subject to an outlier at observation 1997; and the relationship between logged exports from the U.S. to logged imports to Japan of wood chips and particles did not seem to be subject to an outlying observation year. Table 5.7 shows that all of the estimated regressions among bilateral series pairs that were found to be stationary in levels when the U.S. is the exporting country, those for the U.S. to China bilateral trade quantities of plywood, and U.S. to the U.K. bilateral trade quantities of industrial roundwood passed all model diagnostic tests. Those for U.S. to U.K. plywood, U.S. to U.K. wood chips and particles, and U.S. to Japan sawnwood bilateral trade quantities passed all model diagnostic tests except for the Jarque-Bera test for normally distributed residuals. And those for the U.S. to Japan wood chips and particles satisfied all model

diagnostic tests except for failing the null hypothesis of homoskedasticity and is reported with heteroskedasticity corrected standard errors and p-values.

From Table 5.6 the estimated coefficients for $X_{tip,t}$ when the U.S. is the exporting country reveal that for the bilateral trade pair quantities for U.S. to China plywood and U.S. to Japan wood chips and particles, the U.S. is underreporting, or its partner is overreporting. And for the remainder of the bilateral trade pairs/forestry products when the U.S. is the exporting country the U.S. is overreporting or its partner underreporting. Table 5.6 reports $\widehat{\beta_{x\iota p}}$ for each trade flow/product when the U.S. is the exporting country which indicates that, holding all other factors constant, a $1m^3$increase in U.S. export quantities of plywood leads to about a 1.01m^3 increase in China import quantities of plywood; a 1% increase in U.S. exports of industrial roundwood leads to about a 0.5% increase in U.K. imports of industrial roundwood; a 1% increase in U.S. exports of plywood leads to a 0.39% increase in U.K. imports of plywood; a $1m^3$ increase in U.S. exports of wood chips and particles leads to a 0.05m^3 decrease in U.K. imports of wood chips and particles; a $1m^3$ increase in U.S. exports of sawnwood leads to a 0.95m^3 increase in Japan imports of sawnwood; and a 1% increase in U.S. exports of wood chips and particles leads to a 1.27% increase in Japan exports of sawnwood. All of the estimated coefficients for $X_{tip,t}$ are in the expected sign except for U.S. to the U.K. wood chips and particles; and only $\widehat{\beta_{x\iota p}}$ for U.S. to China plywood (1.01) and U.S. to Japan sawnwood (0.95) seem to be close to unity while the other estimated levels regressions do not indicate a proportional relationship between bilaterally reported import and export quantities.

When the U.S. is the exporting country, none of the estimated discrepancies (residuals) conform to the theoretical discrepancy from literature review. Table 5.8 reports hypothesis tests

conducted on the relationships between import and export quantities and displays that the null hypothesis of $\beta_0 = 0$ was rejected at the 5% significance level for every trade flow/product except for U.S. to Japan's bilateral sawnwood trade quantities. For all other trade flow/products when the U.S. is the exporting country, their respective relationships between bilaterally reported import and export quantities of forestry products are affected by some constant intercept. Table 5.8 also reports the results from testing the null hypothesis that $\beta_{XIp} = 1$ was rejected at the 5% significance level for all trade flows/forestry product estimations except for U.S. to China's plywood trade. Therefore, when the U.S. is the exporting country all but one relationship between reported bilateral trade quantities for forestry products is subject to an intercept term's presence, and all but one relationship between exported and imported quantities for a trade flow are not proportional. Therefore, based on the regressions estimated in levels reported in Table 5.6 and the hypothesis tests reported in Table 5.8, it seems like bilateral trade quantity of forestry product statistics might be more discrepant when the U.S. is the exporting country.

Table 5.6 presented the estimated beta coefficients on export quantities for bilateral trade flow/product pairs that were found to be stationary in levels. When the U.S. is the exporting country, estimated beta coefficients were reflective of the correlation coefficients for U.S. to China plywood trade quantities, U.S. to Japan sawnwood trade quantities, logged trade quantities of wood chips and particles from U.S. to Japan, and logged trade quantities of industrial roundwood traded from the U.S. to the U.K. Where the estimated beta coefficients on export quantities in time t and the correlation coefficients from import and export quantities, respectively for U.S. to China plywood are 1.01 and 0.97; those for U.S. to Japan sawnwood are

0.95 and 0.99; for logged wood chips and particles trade quantities from U.S. to Japan are 1.27 and 0.91; and for logged industrial roundwood trade quantities from the U.S. to U.K. are 0.50 and 0.52. The aforementioned estimated beta coefficients and correlation coefficients which indicate positive relationships and positive linear associations between the bilateral trade quantities of each respective bilateral trade pair/product. Contrarily, the U.S. to the U.K. trade flows of plywood and wood chips and particles exhibited smaller estimated beta coefficients on export quantities in time t than the aforementioned trade flows. From the estimated beta coefficients on export quantities for the U.S. to U.K. plywood trade (0.39) and wood chips and particles trade (-0.05) a weak or negative linear association might be expected. However, the correlation coefficients for U.S. to U.K. import and export quantities of plywood and import and export quantities of wood chips and particles are 0.95 and 0.92 respectively which do imply strong positive linear associations between bilateral statistics for these two trade flows. So, although strong positive linear associations were found for the bilateral series sets for plywood and wood chips and particles trade from U.S. to the U.K. neither estimated relationship is proportional, or even close to unity. The results for the trade flows form the U.S. to the U.K. for plywood and wood chips and particles trade reveal that although a strong positive correlation exists between import and export quantities of trade, it does not mean that a proportional estimated relationship exists between the bilateral statistics, nor that export quantities cause import quantities.

5.2.5. VECM Models

This section reports the results from the VECMs estimated with bilateral trade flow forestry products in which both import and export quantities were I(1) from ADF tests and had

Johansen cointegrated mirrored statistics[33]. Results from the Johansen cointegration tests are in the Appendix in Table A.2. Table 5.9 reports the final VECM parameter estimates for the normalized co-integrating vectors, Table 5.10 reports hypothesis tests conducted on the cointegrated vector, Table 5.11 displays the final VECM parameter estimates, and Table 5.12 presents the model diagnostic tests performed on each estimated model.

Table 5.9. Normalized Co-Integrating Vectors Parameter Estimates for Bilateral Trade Pairs.

ORIG (i):	DEST (j):	PROD (p):	Variable:	$\beta_{Mjp,t-1}$	Constant	$\beta_{Xip,t-1}$
Canada	U.S.	SW	$ECT_{ijp,t-1}$	1	-8764605.66	-0.96
Canada	U.S.	PW	$ECT_{ijp,t-1}$	1	-590815.33	-1.05
China	U.S.	PW	$ECT_{ijp,t-1}$	1	-13451.73	-0.99
U.S.	China	SW	$ECT_{ijp,t-1}$	1	-150950.87	-0.94
U.S.	China	WCP	$ECT_{ijp,t-1}$	1	4519.56	-0.86
U.S.	Japan	PW	$ECT_{ijp,t-1}$	1	-463.67	-0.55

Note: ORIG. indicates the origin country (exporter), DEST. indicates the destination country (importer), and PROD. indicates the forestry product. The $ECT_{ijp,t-1}$ can be read as the error correction term estimated for the trade flow from an exporter (origin) to the importer (destination) for the forestry product (p). Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

Table 5.10. Hypothesis Tests on the Co-Integrated Vectors and Alpha for Bilateral Trade Pairs.

			Null Hypothesis:				
ORIG.	DEST.	PROD.	$\boldsymbol{\beta}' = (1, \beta_X, 0)$	$\boldsymbol{\beta}' = (1, -1, 0)$	$\boldsymbol{\beta}' = (1, -1, \beta_0)$	$\alpha_M = 0$	$\alpha_X = 0$
Canada	U.S.	SW	25.60***	23.51***	2.64	12.46***	0.59
Canada	U.S.	PW	45.94***	48.30***	1.54	1.70	24.52***
China	U.S.	PW	1.03	0.32	0.81	1.52	0.90
U.S.	Japan	PW	0.53	16.43***	19.17***	28.55***	0.01

Note: Likelihood ratio test statistics with chi-squared distributions are used to test the above hypotheses where (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, indicate 1% statistical significance. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

[33] Residual plots of each estimated model were examined in unison with model diagnostic checks. Two examples of residual plots sets are provided in the Appendix: (1) for a trade flow/forestry product when the U.S. is the importing country and (2) for a trade flow/forestry product when the U.S. is the exporting country.

Table 5.11. Parameter Estimates for Forestry Products Bilateral Trade Statistics Vector Error Correction Models.

ORIG (i):	DEST (j):	PROD (p):	Variable:	$ECT_{ijp,t-1}$	$D_{1,t}$ (i)	$D_{2,t}$ (j)	$\Delta M_{jp,t-1}$	$\Delta X_{ip,t-1}$	$\Delta M_{jp,t-2}$	$\Delta X_{ip,t-2}$	R^2	F Value
Canada	U.S.	SW	ΔM_{jp}	-0.99***	13836570.00	-8671467.31	0.06	0.05			0.72	6.77***
				(-11.16)			(0.41)	(0.28)				
			ΔX_{ip}	0.24*	14026376.45	1114214.14	-0.15	0.26			0.34	1.39
				(2.08)			(-0.80)	(1.18)				
Canada	U.S.	PW	ΔM_{jp}	-0.17***	-74443.84	-447404.10					0.35	2.54*
				(-3.99)								
			ΔX_{ip}	0.78***	533686.76	-457870.50					0.72	12.11***
				(13.87)								
China	U.S.	PW	ΔM_{jp}	-2.37	-1361006.35						0.24	2.15
				(-1.35)								
			ΔX_{ip}	-1.83	-1626213.07						0.28	2.59*
				(-1.03)								
U.S.	China	SW	ΔM_{jp}	-1.15***	1512346.99	-606287.27	-0.08	-0.09			0.66	5.16***
				(-3.04)			(-0.32)	(-0.32)				
			ΔX_{ip}	-0.38	1612887.47	-1900834.64	-0.02	-0.19			0.73	7.34***
				(-1.00)			(-0.09)	(-0.69)				
U.S.	China	WCP	ΔM_{jp}	-5.47***	69660.28						0.47	6.00***
				(-5.92)								
			ΔX_{ip}	-7.27	98142.43						0.45	5.51***
				(-6.06)								
U.S.	Japan	PW	ΔM_{jp}	-0.89***	1078.67	8904.84					0.84	24.84***
				(-8.70)								
			ΔX_{ip}	0.01	9989.07	-30956.26					0.73	12.91***
				(0.09)								

Note: ORIG. indicates the origin country (exporter), DEST. indicates the destination country (importer), and PROD. indicates the forestry product. M indicates the quantities reported by the importing side and X indicates quantities reported by the exporting side. The estimated coefficients for the ECTs are referred to as α_M for the (DEST.) side and α_X for the (ORIG.) side's reported quantities during interpretation. T-statistics are reported in parenthesis where (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, indicate 1% statistical significance.

Table 5.12. Model Diagnostic Tests for Bilateral Trade Pairs Estimated as Vector Error Correction Models.

				Diagnostic Tests:					
				Bivariate		Univariate			
						Jarque–Bera		ARCH	
ORIG.	DEST.	PROD.	Lags	Portmanteau Lags + 1	Portmanteau Lags + 2	M	X	M	X
Canada	U.S.	SW	2	6.97	9.19	1.28	8.45**	1.31	0.32
Canada	U.S.	PW	1	2.44	3.96	3.13	2.90	0.13	0.16
China	U.S.	PW	1	3.25	7.93	0.21	0.15	0.66	0.66
U.S.	China	SW	2	9.96**	10.69	1.95	3.66	0.01	0.03
U.S.	China	WCP	1	9.53**	11.09	2.01	5.67	0.59	0.43
U.S.	Japan	PW	1	8.09*	11.13	0.61	3.04	0.41	2.74

Note: One asterisk (*) indicates 10% statistical significance, two asterisks (**) indicate 5% statistical significance, and three asterisks (***) indicate 1% statistical significance. ARCH means the autoregressive conditional heteroskedasticity test. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

United States as Importer[34]

When the U.S. is the importing country bilateral trade flows/products in which both side of the mirrored statistics were found to be integrated of the first order and Johansen cointegrated includes Canada to the U.S. sawnwood, Canada to the U.S. plywood, and China to the U.S. plywood. Final estimated models for the applicable bilateral trade quantities of forestry products when the U.S. is the importing country were found to be subject to outliers and included dummy variables to account for outlying observation years: Canada to the U.S. sawnwood trade flow included two additive outliers at years 2002 and 2018, Canada to the U.S. plywood trade flow included two outliers at years 2009 and 2019, and China to the U.S. plywood trade flow included one outlier at year 2018.

Table 5.9 reports the parameter estimates for the normalized co-integrating vector for each estimated VECM when the U.S. is the importing country. The error correction term, or estimated discrepancy takes on the form $\widehat{ECT_{t-1}} = M_{jp} - 8764605.66 - 0.96X_{ip}$ for exports of sawnwood reported by Canada and imports reported by the U.S; $\widehat{ECT_{t-1}} = M_{jp} - 590815.33 - 1.05X_{ip}$ for exports of plywood reported by Canada and imports reported by the U.S; and $\widehat{ECT_{t-1}} = M_{jp} - 13451.73 - 0.99X_{ip}$ for plywood exports reported by China and imports reported by the U.S. Table 5.9 indicates that U.S imports of sawnwood from Canada are overreported or Canada exports of sawnwood are underreported in the long-run; and that the U.S. import quantities are underreported or its partner's export quantities are overreported in the long-run for Canada to the U.S. plywood trade as well as China to the U.S. plywood trade.

[34] Residual plots are provided for one estimated VECM when the U.S. is the importing country. Residual plots for Canada to the U.S. sawnwood is displayed in Figures A.10 and A.11 in Appendix.

Table 5.10 reports results from hypothesis tests on the co-integrating vectors where the null hypothesis that the constant within the long-run equilibrium is equal to zero is rejected at the 5% significance level for Canada to U.S. trade for sawnwood and plywood and cannot be rejected at the 5% significance level for China to the U.S. trade of plywood. The joint null hypothesis that the estimated coefficient for X_{ip} is equal to -1 and the constant term is equal to an unknown constant cannot be rejected for any of the VECMs estimated for trade flows/products when the U.S. is the importing country. And only China to the U.S. bilateral trade statistics for plywood fails to reject the joint null hypothesis that within the long-run equilibrium the constant term is equal to zero and the estimated X_{ip} coefficient is equal to -1. Therefore, at the 5% significance level it is possible to infer that in the long-run, sawnwood and plywood bilateral trade quantities are proportional, but their estimated discrepancy is subject to a constant drift. It is also plausible to infer the 5% significance level that bilateral plywood quantities traded from China to the U.S. are proportional and their estimated discrepancy is not subject to a constant drift.

Table 5.11 displays the final estimated VECM results for all I(1) bilateral series pairs that were found to be Johansen cointegrated. Table 5.12 shows that all of the estimated VECMs when the U.S. is the exporting country satisfy model diagnostic tests except for Canada to the U.S. trade for sawnwood's univariate exporter side model which rejects normally distributed residuals at the 5% significance level. From table 5.11 the speed of adjustment parameters are for sawnwood traded from Canada to the U.S, $\widehat{\alpha_M} \approx$ -0.99 and $\widehat{\alpha_X} \approx 0.24$; so, the U.S. imports have a quicker speed of adjustment after periods of estimated discrepancy comparted to Canada's

exports[35]. For plywood traded from Canada to the U.S, $\widehat{\alpha_M} \approx$ -0.17 and $\widehat{\alpha_X} \approx 0.78$; so, the U.S. imports have a slower speed of adjustment after periods of estimated discrepancy comparted to Canada's exports[36]. For plywood traded from China to the U.S, $\widehat{\alpha_M} \approx$ -2.37 and $\widehat{\alpha_X} \approx -1.83$; so, the U.S. imports have a slower speed of adjustment after periods of estimated discrepancy comparted to China's exports, but both reporters are overcorrecting after deviations from their estimated long-run equilibrium[37]. Table 5.10 reports hypothesis tests on the significance of the speed of adjustment parameter for each VECM and indicates that for trade from Canada to the U.S. of sawnwood only the U.S's reported import quantities significantly respond to periods of estimated discrepancy at the 5% level. For plywood trade from Canada to the U.S. only Canada's reported export quantities significantly adjust after periods of estimated discrepancy at the 5% significance level. And for plywood trade from China to the U.S. neither China's export quantities nor the U.S' import quantities significantly adjust after periods of estimated discrepancy at the 5% significance level.

The results from the significance tests done on the speed of adjustment parameters are interesting because for Canada to the U.S. plywood trade, only Canada significantly adjusts after periods of disequilibrium and plywood is the only forestry products from this sample set that is subject to tariff restrictions. Tariffs are collected by the importing country from the person(s) who import the restricted product which might give an incentive for the U.S. to misreport their statistics vs Canada who has no incentive to misreport. For the sawnwood traded from Canada

[35] To illustrate the cointegrating relationship, a plot of the cointegrating equation for Canada to U.S. sawnwood trade is in the Appendix Figure A.16.
[36] A plot of the cointegrating relationship for Canada to the U.S. plywood trade is in the Appendix Figure A.17.
[37] A plot of the cointegrating relationship for China to the U.S. plywood trade is in the Appendix Figure A.18.

to the U.S, only the U.S. importing side statistics adjust significantly after periods of disequilibrium which could be evidence of correction after some processing lag due to time, distance or transshipping. For example, if Canada sent its sawnwood exports out at the end of the year, the U.S. might not receive its imports until the beginning of next year which may be the reason that the U.S. imports almost completely adjust ($\widehat{\alpha_M} \approx$ -0.99) after periods of disequilibrium. Another interesting observation is that for the plywood trade from China to the U.S. neither bilateral reporter's statistics significantly respond to periods of disequilibrium, and China on average has the weakest control of corruption, government effectiveness and rule of law (Figure 2.2) of all the countries included in this thesis.[38]

United States as Exporter[39]

When the U.S. is the exporting country bilateral trade flows/products in which both side of the bilateral statistics were found to be integrated of the first order and Johansen cointegrated include U.S. to China sawnwood, U.S. to China wood chips and particles, and U.S. to Japan plywood. Final estimated models for the applicable bilateral trade quantities of forestry products when the U.S. is the exporting country were found to be subject to outliers and included dummy variables to account for outlying observation years: U.S. to China sawnwood trade flow included two additive outliers at observation years 2011 and 2018, U.S. to China wood chips and particles trade flow included one outlier at observation year 2021, and U.S. to Japan plywood trade flow included two outliers at observation years 1998 and 2013. Table 5.12 indicates that the

[38] From literature review, it was expected that weaker control of corruption, government effectiveness, and rule of law lead to more discrepant bilateral trade statistics.

[39] Residual plots are provided for one estimated VECM when the U.S. is the exporting country. Residual plots for the U.S. to China sawnwood is displayed in Figures A.12 and A.13 in Appendix.

estimated VECMs in Table 5.11 for U.S. to China sawnwood and wood chips and particles the Portmanteau test for no autocorrelated residuals is rejected. Table 5.12 also shows that the estimated VECM for import and export quantities of plywood from the U.S. to Japan satisfies all model diagnostic tests and fails to reject the Portmanteau test of no autocorrelated residuals.

Table 5.9 reports the parameter estimates for the normalized co-integrating vector for plywood traded from the U.S. to Japan. The error correction term, or estimated discrepancy takes on the form $\widehat{ECT_{t-1}} = M_{jp} - 463.67 - 0.55X_{ip}$ for exports of plywood reported by the U..S. and imports reported by Japan. The error correction term shows that in the long-run the U.S. is overreporting exports of plywood or Japan is underreporting imports of plywood. Table 5.10 reports results from hypothesis tests on the co-integrating vectors where the null hypothesis that the constant within the long-run equilibrium is equal to zero is not rejected at the 5% significance level, and the joint null hypothesis that within the long-run equilibrium the constant term is equal to zero and the estimated X_{ip} coefficient is equal to -1 is rejected at the 5% significance level. Therefore, at the 5% significance level it is possible to infer that the long-run equilibrium between import quantities of plywood reported by Japan an export quantities of plywood reported by the U.S. is not subject to a constant trend; and the bilateral trade quantities of plywood trade from the U.S. to Japan are not proportional in the long-run.[40]

One of the parameters of interest for this analysis is the speed of adjustment parameter, α, or the estimated coefficient on the error correction term. From table 5.11 for plywood traded from the U.S. to Japan, $\widehat{\alpha_M} \approx$ -0.89 and $\widehat{\alpha_X} \approx 0.01$; so, the Japan imports have a quicker speed

[40] The plot of the cointegrating relationship for Japan to the U.S. plywood trade is in the Appendix Figure A.19 which illustrates the lack of a constant trend within the long-run equilibrium.

of adjustment after periods of estimated discrepancy comparted to the U.S' exports. Table 5.10 reports hypothesis tests on the significance of the speed of adjustment parameter for each VECM and indicates that for trade from the U.S. to Japan of plywood only Japan's reported import quantities significantly respond to periods of estimated discrepancy at the 5% level.

The results of the significance testing done on the alpha parameters in Table 5.10 is interesting because for U.S. to Japan trade for plywood only Japan's import statistics adjust from disequilibrium in the following time period. This is similar to the results in Table 5.10 for Canada to the U.S. sawnwood trade in which only the U.S's import statistics adjust back from periods of disequilibrium in the next time period. Both scenarios might be evidence for some sort of processing delay due to a time-distance lag, or transshipment. In other words, exports that are shipped later in the year may not arrive to the destination until the following year. Similarly, for the case of transshipment, goods that are re-exported through intermediary countries prior to reaching their final destination may not arrive until the next year depending on the speed of the transshipment process. The speed of adjustment parameter for Japan imports of plywood from the U.S. (-0.89) shows a slower adjustment than the U.S's imports of sawnwood from Canada (-0.99) which is not surprising, because if the adjustment is due to a delay in processing or shipment, it could be expected that imports from Canada to the U.S. will arrive quicker than those to Japan from the U.S. because Japan is farther from the U.S. than Canada.

5.2.6. ARDL Unidirectional ECM Models

The following bilateral pairs/products were found to be integrated of different orders and cointegrated by means of the bounds tests and then estimated as an ARDL with error correction term. ARDL cointegration tests are found in the Appendix in Table A.3. ARDL LR hypothesis test

table is in Appendix Table A.4. Table 5.13 presents the final long-run ARDL estimates, Table 5.14 reports the results from hypothesis tests, Table 5.15 reports the ARDL error correction model estimates, and Table 5.16 is the results from model diagnostic tests.

Table 5.13. Long-Run ARDL Estimates on Mixed Integration Bilateral Trade Statistic Pairs.

ORIG (i):	DEST (j):	PROD (p):	Funct. Form	Variable:	$\beta_{Mjp,t-1}$	Constant	$\beta_{Xip,t-1}$
Canada	U.S.	IR	Log-Log	$ECT_{ijp,t-1}$	1	-0.71 (0.43)	1.05*** (4.11)
Brazil	U.S.	IR	Linear	$ECT_{ijp,t-1}$	1	8272.14*** (5012.78)	0.65** (2.56)
China	U.S.	IR	Linear	$ECT_{ijp,t-1}$	1	15001.09 (1.52)	11.15* (2.00)
U.S.	Japan	IR	Linear	$ECT_{ijp,t-1}$	1	223613.53 (1.17)	0.83*** (8.21)
U.S.	U.K.	SW	Linear	$ECT_{ijp,t-1}$	1	91386.83*** (3.74)	0.26* (2.01)

Note: The $ECT_{ijp,t-1}$ can be read as the error correction term estimated for the trade flow from an exporter (origin) to the importer (destination) for the forestry product (p). The estimated coefficient on $\beta_{Xip,t-1}$ is referred to as θ in the discussion of results following the tables. T-statistics are reported in parenthesis (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

Table 5.14 Hypothesis Tests on Bilateral Trade Statistic Pairs Estimated as ARDL ECMs.

			Null Hypothesis:					
			Long Run Tests			Granger Non-Causality Tests		
ORIG.	DEST.	PROD.	$\beta_0 = 0$	$\theta = 0$	$\theta = -1$	$\delta_i = 0$	$\alpha = 0$	$\delta_i = \alpha = 0$
Canada	U.S.	IR	-0.43	4.11***	0.18	5.59***	28.84***	15.33***
Brazil	U.S.	IR	5012.78***	2.56**	-1.37	0.99	37.41***	733.47***
China	U.S.	IR	1.52	2.00*	1.82*	1.50	35.41***	314.52***
U.S.	Japan	IR	1.17	8.21***	-1.66	20.42***	21.72***	145.40***
U.S.	U.K.	SW	3.74***	2.01*	-5.71***			

Note: Hypothesis tests 1-5 are conducted with t-statistics while the last hypothesis test is conducted using an F-statistic where (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance.

Table 5.15. ARDL Error Correction Model Estimates for Bilateral Trade Statistic Pairs.

				Short-Run								Long-Run	
ORIG (i):	DEST (j):	PROD (p):	Variable:	ΔX_{ip}	$\Delta X_{ip,t-1}$	$\Delta X_{ip,t-2}$	$\Delta X_{ip,t-3}$	$\Delta M_{jp,t-1}$	$\Delta M_{jp,t-2}$	$\Delta M_{jp,t-3}$	D_t	$ECT_{ijp,t-1}$[41]	R^2
Canada	U.S.	IR	ΔM_{jp}*	0.50**	0.11	0.10					-1.06**	-1.06***	0.82
				(2.17)	(1.60)	(1.59)					(2.26)	(-5.23)	
Brazil	U.S.	IR	ΔM_{jp}	0.19	0.07	-0.01	0.16				596700***	-1.02***	
				(1.19)	(0.51)	(-0.05)	(1.51)				(37.19)	(-37.41)	
China	U.S.	IR	ΔM_{jp}	3.41	-5.06	-2.37		-0.07***	-0.02		1509000***	-0.92***	0.99
				(1.02)	(-1.56)	(-0.74)		(-3.39)	(-1.40)		(63.84)	(-35.41)	
U.S.	Japan	IR	ΔM_{jp}	0.69***	-0.009	-0.02					133300	-1.08***	0.57
				(7.32)	(-0.09)	(-0.23)					(0.14)	(-4.65)	
U.S.	U.K.	SW	ΔM_{jp}	0.42***				0.17*	0.12	0.08	202100***	-1.09***	0.98
				(8.07)				(1.87)	(1.45)	(1.27)	(11.58)	(-8.13)	

Note: The parameter estimate on $ECT_{ijp,t-1}$ is referred to as α in the analysis following the tables. T-statistics are reported in parenthesis where (*) indicates 10% statistical significance, (**) indicate 5% statistical significance, and (***) indicate 1% statistical significance. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

Table 5.16. Diagnostic Tests for Bilateral Trade Statistic Pairs Estimated as ARDL ECMS.

			Diagnostic Tests:			
ORIG.	DEST.	PROD.	Ramsey Reset	Jarque-Bera	Breusch-Pagan	Ljung-Box
Canada	U.S.	IR	2.59*	2.09	12.81**	1.05
Brazil	U.S.	IR	1.33	5.69*	0.98	3.07*
China	U.S.	IR	1.90	12.48***	0.94	2.85*
U.S.	Japan	IR	0.07	13.54***	5.37	0.21
U.S.	U.K.	SW	0.94	0.50	2.66	0.11

Note: One asterisk (*) indicates 10% statistical significance, two asterisks (**) indicate 5% statistical significance, and three asterisks (***) indicate 1% statistical significance. Abbreviations for are industrial roundwood (IR), sawnwood (SW), plywood (PW), and wood chips and particles (WCP).

[41] The parameter estimate on $ECT_{ijp,t-1}$ is the estimated alpha parameter (speed of adjustment).

United States as Importer[42]

When the U.S. is the importing country, bilateral trade flow series sets that were found to have mixed orders of integration and rejected the F-Bounds and T-Bounds tests for no cointegration include Canada to the U.S. industrial roundwood; Brazil to the U.S. industrial roundwood; and China to the U.S. industrial roundwood. Table 5.14 shows that both Canada to the U.S. industrial roundwood and Brazil to the U.S. industrial roundwood reject the null hypothesis that $\theta = 0$; and China to the U.S. industrial roundwood fails to reject the null at the 5% significance level. As discussed in Section 3.2.3, if the null hypothesis $\theta = 0$ cannot be rejected then a long-run relationship does not exist. So, the bilateral import and export quantity statistics for China to the U.S. industrial roundwood trade do not share a long-run relationship and are tested for Granger non-causality in Section 5.2.7.

Table 5.13 reports the long-run ARDL estimates for the error correction term. The relationship for Canada to the U.S. bilateral series set for industrial roundwood trade was best represented as a log-log relationship and its error correction term is estimated as $\widehat{ECT_{t-1}} = \log(M_{jp}) + 0.71 - 1.05\log(X_{ip})$. In the long-run industrial roundwood import quantities reported by the U.S. are potentially overreported or exports quantities reported by Canada are potentially underreported. Table 5.13 reports the estimated error correction term for Brazil to the U.S. industrial roundwood bilateral trade quantities as $\widehat{ECT_{t-1}} = M_{jp} - 8272.14 - 0.65X_{ip}$. Therefore, in the long-run it is possible that imports reported by the U.S. of industrial roundwood

[42] Cumulative sum of recursive residuals (CUMSUM) and CUMSUM of square (CUSUMSQ) plots are examined to test the model's stability. One example of CUMSUM and CUMSQ plots when the U.S. is the importing country is shown in Figure A.14 for Canada to U.S. industrial roundwood. The CUMSUM plot looks good, but CUMSQ points to instability.

are underreported or exports reported by Brazil are overreported. Both error correction terms are subject to a constant term, but Table 5.14 shows that the null hypothesis of $\beta_0 = 0$ cannot be rejected for the final ARDL ECM for Canada to the U.S. industrial roundwood. And the null hypothesis is rejected for Brazil to the U.S. industrial roundwood. The null hypothesis of $\theta = 1$ cannot be rejected for Canada to the U.S. industrial roundwood nor Brazil to the U.S. industrial roundwood. Therefore, within the long-run relationship between Canada export and U.S. import quantities of industrial roundwood is not significantly affected by a constant term and is proportional at the 5% significance level. Within the long-run relationship between Brazil export and U.S. import quantities of industrial roundwood, the quantity statistics are significantly proportional at the 5% level but is subject to a significant constant drift. So that Canada to U.S. industrial roundwood trade seems to be less discrepant than Brazil to the U.S. industrial roundwood trade in the long-run.

Table 5.15 shows the ARDL ECM estimation where the parameter of interest is the alpha parameter or estimated coefficient on the error correction term. For Canada to the U.S. industrial roundwood $\alpha = -1.06$, and for Brazil to the U.S. industrial roundwood $\alpha = -1.02$. Both alpha parameters for ARDL ECMs when the U.S. is the importing country are significant at the 5% level and can be interpreted as the speed of adjustment. When M>X+C or M-X-C=1 within the error correction term U.S. import quantities of industrial roundwood from Canada decrease in the next time period to adjust back to long-run equilibrium or a non-discrepant state. U.S. reported import quantities of industrial roundwood from Brazil will decrease by in the next time period following years of deviations from its long-run equilibrium. Both ARDL ECMs when the U.S. is the exporting country demonstrate overcorrection by U.S. imports when deviations from respective long-run

equilibriums occur with U.S. imports from Canada experiencing the quicker speed of adjustment. Both speed of adjustment parameters when the U.S. is the importing country might be evidence of a processing or shipment delay occurring with U.S from Canada's speed of adjustment parameter being quicker than U.S. from Brazil's. If the speed of adjustment parameter is correcting for a processing or shipment delay, it makes sense that U.S. imports from Canada adjust quicker after disequilibrium than U.S. imports from Brazil because Brazil is farther from the U.S. than Canada. Table 5.16 reports model diagnostic checks where the estimated ARDL ECM for Canada to the U.S. industrial roundwood passes all model diagnostics except for the null of homoskedasticity; so, heteroskedasticity consistent t-values are reported in Table 5.15. The estimated ARDL ECM for industrial roundwood trade from Brazil to the U.S. satisfies all model diagnostic conditions.

Table 5.14 reports hypothesis tests for Granger non-causality which can be interpreted as in the short run Canada exports Granger cause U.S. imports of industrial roundwood, Canada exports also Granger cause U.S. imports of industrial roundwood in the short-run, and Canada exports strongly Granger cause U.S. imports of industrial roundwood. For the trade flow from Brazil to the U.S. for industrial roundwood, Brazil export quantities of industrial roundwood do not Granger cause U.S. import quantities of industrial roundwood in the short run but do Granger cause U.S. import quantities of industrial roundwood in the long-run, and strongly Granger cause U.S. import quantities of industrial roundwood.

United States as Exporter[43]

When the U.S. is the exporting country, bilateral trade flow series sets that were found to have mixed orders of integration and rejected the F-Bounds and T-Bounds tests for no cointegration include the U.S. to Japan industrial roundwood and the U.S. to the U.K. sawnwood. Table 5.14 shows that the U.S. to Japan industrial roundwood rejects the null hypothesis that $\theta = 0$; and the U.S. to the U.K. sawnwood fails to reject the null at the 5% significance level. As discussed in Section 3.2.3, if the null hypothesis $\theta = 0$ cannot be rejected then a long-run relationship does not exist. So, the bilateral import and export quantity statistics for the U.S. to the U.K. sawnwood trade do not share a long-run relationship and are tested for Granger non-causality in Section 5.2.7.

Table 5.13 reports the estimated error correction term for Brazil to the U.S. industrial roundwood bilateral trade quantities as $\widehat{ECT_{t-1}} = M_{jp} - 223613.53 - 0.83X_{ip}$. Therefore, in the long-run imports reported by Japan of industrial roundwood is underreported or exports reported by U.S. are overreported. The error correction term is subject to a constant term, but Table 5.14 shows that the null hypothesis of $\beta_0 = 0$ cannot be rejected for the final ARDL ECM for the U.S. to Japan industrial roundwood. The null hypothesis of $\theta = -1$ cannot be rejected for the U.S. to the Japan industrial roundwood. Therefore, it can be inferred that the long-run relationship between U.S. export and U.K. is not subject to a constant trend and is proportional.

Table 5.15 shows the ARDL ECM estimation where the parameter of interest is alpha, the estimated coefficient on the error correction term. For the U.S. to Japan industrial roundwood

[43] Cumulative sum of recursive residuals (CUMSUM) and CUMSUM of square (CUSUMSQ) plots are examined to test the model's stability. One example of CUMSUM and CUMSQ plots when the U.S. is the exporting country is shown in Figure A.14 for U.S. to Japan industrial roundwood. Both plots point to model stability.

$\alpha = -1.08$, is significant at the 5% level, and can be interpreted as the speed of adjustment. Japan reported import quantities of industrial roundwood from the U.S. will decrease and overcorrect in the next time period following years of deviations from its long-run equilibrium. The estimated ARDL ECM for industrial roundwood trade from the U.S. to the Japan satisfies all model diagnostic conditions except for the Jarque-Bera test for normally distributed residuals. Table 5.14 reports hypothesis tests for Granger non-causality which can be interpreted as in the short run U.S. export quantities of industrial roundwood Granger cause import quantities of industrial roundwood, Granger cause import quantities in the long-run, and strongly Granger cause import quantities.

5.2.7 Granger Non-Causality Tests

The Granger non-causality tests were performed on bilateral series sets that were found to not be cointegrated. The tests can be performed on series without checking for cointegration first, but for the purpose of this thesis all I(1) series sets were checked for Johansen cointegration, and series sets that exhibited different orders of integration were checked for Bounds cointegration. The Granger non-causality test results can be found in Table 5.17 are briefly discussed as follows.

Table 5.17 implies that export quantities of sawnwood reported by Brazil do not Granger cause imports reported by the U.S. of sawnwood quantities. It was also determined that imports reported by the U.S. of sawnwood from Brazil do not Granger cause exports. These results are not a surprise because about 76% of the years from 1997-2021 were discrepant.[44]

[44] See GAP Analysis on Pg. 55.

Table 5.17. Granger Non-Causality Tests for Bilateral Trade Statistics.

				Null Hypothesis: Exports do not Granger cause Imports		Reverse Null Hypothesis: Imports do not Granger cause Exports	
ORIG:	DEST:	PROD:	(k + dmax)	MWALD Statistic	Conclusion	MWALD Statistic	Conclusion
Brazil	U.S.	SW	(1+1) =2	1.35 (0.245)	Yes	0.02 (0.892)	Yes
Brazil	U.S.	PW	(2+1)=3	1.50 (0.471)	Yes	1.70 (0.427)	Yes
China	U.S.	PW	(1+1)=2	2.20 (0.138)	Yes	1.69 (0.194)	Yes
U.S.	China	IR	(1+1)=2	1.34 (0.247)	Yes	0.51 (0.475)	Yes
U.S.	China	SW	(1+1)=2	5.20** (0.025)	No	0.24 (0.626)	Yes
U.S.	China	WCP	(1+1)=2	5.93** (0.015)	No	12.09*** (0.001)	No

Note: ORIG. refers to the origin country (exporter), DEST. refers to the destination country (importer), and PROD. refers to the forestry product. k is the optimal lag order and dmax is the maximum integration order. MWALD refers to the modified Wald statistic. P-values for the MWALD statistic are reported in parenthesis where (*) indicates 10% statistical significance, (**) indicates 5% statistical significance, (***) indicates 1% statistical significance.

From Table 5.17 it can also be inferred that exports of plywood reported by Brazil do not Granger cause imports reported by the U.S, and imports reported by the U.S. do not Granger cause exports. About 48% of the years form 1997-2021 were reported as discrepant which is less than those for plywood between this trade pair but is still a substantial amount. Table 5.17 indicates that exports of plywood reported by China do not Granger cause imports reported by the U.S, and imports reported by the U.S. do not Granger cause exports reported by China. From the GAP analysis, 48% of the years form 1997-2021 were reported as discrepant which is less than those for plywood between this trade pair but is still a substantial amount. Table 5.17 reveals that exports reported by the U.S. do not Granger cause imports reported by China of

industrial roundwood quantities, and imports reported by China do not granger cause exports. This result is not surprising because 100% of the reported years from 1997-2021 for industrial roundwood trade between from the U.S. were discrepant.

Table 5.17 indicates that the only trade flow/products that reject the null hypothesis of Granger non-causality is U.S. to China sawnwood and wood chips and particles. Where U.S. export quantities of sawnwood Granger cause China imports of sawnwood; but China import quantities of sawnwood do not Granger cause U.S. export quantities of sawnwood. Table 5.17 shows that U.S. export quantities of wood chips and particles do Granger cause China import quantities of wood chips and particles; and China import quantities of wood chips and particles do Granger cause U.S. export quantities of wood chips and particles. Because sawnwood's bilateral trade quantities only experienced unidirectional Granger causality it might be expected that wood chips and particle's bilateral trade quantities are less discrepant than sawnwood; however, sawnwood only experiences discrepancies 44% out of 25 years and wood chips and particles is slightly higher at 48% out of 25 years. Further, 48% of the reported observation years were found to be discrepant for wood chips and particles which is interesting because, 48% is the same number of discrepancies found in the trade flows from Brazil to the U.S. of plywood and China to the U.S. of plywood both of which were not found to experience Granger causation between the reported export and import quantities. Therefore, Granger causality might not be a reliable indication of how discrepant a bilateral trade quantity is. Another explanation could be that the non-causality test using AVARs might not have enough power in small sample sizes of 50 observations or less (Zapata and Rambaldi, 1997).

5.2.8. Summary of Bilateral Trade Partners and Forestry Products

The results in this Chapter revealed that discrepancies do exist for every forestry product and trade flow explored. ADF and cointegration tests implied that each unique time series bilateral trade quantity pair exhibits unique data generating processes. Therefore, the results from the time series analysis indicated that estimating the discrepant relationship between bilateral import and export quantities of forestry products does not follow one clear cut method and that discrepancies can be estimated through a variety of models including simple linear regressions (SLR), multiple linear regressions (MLR), Vector Error Correction Models (VECM), Autoregressive Distributed Lag Error Correction Models (ARDL ECM), and Augmented Vector Autoregression Models (AVAR)s. Through hypothesis testing, whether the estimated relationship between import and export statistics are proportional and if discrepancies could be attributed by a constant factor were explored which provides more insight to the contribution of each side of a trade flow to discrepancies. By utilizing error correction models, when appropriate based on the DGP of the bilateral series pair, it was possible to explore whether the import and/or export side's statistics adjust after periods of discrepant reporting. Table 5.18 is a summary of the bilateral forestry product trade data discrepancy analysis conducted in this chapter. The remainder of this summary references factors from literature review that were previously found to affect the size and quantity of discrepant reporting years within a bilateral trade flow.[45]

[45] Figures 2.1, 2.2, 2.3, and 2.4 in Section 2.3.

Table 5.18. Summary of Bilateral Trade Data Discrepancy Analysis.

ORIG.	DEST.	PROD.	ADF Results	Cointegration: Test Type	Result	Final Estimated Model Type	Funct. Form	H_0: $\beta_0 = 0$	$\beta_X = \beta_M$	$\beta_{MX,i} = 0$	$\beta_{XM,i} = 0$	% of GAP Years
Canada	U.S.	IR	I(1) ; TS	Bounds	Yes	ARDL ECM (3,1)	Log-Log	No	Yes			80
Canada	U.S.	SW	I(1) ;I(1)	Johansen	Yes	VECM(2)		No	Yes			44
Canada	U.S.	PW	I(1); I(1)	Johansen	Yes	VECM(1)		No	Yes			48
Canada	U.S.	WCP	I(0); I(0)			MLR	Log-Log	Yes	Yes			48
Brazil	U.S.	IR	I(0); I(1)	Bounds	Yes	ARDL ECM (4,1)	Linear	No	Yes			100
Brazil	U.S.	SW	I(1); I(1)	Johansen	No	AVAR(2)				Yes	Yes	72
Brazil	U.S.	PW	I(1); I(1)	Johansen	No	AVAR(3)				Yes	Yes	48
Brazil	U.S.	WCP	I(0); I(0)			SLR	Linear	Yes	Yes			8
China	U.S.	IR	I(0);I(1)	Bounds	Yes	ARDL ECM (3,3)	Linear	Yes	Yes			100
China	U.S.	SW	I(0); I(0)	CB	No							96
China	U.S.	PW	I(1); I(1)	Johansen	Yes	AVAR(2)				Yes	Yes	48
China	U.S.	WCP	I(0); I(0)			SLR	Linear	Yes	No			64
U.S.	China	IR	I(1); I(1)	Johansen	No	AVAR(2)				Yes	Yes	100
U.S.	China	SW	I(1); I(1)	Johansen	Yes[46]	AVAR(2)				No	Yes	44
U.S.	China	PW	I(0); I(0)			MLR		No	Yes			48

(Table cont'd.)

[46] Estimating a VECM resulted in autocorrelated residuals based on Portmanteau tests. AVAR did not suffer from autocorrelated residuals.

				Cointegration:		Final Estimated Model		H_0:				
ORIG.	DEST.	PROD.	ADF Results	Test Type	Result	Type	Funct. Form	$\beta_0 = 0$	$\beta_X = \beta_M$	$\beta_{MX,i} = 0$	$\beta_{XM,i} = 0$	% of GAP Years
U.S.	China	WCP	I(1); I(1)	Johansen	Yes[47]	AVAR(2)				No	No	48
U.S.	Japan	IR	I(0); I(1)	Johansen	Yes	ARDL ECM (1,3)	Linear	Yes	Yes			88
U.S.	Japan	SW	I(0); I(0)			MLR		Yes	No			44
U.S.	Japan	PW	I(1); I(1)	Johansen	Yes	VECM(1)		Yes	No			52
U.S.	Japan	WCP	I(0); I(0)			SLR		No	No			48
U.S.	U.K.	IR	I(0); I(0)			MLR		No	No			88
U.S.	U.K.	SW	I(1); I(0)	Bounds	Yes	ARDL ECM (4,1)	Linear	No	Yes			44
U.S.	U.K.	PW	I(0); I(0)			MLR	Log-Log	No	No			56
U.S.	U.K.	WCP	I(0); I(0)			MLR	Linear	No	No			76

[47] Estimating a VECM resulted in autocorrelated residuals based on Portmanteau tests. AVAR did not suffer from autocorrelated residuals.

United States as Importer[48]

For the trade flow from Canada to the U.S. industrial roundwood bilateral statistics had the most discrepant reporting years out of the forestry products analyzed; and was also found to be the cheapest forestry product based on average unit values. Unit prices reported by the U.S. were about $91.28/$m^3$ and about $91.62/$m^3$ reported by Canada on average each year. Plywood is the only forestry product in this trade relationship that was subject to ad-valorem taxes; and because tariffs are paid for by the importing individual rather than the exporter, U.S. importers could be inclined to misreport statistics to avoid tariff charges. This observation is consistent with the VECM results for plywood traded from Canada to the U.S. of which only Canada's export significantly adjusted back to a long-run proportional relationship after discrepant reporting years. Sawnwood had the most aggregated non-tariff measures (NTMs) at 487 total restrictions and experienced the least amount of discrepancies among the four forestry products based on the GAP analysis which is contrary to what previous literature found. Canada consistently had a higher rating than the U.S. for control of corruption, government effectiveness and rule of law and from the GAP analysis it was found that the U.S. reported larger trade quantities than Canada for all forestry products.

Contrary to the results from the U.S. to Canada's forestry trade, among the forestry products traded from Brazil to the U.S, the cheapest product based on average unit price each year experienced the least number of discrepancies: wood chips and particles where the U.S. reports an average unit price of about $99.63/$m^3$ and Brazil reports an average unit price of about $108.61/$m^3$ each year. Plywood again is the only forestry product here that was found to

[48] All references (NTMs, WGIs, etc.) made in this section come from Figures 2.1-2.4.

be subject to tariffs but does not seem to have a relative influence on the amount of discrepancies or the estimated relationship between imports and exports as compared to the other three forestry products not subject to tariff restrictions. Wood chips and particles was subject to the least amount of NTMs and reported the least amount of discrepancies which aligns with expectations from literature review. Among the five HS code revisions that took place in the time period covered by this analysis, on average, wood chips and particles was subject to division into the least amount of codes and had the least number of discrepant reporting years. Brazil is consistently rated much weaker than the U.S. for control of corruption, government effectiveness and rule of law. There does not seem to be a clear association between World Governance Indicators (WGIs) and discrepant reporting because although Brazil is weaker in its control of corruption, government effectiveness and rule of law, the U.S. still seems to overreport on average for all forestry products or Brazil underreports. Whereas for the trade flow from Canada to the U.S, the U.S. had weaker control and was found to still be the source of overreporting or its partner underreporting.

For the trade flow from China to the U.S, plywood and wood chips and particles had the least number of reported discrepancy years at 64% each and were also the cheaper forestry products which is the same observation made for the trade flow from Brazil to the U.S.[49] Again, plywood is the only forestry product of the four subject to tariff restrictions, and had one of the lower discrepancy percentages compared to the other forestry products which is contrary to expectations based on literature review. Wood chips and particles is subject to the least amount

[49] Plywood's average unit price was \$494.86/m^3 reported by the U.S. and about \$487.18/m^3 reported by China and wood chips and particles had average unit prices of about \$50.48/m^3 reported by the U.S. and \$64.64/m^3 reported by China.

of NTMs, and had one of the lowest discrepancy reporting percentages, but plywood was subject to the second highest number of NTMs and had the same percentage of discrepant reporting years; so, NTMs might not be a valid explanation for discrepant reporting for this trade pair flow. The two lowest reported discrepancy percentages, wood chips and particles and plywood are divided on average into 2 and 6 HS codes each year respectively; so, the average number of harmonized system (HS) codes that each forestry product is subject to for reporting purposes might be a contender in explaining the differences in discrepancy percentages among these traded forestry products.[50] The WGIs determine that on average, China has a weaker control over corruption than the U.S. and its rule of law and government effectiveness indicators are substantially lower than those of the U.S. Similarly, to the trade flow from Brazil to the U.S, for the trade flow from China to the U.S on average over all four forestry products, the U.S. overreports or China underreports. It seems to be the case that overall, for trade flows when the U.S. is the importer, the United States is overreporting or the exporter is underreporting.

United States as Exporter[51]

For the trade flow from the U.S. to China, the cheapest (most expensive) forestry product was not as easily identifiable because the variation between average unit price reported by China and the U.S. was quite large compared to trade flows when the U.S. is the importing country. Sawnwood had the least number of discrepant reporting years, and its average unit prices were the closest to one another with about \$396.85/m^3 reported by China and about \$325.48/m^3 reported by the U.S. Other forestry products reported large differences in the average unit price

[50] Averages of HS codes are attainted over 5 revision year. See Table 4.2.

[51] All references (NTMs, WGIs, etc.) made in this section come from Figures 2.1-2.4.

reports per year with industrial roundwood's being the most substantial which reflected its GAP analysis of 100% discrepancy: about $\$224.17/m^3$ reported by China and about $\$872.87/m^3$ reported by the U.S. Plywood and wood chips and particles both had 48% of reported years as discrepant and both experienced average unit price per year where China reported almost double the price of the U.S.[52] For this trade flow pair all four forestry products are subject to tariffs with plywood's being the largest (an ad-valorem tax of 86%) followed by industrial roundwood's (70%), wood chips and particles' (40%), and sawnwood's (14%). Based on literature review, this information seems to translate into the results of the GAP and estimation analysis because the highest GAP percentages are ranked as highest being industrial roundwood, then a tie between plywood and wood chips and particles with the lowest being for sawnwood. In this trade flow sawnwood is also subject to the most amount of total NTMs at 1026 which contradicts expectations from literature review. Further, sawnwood's HS code average per year across different revisions is the largest at 12 codes on average per year which contradicts observations made for trade flows when the U.S. is the importing country. WGIs as discussed in the last subsection are the same, where on average, China had a weaker control over corruption, rule of law, and government effectiveness than the U.S. Contrary to observations made on the trade flow from the China to the U.S., in the trade flow from the U.S. to China, China is overreporting and the U.S. is underreporting. This observation of WGIs does align with the observations made for the trade flow from Canada to the U.S.

[52] For plywood China reports about $\$694.47/m^3$ and the U.S. reports about $\$389.39/m^3$. For wood chips and particles China reports about $\$253.68/m^3$while the U.S. reports about $\$125.85/m^3$.

For the trade flow from the U.S. to Japan, unit price information did not seem very fruitful in relation to understanding the potential underlying causes of GAP. For example, sawnwood had the least number of discrepant reporting years, but the second to largest average reported unit prices \$444.21/m^3 and for export average reported unit price was \$445.94/m^3 per year; and industrial roundwood had the greatest number of discrepant reporting years and the second lowest average unit price where Japan reported an average unit price of about \$224.74/m^3 and U.S. reported about \$196.83/m^3 each year. Industrial roundwood was subject to the largest ad-valorem tax rate at 90% and is the most discrepant data series, and plywood had the next largest tax rate of 60% and experienced the next largest number of discrepant years which adheres to expectations from literature review. Non-tariff measures (NTMs) do not seem to offer much guidance in understanding GAP because industrial roundwood is subject to a total of 8 NTMs and had the largest number of discrepant reporting years; plywood is subject to a total of 22 and had the second largest number of discrepant reporting years; wood chips and particles have a total of 2; and sawnwood had the least number of discrepant reporting years and was subject to 15 NTMs[53]. For this trade flow it seems that the highest number of HS codes, which is for sawnwood with an average of 12 codes per year over different revisions did have the lowest number of discrepancies which is contradictory of the observations made on previous trade flows. Based on the WGI the U.S. has a higher average score for control of corruption, government effectiveness, and rule of law compared to Japan: control of corruption is 1.44 for the U.S. and 1.39 for Japan, for government effectiveness, on average the U.S. score is 1.57 and Japan's is 1.45, and rule of

[53] It is important to note that the NTM data for the trade flow from the U.S. to Japan was not as easily accessible and might contain erroneous errors.

law average for the U.S. is 1.57 compared to Japan's 1.38. On average across the four forestry products, the U.S. seems to overreport and Japan underreport more often than the opposite which aligns with the observations made on Brazil to the U.S. and China to the U.S; but contradicts those made for Canada to the U.S. and the U.S. to China.

For the trade flow from the U.S. to the U.K, industrial roundwood had the cheapest reported average unit prices and the largest number of discrepant reporting years where average reported unit price by the U.K. was about \$379.39/m^3 per year and the U.S. reported \$201.08/m^3 on average. This observation of industrial roundwood is consistent with the same observation made from the Canada to U.S. trade relationship. All four forestry products are subject to Value Added Taxes (VAT) implemented by the U.K. with the largest being for wood chips and particles at 20% followed by sawnwood at 19.20%, then plywood of 17.50% and finally industrial roundwood with 12.4%. The largest ad-valorem tax rate is attributed to sawnwood for 80% which is also the forestry product with the least reported discrepancies which contradicts expectations from literature review. Total NTMs are the most for industrial roundwood with 17 who also has the most reported discrepancies which does align with literature review, but for the other three forestry products NTMs do not shed much light on the relationships.[54] On average, the U.K. consistently scores higher than the U.S. on WGIs: control of corruption for U.K. is 1.83 while it is 1.44 for the U.S, government effectiveness is 1.66 for the U.K. and 1.57 for the U.S, and rule of law scores are 1.7 for the U.K. and 1.57 for the U.S. This is interesting because across the four forestry products the U.S. overreports and the U.K. underreports more often than the alternative

[54] Important to note is that non-tariff measures for the trade flow from the U.S. to the U.K. are not as easily accessible as other country trade flow pairs so there may be some inconsistencies with the number of measures.

and is consistent with observations made for the trade flow from Canada to the U.S. and the U.S. to China.

Overall based on analysis of bilateral trade flows it seems like the trade flow from China to the U.S. was the most discrepant on average, followed by U.S. to U.K, then U.S. to China. The trade flow with the least discrepant data came from Canada to the U.S. followed by Brazil to the U.S. then U.S. to Japan. This seems to mostly align with WGI except for the trade flow between Brazil to the U.S. where Brazil had the lowest average scores from the observed countries but was the second least discrepant trade flow. Two of the most discrepant trade flows also experienced the highest number of NTMs where China to the U.S. had a total of 897, and the U.S. to China had a total of 1882; however, the one of the least discrepant trade flow, Brazil to the U.S. had the third most NTMs of 997. U.S. to China, U.S. to Japan and U.S. to U.K. had the highest tariff rates on forestry products of which U.S. to China and U.S. to U.K. were in the top three most discrepant trade data flows. It seems like it is the case among all trade flows except for U.S. to China, that the United States was found to overreport or its partner underreport more often than the alternative. Industrial roundwood is the forestry product with the most discrepant reported data for every trade flow. The least discrepant data is mixed among the other products with sawnwood being the least discrepant for Canada to the U.S, U.S. to the U.K, and U.S. to Japan; and wood chips and particles being the least discrepant for the flows Brazil to the U.S, and China to the U.S. Plywood is tied as the least discrepant traded forestry product with sawnwood for the trade flow from U.S. to China.

Chapter 6. Conclusion

Based on previous research, covered in Chapter 2, it is reasonable to expect discrepancies in forestry trade data databases exist, as they do in most bilateral trade statistic sets. Discrepancies can be due to unavoidable, measurement, and reporting errors. In the collection of literature reviewed that focuses on the discrepancies present in trade data, few focus on forestry. This thesis aims to help fill the gap of trade data discrepancy analysis targeting the forestry sector, all the while exploring the underlying econometric time-series modeling structure of discrepancies.

This research aimed to use simple mathematical equations to identify discrepancies within forestry trade data for the unilateral and bilateral trade data for the U.S. and its top trade partners of forestry products. Through the Excess Trade (EX1) and Excess Net Trade (EX2) analysis there were no discrepancies present in the unilateral forestry product trade quantity statistics for the six countries explored except for one discrepant reporting year found for Brazil's wood chips and particles trade. While the GAP analysis performed on bilateral forestry product trade statistics for the six trade flows explored revealed discrepancies for each trade flow/product.

One potential reason for this occurrence is that unilateral import, export, and production statistics are more easily balanced for reporting purposes by the reporting country and/or trade data collection and dissemination agencies. For example, in the case of FAOSTAT Forestry Trade Data where reported trade statistics are revised by Joint Forest Sector Questionnaire (JFSQ) practitioners to alleviate any reporting mistakes (discrepancies), it is less time and resource intensive to ensure that import, export, and production data for the U.S. industrial roundwood trade align with one another compared to cross-checking the import statistics reported by the

U.S. with the export statistics reported by Canada for industrial roundwood trade from Canada to the U.S. Goal and linear programming methods to identify discrepancies may provide a more precise indication of discrepancies present in unilateral trade statistics. Because rather than comparing unilateral import, export, and production statistics for a traded product (EX1 and EX2 methods) which may have been artificially corrected, linear and goal programming methods can compare the production of forestry products to the consumption of the raw wood materials that are needed to manufacture them (Kallio and Solberg, 2018; Buongiorno, 2018).

Alongside discrepancy identification, this thesis aimed to apply a time-series econometric approach to estimate the nature of discrepancies within a bilateral trade flow for the forestry products industrial roundwood, sawnwood, plywood, and wood chips and particles. From Augmented Dickey-Fuller tests (ADF) it was revealed that each unique trade flow/forestry product bilateral series set were comprised of different data generating processes. For example, some series sets had bilateral import and export statistics that were both stationary in levels, some were integrated of different orders, and some were both integrated of the first order. Previous trade data discrepancy analyses have not considered whether reported trade data statistics are nonstationary and have utilized regression analyses to explain the GAP (discrepancy) through tariff costs and other proxy variables for barriers to trade. The results from this thesis provide new evidence on the underlying data generating processes of bilaterally reported trade statistics, their discrepancies, and offers more insight on how bilateral trade data can/should be treated for analyses.

From the ADF results, discrepancies could be estimated through simple or multiple linear and log-log regressions, Vector Error Correction Models (VECMs), Autoregressive Distributed Lag

Error Correction Models (ARDL ECMs), and Augmented Vector Autoregression Models (AVARs). The estimated discrepancies were used to analyze the relationship between import and export quantities reported by a bilateral trade partner pair, identify whether a constant or trend term affected reported trade quantities, and whether import and export statistics were proportional. Previous data discrepancy analyses have used the GAP (discrepancy) as a dependent variable for a regression in attempts to explain the source of GAP by tariffs, barriers to trade, and corruption. A limitation to that approach is that most of the dependent variables used are proxy variables that might lack the power to conclusively arrive at a discrepancy's source. Whereas previous studies found that higher tariff rates, more barriers to trade, and higher corruption led to more discrepantly reported data, the subsections within Section 5.2.8 compared the number of discrepancies to the aforementioned explanatory factors and found that these ideas contradicted themselves often. For example, the forestry product with highest tariff rates or the most barriers to trade may have been the most discrepant product for one trade flow, but the least discrepant within another. While this thesis did little to pinpoint which factors attribute to data discrepancies, it presented a new methodology that can be used by subsequent researchers to check the quality of bilateral trade data statistics. It also provided new information on the nature of discrepancies, and the data generating processes of bilateral trade quantity data which leads to new insight of how bilateral trade quantities, and trade quantity data in general should be modeled from a statistical perspective.

Several implications can be made from the trade data discrepancy analysis performed in this thesis. First, as mentioned in the previous paragraph, the import and export quantities reported within a bilateral trade pair do not seem to accept one general modeling criteria. Each

reported import and export quantity exhibited unique data generating processes, some were subject to outlying observations that could have been causes by data entry errors, and all required individual modeling considerations across different trade partners and forestry products. In respect to policy analyses, forecasting for supply and demand, and forecasting future developments in the forestry sector, the results of this research demonstrate that special care should be taken when determining the appropriate modeling structure for bilateral trade flows, and trade data in general especially when utilizing trade data across different trade flows, products, and over time.

Second, an economic and trade policy implication that can be derived from this thesis is that the discrepancies found within bilateral trade flow country/product pairs could be leading to product and revenue loss during the process of trade. For this thesis, it was expected that for non-discrepant trade flows imports are equal to exports as a proportional relationship, without being subject to any constant trends (intercepts). However, the results from this thesis indicated that import and export statistics where not always proportional, and sometimes affected by some sort of constant trend. The absence of proportionality between reported import and export quantities could serve as an indication of lost revenue for one or both sides of a bilateral trade flow. For example, the wood chips and particles trade from Brazil to the U.S. was estimated as a simple linear regression, and results were as expected for a non-discrepant trade flow. Based on hypothesis testing conducted, it could not be rejected that the relationship between import and export quantities took on the form $M = 1X + e$; so that a 1 m^3 increase in exports of wood chips and particles from Brazil led to a 1 m^3 increase in imports of wood chips and particles reported by the U.S implying that no trade units nor revenue should be loss during the course of trade.

Wood chips and particles trade from China to the U.S. was also estimated as a simple linear regression, but results were not as expected for a non-discrepant trade flow. For the case of wood chips and particles traded from China to the U.S, hypothesis tests could not reject that that the relationship between import and export quantities took on the form $M = 0.24X + e$; so, whenever China reports 1 m^3 of wood chips and particles being exported to the U.S. the U.S. only reports receiving 0.24 m^3. This result indicates that the U.S. is missing about 0.76m^3 of wood chips and particles. The missing products could be a result of discrepancies due to reporting errors, but it could also indicate that China exporters are losing revenue. On average, China reports its unit price of wood chips and particles when exporting to the U.S. as about \$2,348.38/m^3 and the U.S. reports its average unit price of wood chips and particles when importing from China as about \$8,306.00/m^3. Assuming that the cost, insurance, and freight leads to about a 10% difference in reporting values from exporters to importers, it can be inferred that the U.S's average unit price of wood chips and particles less the cost of insurance and transportation is about $\frac{\$8,306.00}{1.10} = \$7550.91$.[55] From the average unit prices, it can be deduced that on average when China is exporting about $1m^3 * \$2,348.38 = \$2,348.38$ of wood chips and particles the U.S. is only importing about $0.24m^3 * \$7550.91 = \$1,812.22$. Therefore, on average, China is losing about $\$2,348.38 - \$1,812.22 = \$536.16/m^3$ of wood chips and particles exported to the U.S. This is a significant economic implication to consider for policy makers and government officials for both the U.S. and China. While China and Brazil both seem to have substantially lower controls over corruption than the U.S, wood chips and particles trade

[55] In Section 2.2 discusses that in previous research, researchers accept about a 6-10% difference in costs between import and export statistics due to the CIF-FOB differences in reporting values.

with only China seems to have issue. One difference in the trade between Brazil to the U.S. and China to the U.S. is that Brazil to the U.S's wood chips and particle trade is subject to more non-tariff measures restrictions to trade than China to the U.S's. Policy makers and trade officials for both the U.S. and China might benefit by requiring more or revised trade regulations for wood chips and particles following the regulations in place for wood chips and particles traded from Brazil to the U.S. The results from the estimation on wood chips and particles traded from China to the U.S. could also be evidence that transshipping is occurring and not being recorded properly. The U.S. and China could both benefit from increased audits and increased training on the reporting practices of customs officials of each respective country's customs agencies.

Although this thesis found that discrepancies existed among all of the observed forestry product trade flows, it does not provide a solution of how to fix these issues. Previous researchers have suggested improving data collection methods with features like GIS or GPS technologies, more auditing and better training for customs officials, or country tax exchange partnerships; but there is still no clear solution for the trade data discrepancy issue. One opportunity to improve trade data quality as well as deter illegal activities during the process of trade could lie in the adoption of a universal labeling system. By adopting universal product labels that are assigned to traded products during their manufacturing or harvesting processes, forestry products and other traded goods can be traced from their origin to their final destination. A digitalized universal label system could more easily track products as they are traded, and could serve as a tool for customs and data reporting agencies to cross-check reported trade quantities and values; or potentially serve as the primary source of reported trade data. Utilizing a universal labeling system might also deter illegal activities by increasing accountability and traceability of products that are "lost"

during the course of trade whether it be due to transshipping, reporting errors, or illegal activities. It would also help alleviate illegally harvested products from entering or departing a country through trade. This suggestion requires agreements across different industry and government players between countries and does not account for differences in conversion factors across countries; but in this age of digitalization, it is a real potential solution for globally mending the trade data reporting process.

Something to consider is that even if discrepancies are “fixed”, or a new and improved reporting system is put in place, historical trade data quantities will still be reflective of discrepancies. And “fixing” historical trade data will prove to be complex, will require many resources and may not be feasible. One alternative for researchers to attain discrepancy free trade data could lie in machine learning techniques. Machine learning could be useful in estimating trade flow values and quantities because it is centered on prediction by uncovering complex patterns within data; and it is also technically easy to utilize with statistical software packages (Mullainathan and Spiess, 2017). Several studies have utilized machine learning to predict trade flow values based on factors like transportation modes, GDP, land area, and common languages, and ship radio signals among others (Circlaeys et al., 2017; Kottou et al., 2020; Stamer, 2022; Gopinath et al., 2020). Therefore, estimating trade flows with machine learning techniques might provide a solution to attain discrepancy free trade statistics instead of researchers having to rely on discrepantly reported data. Future researchers should focus their efforts in exploring methods like machine learning to generate discrepancy free trade data rather than continuing to identify discrepancies within the trade data or attempts to explain them by use of proxy variables.

Other future research might consider Monte Carlo simulation studies utilizing the results found from this thesis. Because the data generating processes for bilateral trade pairs for different forestry products were revealed through the course of this study, the processes can be replicated as larger sample sizes at annual and monthly frequencies. The research methodology conducted in this thesis can then be replicated for the same data generating processes utilizing larger sample sizes to explore whether the results in this thesis hold in larger sample sizes and with higher time frequencies. Specifically, for Granger non-causality testing which has been proven to have less power for smaller sample sizes of 50 or less by Zapata and Rambaldi (1997), a Monte Carlo simulation study utilizing the specific data generating processes uncovered within this thesis for bilateral trade quantities could provide further information for whether or not export quantities reported by an origin country Granger cause import quantities reported by a destination country. In addition to Monte Carlo simulation work, an impulse response analysis would be a viable addition to the trade data discrepancy analysis conducted in this thesis. By conducting an impulse response analysis whether on reported or simulated trade data, future researchers can visually identify the behavior of imports and export quantities in response to shocks. An impulse response analysis can shed light on what triggers one side of a bilateral trade statistic pair (for example, imports) to adjust after shocks are experienced by its partner's reported statistics (exports). It could be expected that bilateral trade pairs with good quality reporting systems would respond quicker to shocks in the system as opposed to bilateral trade pairs with faulty or inconsistent reporting practices.

Further future research might also benefit from a more in depth look into the customs practices used by each country involved in forestry product trade through interviews with

customs officials. Interviews with customs officials could help to shed light on discrepancy results found between different trade partners/products and give a better indication of any improvements that could be implemented to recording and processing operations and how changes might be implemented or perceived.

There is a limitation of this study that is important to consider. First, in order to examine and analyze forestry trade data in a time-series domain, a reasonable number of observations is necessary. In the case of partner flow data from the FAOSTAT Forestry Database, its annual data is available from 1996-2017 which means that even if every year is reported for each country for each forestry product, only 21 observations are available for each. This data was supplemented by bilateral trade data statistics obtained from ITC Trade Map for trade data from 2018-2021. Using two different sources for trade data quantities might have resulted in an unintentional estimation bias. Future researchers who apply this method might benefit from including a dummy variable when the data sources change or performing Chow Tests for structural break to determine if there is a difference between the statistics obtained from two sources.

This thesis aimed to provide a framework for proceeding researchers to apply in order to analyze time-series trade data discrepancies across any product and country. It also aspired to fill in gaps in literature of trade data discrepancy analyses that examine forestry product data. Identifying and analyzing discrepancies in trade data is the first step in working to mend the mechanisms that lead to discrepancies or determining a more efficient method for accessing discrepant free trade statistics

Appendix. Supplementary Tables and Figures

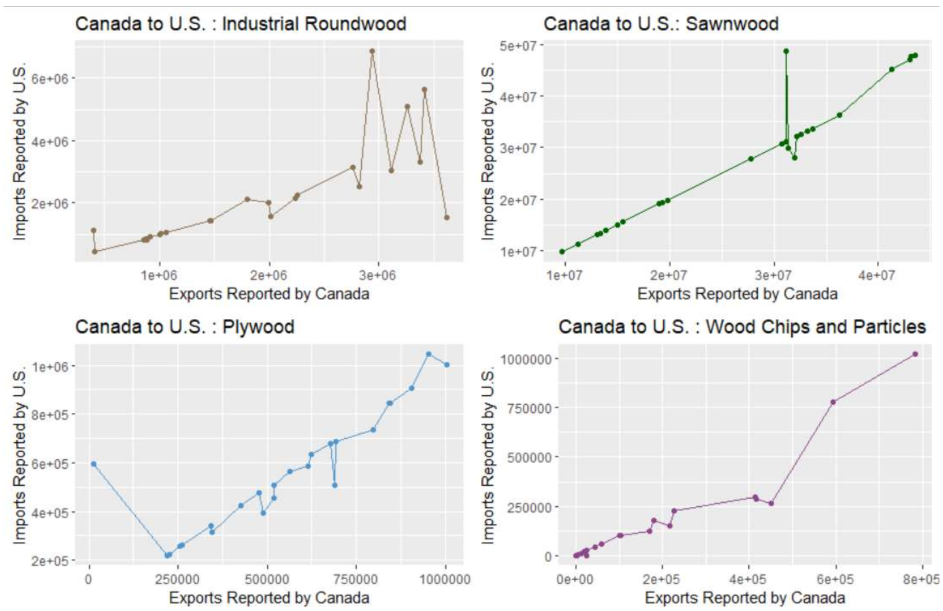

Figure A.1. Canada to U.S. Plots of Imports over Exports.

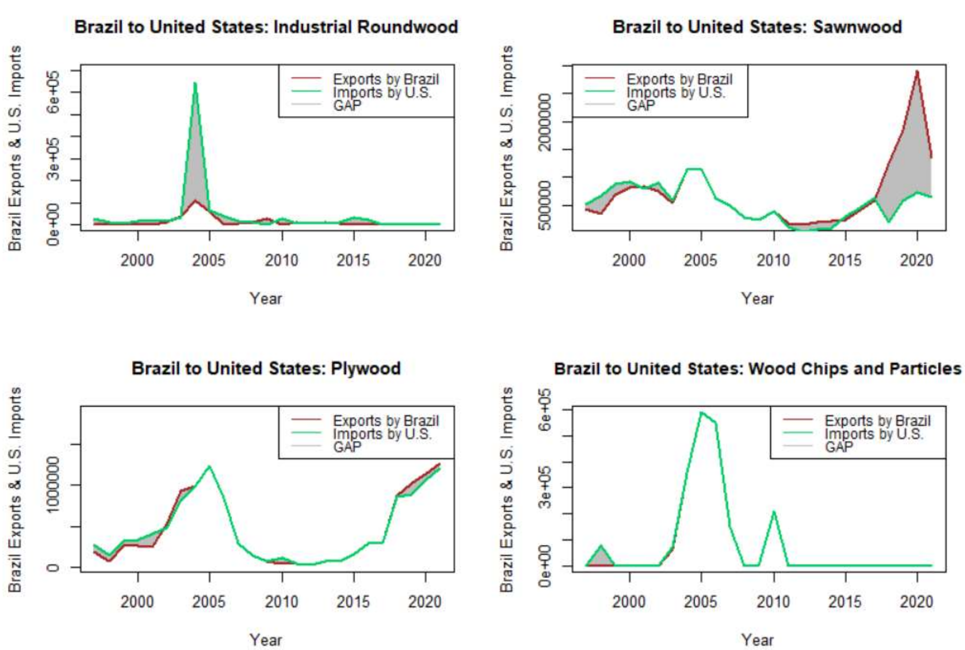

Figure A.2. Brazil to U.S. Time Series Plots.

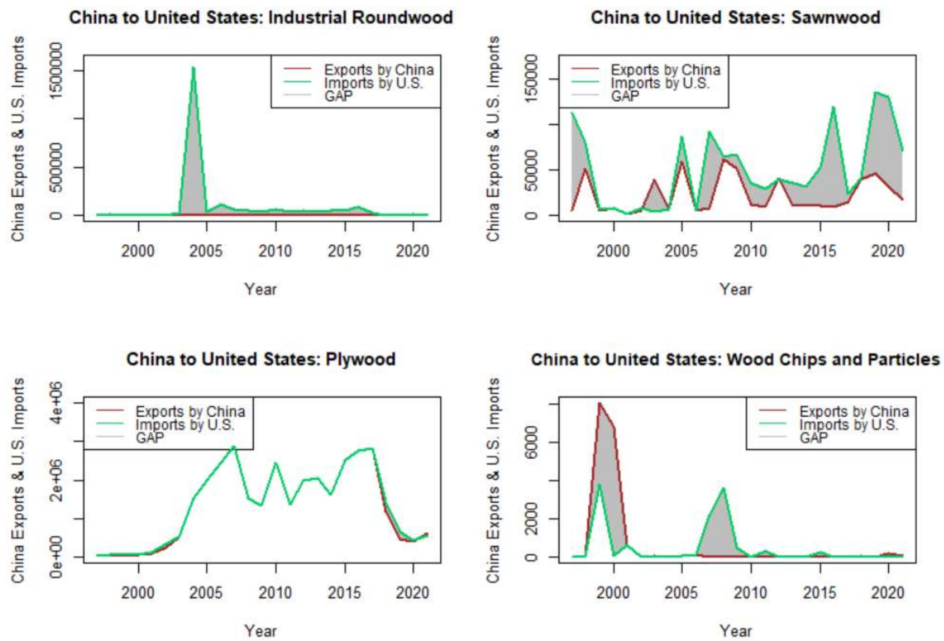

Figure A.3. China to U.S. Time Series Plots.

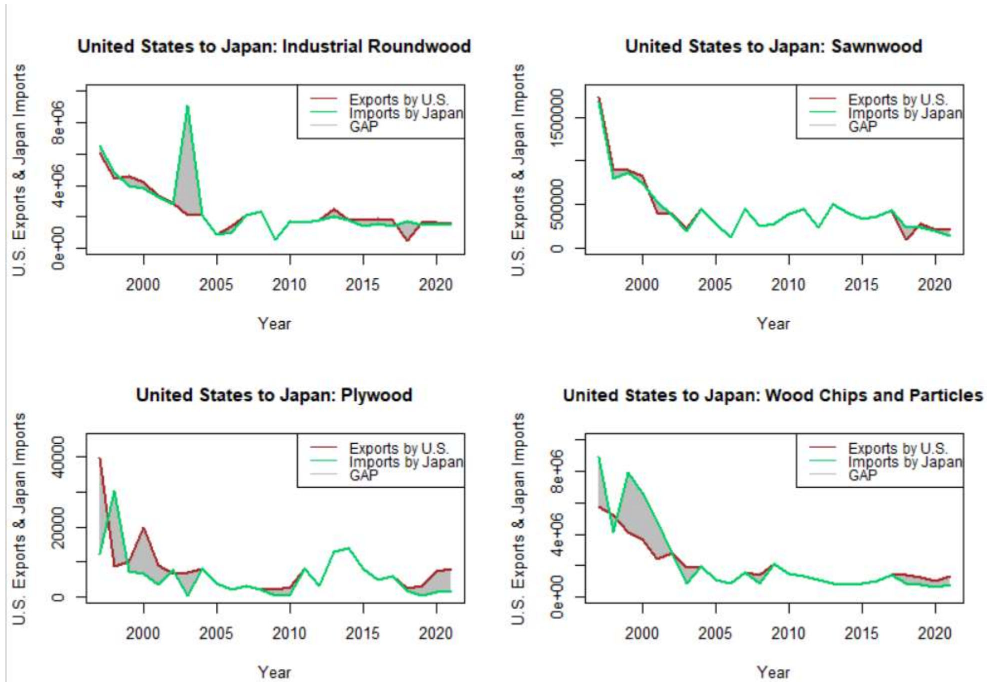

Figure A.4. U.S. to Japan Time Series Plots.

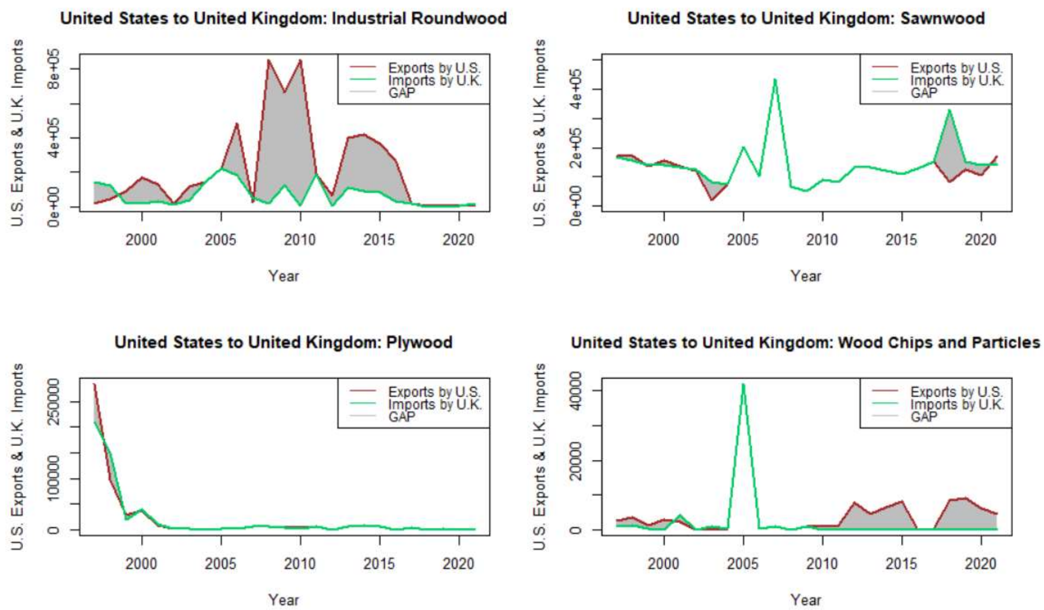

Figure A.5. U.S. to U.K. Time Series Plots.

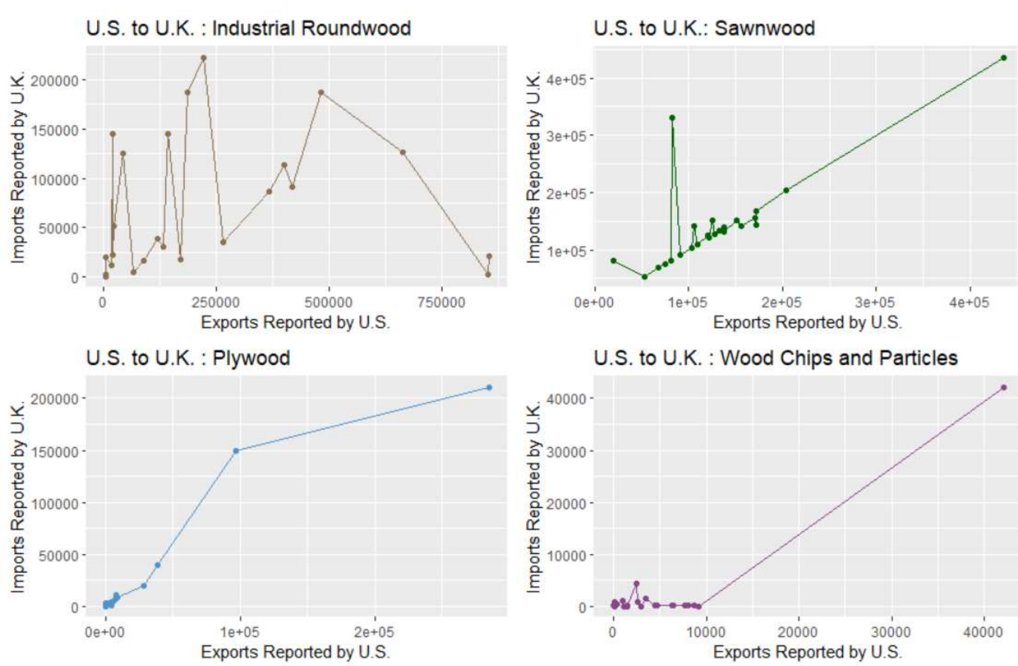

Figure A.6. U.S. to U.K. Plots of Imports over Exports.

Table A.1. Augmented Dickey Fuller Tests.

Critical Values:	1%		-4.38	10.61	-3.75	7.88	-2.66	3.74	3.41		
	5%		-3.60	7.24	-3.00	5.18	-1.95	2.85	2.61		
	10%		-3.24	5.91	-2.62	4.12	-1.60	2.39	2.20		
Null Hypothesis:			$\gamma = 0$	$\gamma = \alpha_2 = 0$	$\gamma = 0$	$\gamma = \alpha_0 = 0$	$\gamma = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\gamma = 0$ [56]	
Trade Flow: Direction	Prod.	Outlier:	Test Statistic: τ_t	Test Statistic: ϕ_3	Test Statistic: τ_μ	Test Statistic: ϕ_1	Test Statistic: τ_t	Test Statistic: $\tau_{\beta t}$	Test Statistic: $\tau_{\alpha\mu}$	Test Statistic: T	Inference:
Canada to U.S: U.S. Imports	IR_M	Sh. 2019 (23)	-2.31	2.83	-1.57	1.24	-0.50				
	ΔIR_M	Ad. (23)	-2.52	3.59	-2.77*	3.83	-2.85***				I(1)
	SW_M	Sh. 2002 (6)	-0.74	1.14	-1.10	0.63	-0.19				
	ΔSW_M	Ad. (6)	-2.56	3.92	-1.62	2.10	-1.85*				I(1)
	PW_M	Sh. (13)	-3.00	5.46	-2.81*	4.06	-0.78				
	ΔPW_M	Ad. (13)	-2.51	3.24	-2.63*	3.65	-2.77***				I(1)
	WCP_M	Sh. (3)	-2.60	4.44	-2.58	3.13	-2.58**				I(0)
Canada to U.S: Canada Exports	IR_X	Sh. (22)	-3.47	7.75***				-3.12**		-3.47*** (0.003)	Trend Stationary
	SW_X	Sh.(12)	-0.91	0.88	-0.48	0.85	-1.33				
	ΔSW_X	Ad. (22)	-2.76	4.06	-2.46	3.02	-2.33**				I(1)
	PW_x	Ad. (23)	-1.70	1.46	-1.33	0.92	-0.23				
	ΔPW_X	Ad. (14)*	-2.30	2.72	-2.41	3.08	-2.54**				I(1)
	WCP_x	Ad. (7)	-2.71	11.7***	-4.97***			0.218			I(0)
(Table cont'd.)											

[56] Test utilizing standard critical t-values (trend stationary).

Critical Values:	1%		-4.38	10.61	-3.75	7.88	-2.66	3.74	3.41		
	5%		-3.60	7.24	-3.00	5.18	-1.95	2.85	2.61		
	10%		-3.24	5.91	-2.62	4.12	-1.60	2.39	2.20		
Null Hypothesis:			$\gamma = 0$	$\gamma = \alpha_2 = 0$	$\gamma = 0$	$\gamma = \alpha_0 = 0$	$\gamma = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\gamma = 0$ [56]	
Trade Flow: Direction	Prod.	Outlier:	Test Statistic: τ_t	Test Statistic: ϕ_3	Test Statistic: τ_μ	Test Statistic: ϕ_1	Test Statistic: τ_t	Test Statistic: $\tau_{\beta t}$	Test Statistic: $\tau_{\alpha\mu}$	Test Statistic: T	Inference:
Brazil to U.S: U.S. Imports	IR_M	Ad. (8)	-3.72**								I(0)
	SW_M	Ad. (7)	-2.10	2.45	-2.20	2.46	-1.35				
	ΔSW_M	Ad. (8)	-2.40	3.90	-2.57	4.21*	-2.71***				I(1)
	PW_M	Ad. (9)	0.73	0.29	0.78	2.97	0.109				
	ΔPW_M	Ad. (21)*	-3.47*	6.45*	-3.56**						I(1)
	WCP_M	Ad. (14)	-3.25*	5.28	-3.15**						I(0)
Brazil to U.S: Brazil Export	IR_X	Ad. (8)	-1.71	1.84	-1.57	1.51	CB				
	ΔIR_X	Ad. (8)	-3.99***								I(1)
	SW_X	Ad. (24)	-1.70	1.66	-1.82	-1.82	0.147				
	ΔSW_X	Ad. (25)	-2.26	2.99	-2.05	2.31	-1.99**				I(1)
	PW_x	Sh. (22)	-2.80	4.11	-2.35	2.80	-1.44				
	ΔPW_X	Ad. (22)	-2.35	3.02	-2.45	3.23	-2.54**				I(1)
	WCP_x	Ad. (14)	-3.50*	6.17*	-3.33**						I(0)
China to U.S: U.S. Imports	IR_M	Ad. (8)	-1.88	1.94	-2.03	2.72	-2.14				I(0)
	SW_M	Sh. (23)	-3.85**								I(0)
	PW_M	Ad. (14)	-1.69	2.10	-2.11	2.22	-1.23				
	ΔPW_M	Ad. (22)	-3.79**								I(1)
	WCP_M	Ad. (3)	-3.10	4.80	-3.05**						I(0)
(Table cont'd)											

Critical Values:	1%		-4.38	10.61	-3.75	7.88	-2.66	3.74	3.41		
	5%		-3.60	7.24	-3.00	5.18	-1.95	2.85	2.61		
	10%		-3.24	5.91	-2.62	4.12	-1.60	2.39	2.20		
Null Hypothesis:			$\gamma = 0$	$\gamma = \alpha_2 = 0$	$\gamma = 0$	$\gamma = \alpha_0 = 0$	$\gamma = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\gamma = 0$ [56]	
Trade Flow: Direction	Prod.	Outlier:	Test Statistic: τ_t	Test Statistic: ϕ_3	Test Statistic: τ_μ	Test Statistic: ϕ_1	Test Statistic: τ_t	Test Statistic: $\tau_{\beta t}$	Test Statistic: $\tau_{\alpha\mu}$	Test Statistic: T	Inference:
China to U.S: China Exports	IR_X	Ad. (15)	-0.74	0.38	-0.49	0.33	-0.84				
	ΔIR_M	Ad. (16)	-3.67**								I(1)
	SW_X	None	-3.97***								I(0)
	PW_x	Ad. (14)	-1.83	2.22	-2.16	2.33	-1.26				
	ΔPW_X	Ad. (22)	-3.62**								I(1)
	WCP_x	Ad. (3)	-3.88								I(0)
U.S. to China: China Imports	IR_M	Sh. (15)	-1.92	1.84	-1.69	1.91	-0.75				
	ΔIR_M	Ad. (15)	-4.50**								I(1)
	SW_M	Sh. (15)	-1.03	1.02	-1.01	0.63	-0.20				
	ΔSW_M	Ad. (15)	-2.86	5.70	-3.25**						I(1)
	PW_M	Ad. (17)	-2.05	2.77	-2.43	3.17	-2.47**				I(0)
	WCP_M	Ad. (25)	-1.13	1.06	-1.55	1.08	-1.08				
	ΔWCP_M	Ad. (25)	-4.20**								I(1)
U.S. to China: U.S. Exports	IR_X	Ad. (21)	-0.75	0.29	-0.58	1.06	-1.71				
	ΔIR_X	Ad. (22)	-3.00	5.25	-3.18**						I(1)
	SW_X	Sh. (15)	-1.09	0.76	-1.06	0.67	-0.47				
	ΔSW_X	Ad. (22)	-2.67	4.22	-2.98*	4.61*	-2.94***				I(1)
	PW_x	Ad. (22)	-1.41	3.40	-2.55	5.15*	-3.30***		0.129		I(0)
	WCP_x	Ad. (25)	-1.28	1.01	-1.55	1.07	-1.04				
	ΔWCP_X	Ad. (25)	-4.25**								I(1)
(Table cont'd)											

Critical Values:	1%		-4.38	10.61	-3.75	7.88	-2.66	3.74	3.41		
	5%		-3.60	7.24	-3.00	5.18	-1.95	2.85	2.61		
	10%		-3.24	5.91	-2.62	4.12	-1.60	2.39	2.20		
Null Hypothesis:			$\gamma = 0$	$\gamma = \alpha_2 = 0$	$\gamma = 0$	$\gamma = \alpha_0 = 0$	$\gamma = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\gamma = 0$ [56]	
Trade Flow: Direction	Prod.	Outlier:	Test Statistic: τ_t	Test Statistic: ϕ_3	Test Statistic: τ_μ	Test Statistic: ϕ_1	Test Statistic: τ_t	Test Statistic: $\tau_{\beta t}$	Test Statistic: $\tau_{\alpha\mu}$	Test Statistic: T	Inference:
U.S. to Japan: Japan Imports	IR_M	Ad. (7)	-3.41*	6.81*	-3.42**						I(0)
	SW_M	Ad. (1)	-2.88	4.12	-2.98*	5.06*	-2.39**				I(0)
	PW_M	Ad. (2)	-1.82	4.76	-1.88	1.76	-1.38				
	ΔPW_M	Ad. (17)	-3.38*	5.74	-3.39**						I(1)
	WCP_M	Sh. (6)	-4.93***								I(0)
U.S. to Japan: U.S. Exports	IR_X	Ad. (13)	-1.92	2.83	-2.40	3.38	-1.89*				
	ΔIR_X	None	-4.76***								I(1)
	SW_X	Ad. (1)	-3.05	5.01	-3.35**						I(0)
	PW_x	Ad. (1)	-2.32	2.49	-2.39	2.57	-1.22				
	$\Delta \mathrm{PW}_X$	Ad. (2)	-4.10**								I(1)
	WCP_x	Sh. (4)	-1.96	3.22	-2.61	8.82**	-4.10***		0.96		I(0)
U.S. to U.K: U.K. Imports	IR_M	Sh. (15)	-2.41	2.94	-2.45	3.20	-1.96**				I(0)
	SW_M	Ad. (11)	-2.99	4.52	-2.80	4.00	-0.51				
	ΔSW_M	Ad. (22)	-3.86***								I(1)
	PW_M	Ad. (1)	-6.58***								I(0)
	WCP_M	Ad. (9)	-4.16**								I(0)
U.S. to U.K: U.S. Exports	IR_X	Ad. (12)	-4.42***								I(0)
	SW_X	Ad. (11)	-3.38*	6.21*	-3.62**						I(0)
	PW_x	Ad. (1)	-5.47***								I(0)
	WCP_x	Ad. (9)	-3.65**								

Note: Source for Critical Values: (Enders, 2009, pg. 488-9, Statistical Tables A and B; Dickey and Fuller, 1981).

Table A.2. Johansen Cointegration Tests.

Test:			Trace Test: $H_0 = r$ $H_A > r$						Max Eigen Test: $H_0 = r$ $H_A = r + 1$					
			Case 1:		Case 2:		Case 3:		Case 1:		Case 2:		Case 3:	
Trade Flow and Product:		p	r = 0	r = 1	r = 0	r = 1	r = 0	r = 1	r = 0	r = 1	r = 0	r = 1	r = 0	r = 1
Canada to U.S.	SW	2	22.35***	3.60*	36.77***	3.85	32.13***	1.29	18.74***	3.60*	32.92***	3.85	30.84***	1.29
			(<0.001)	(0.068)	(<0.001)	(0.436)	(<0.001)	(0.255)	(0.002)	(0.068)	(<0.001)	(0.436)	(<0.001)	(0.255)
	PW	1	33.73***	4.16**	79.99***	4.26	55.13***	3.89**	29.57***	4.16**	72.73***	4.26	51.24	3.89**
			(<0.001)	(0.049)	(<0.001)	(0.375)	(<0.001)	(0.049)	(<0.001)	(0.049)	(<0.001)	(0.375)	(<0.001)	(0.049)
Brazil to U.S.	SW	1	6.78	1.19	9.89	3.48	9.63	3.45*	5.59	1.19	6.41	3.48	6.18	3.45*
			(0.348)	(0.321)	(0.651)	(0.496)	(0.311)	(0.063)	(0.399)	(0.321)	(0.740)	(0.496)	(0.590)	(0.063)
	PW	2	10.34	0.21	16.28	4.34	15.45*	3.53*	10.13*	0.21	11.94	4.35	11.92	3.52*
			(0.105)	(0.705)	(0.161)	(0.364)	(0.051)	(0.061)	(0.077)	(0.705)	(0.189)	(0.364)	(0.113)	(0.061)
China to U.S.	PW	1	22.25***	0.00	26.99***	4.41	23.97***	2.85*	22.25***	0.00	22.57***	4.42	21.12***	2.85*
			(<0.001)	(0.999)	(0.005)	(0.354)	(0.002)	(0.092)	(<0.001)	(0.999)	(0.004)	(0.354)	(0.003)	(0.092)
U.S. to China	IR	1	14.37**	0.003	15.74	1.26	15.08*	0.72	14.36**	0.003	14.48*	1.26	14.36**	0.724
			(0.023)	(0.963)	(0.186)	(0.913)	(0.058)	(0.914)	(0.013)	(0.963)	(0.082)	(0.914)	(0.048)	(0.395)
	SW	2	19.11***	1.81	26.04***	2.88	21.32***	0.31	17.30***	1.81	23.16***	2.88	21.01***	0.308
			(0.003)	(0.210)	(0.007)	(0.604)	(0.006)	(0.579)	(0.004)	(0.210)	(0.003)	(0.604)	(0.003)	(0.579)
	WCP	1	16.09***	5.50**	22.65**	6.61	21.64***	6.11**	10.59*	5.50**	16.04**	6.61	15.35**	6.11**
			(0.011)	(0.022)	(0.023)	(0.148)	(0.005)	(0.013)	(0.065)	(0.022)	(0.047)	(0.148)	(0.031)	(0.013)
U.S. to Japan	PW	1	44.73***	4.09**	46.44***	7.65*	46.33***	7.24***	40.64***	4.09**	40.79***	7.65*	39.09***	7.24***
			(<0.001)	(0.051)	(<0.001)	(0.096)	(<0.001)	(0.007)	(<0.001)	(0.051)	(<0.001)	(0.096)	(<0.001)	(0.007)

Table A.3. Johansen Cointegration LR Test of Restriction.

		Test of the Restriction: $H_0 = Case\ 2$ $H_A = Case\ 3$	
Trade Flow	Product	r = 0	r = 1
Canada to U.S.	SW	4.64*	2.56
		(0.098)	(0.120)
	PW	21.86***	0.37
		(<0.001)	(0.545)
Brazil to U.S.	SW	0.26	0.03
		(0.879)	(0.867)
	PW	0.84	0.82
		(0.658)	(0.365)
China to U.S.	PW	3.02	1.57
		(0.221)	(0.211)
U.S. to China	IR	0.65	0.54
		(0.722)	(0.464)
	SW	4.72*	2.57
		(0.095)	(0.109)
	WCP	1.01	0.50
		(0.605)	(0.479)
U.S. to Japan	PW	2.11	0.41
		(0.348)	(0.521)

Table A.4. ARDL Cointegration Tests.

			Critical Bounds F-Test :									
			Without Deterministic Trend					With Deterministic Trend (Unrestricted Intercept)				
			Sig. Level:	I(0)	I(1)	I(0)	I(1)	Sig. Level:	I(0)	I(1)	I(0)	I(1)
			10%	4.29	5.08	3.30	3.80	10%	6.01	6.78	4.43	4.96
			5%	5.39	6.35	4.09	4.66	5%	7.36	8.27	5.38	5.96
			1%	8.17	9.29	6.03	6.76	1%	10.61	11.65	7.59	8.35
			Case:	Unrestricted Intercept		Restricted Intercept		Case:	Unrestricted Trend		Restricted Trend	
ORIG:	DEST:	PROD:	(p,q):	F-Stat	T-Stat	F-Stat		(p,q):	F-Stat	T-Stat	F-Stat	
Canada	U.S.	IR	(3,1)	17.22***	-5.53	11.49***		(3,1)	15.18***	-5.40***	11.09***	
				(<0.001)	(<0.001)				(<0.001)	(0.004)		
Canada	U.S.	IR	(1,3)	11.30***	-4.75	7.60***		(3,1)	23.30***	-6.08***	16.06***	
		(Log-Log)		(0.006)	(0.004)				(<0.001)	(<0.001)	(<0.001)	
Brazil	U.S.	IR	(1,4))	539.96***	-20.84***	404.28***		(1,4)	549.16***	-20.67***	393.28***	
				(<0.001)	(<0.001)	(<0.001)			(<0.001)	(<0.001)	(<0.001)	
China	U.S.	IR	(3,3)	453.3***	-30.11***	362.28***		(3,3)	404.43***	-28.42	293.56	
				(<0.001)	(<0.001)	(<0.001)			(<0.001)	(<0.001)	(<0.001)	
China	U.S.	SW	(1,4)	358.85***	-26.71***	318.6***		(2,4)	659.08***	-36.21***	481.75***	
				(<0.001)	(<0.001)	(<0.001)			(<0.001)	(<0.001)	(<0.001)	
U.S.	Japan	IR	(1,3)	204.25***	-20.21***	138.69***		(1,3)	187.05***	-19.25***	128.14***	
				(<0.001)	(<0.001)				(<0.001)	(<0.001)	(<0.001)	
U.S.	U.K.	SW	(4,1)	14.47***	-3.46**	10.58***		(4,1)	17.40***	-3.94*	11.63***	
				(0.003)	(0.043)	(0.002)			(0.003)	(0.053)	(0.004)	

Note: Source for Critical Values: (Narayan, 205)

Table A.5. ARDL ECM LR tests for Case.

	Canada to U.S. IR		Brazil to U.S: IR		China to U.S.: IR		China to U.S.: SW		U.S.to Japan: IR		U.S. to U.K.	
	LR (3,1)	Conclusion	LR (1,4)	Conclusion	LR (3,3)	Conclusion	LR (1,4)	Conclusion	LR (1,3)	Conclusion	LR (4,1)	Conclusion
$H_R: Case\ 3$ $H_{UR}: Case\ 5$	2.18 (0.14)	Fail to Reject Case 3	3.5* (0.061)	Fail to Reject Case 3	1.01 (0.293)	Fail to Reject Case 3	0.12 (0.735)	Fail to Reject Case 3	2.4 (0.121)	Fail to Reject Case 3	0.22 (6.399)	Fail to Reject Case 3

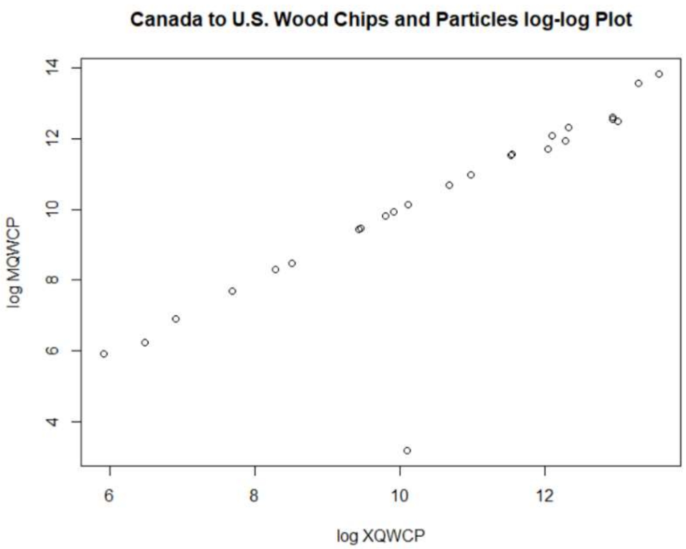

Figure A.7. Canada to U.S. Plot of log(imports) over log(exports).

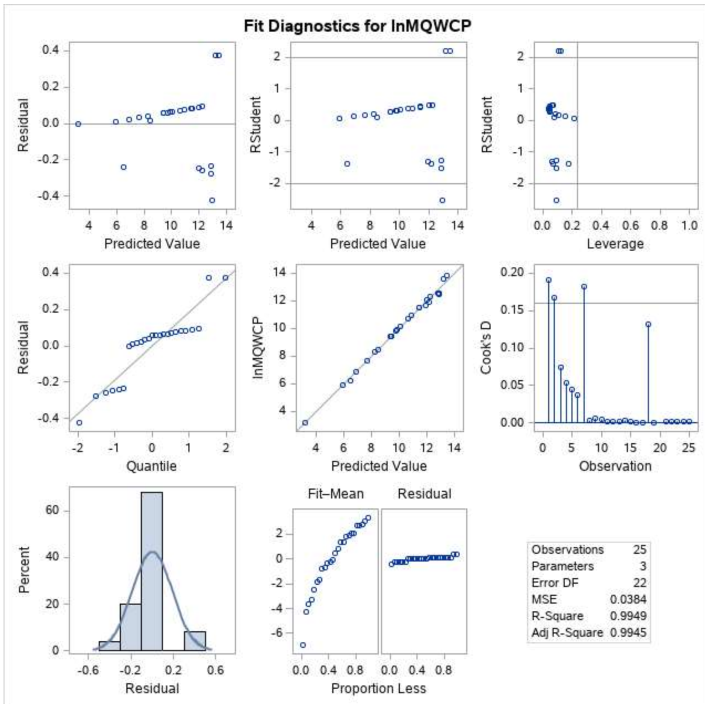

Figure A.8. Residual Plots for Canada to U.S. WCP.

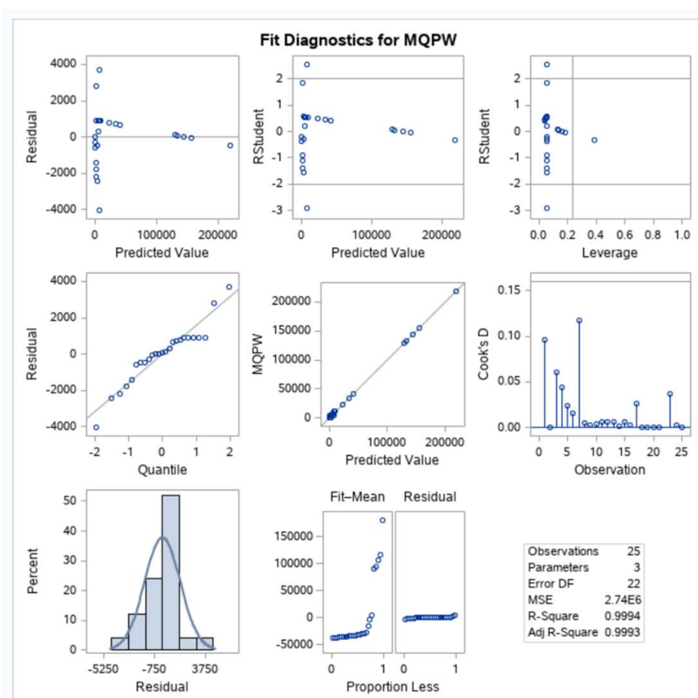

Figure A.9. Residual Plots for U.S. to China PW.

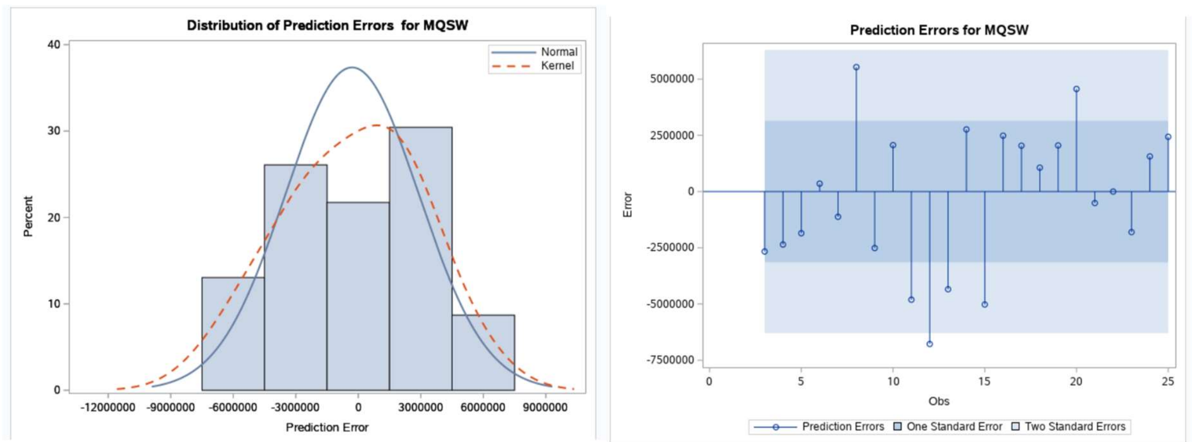

Figure A.10. Residual Plots for Canada to U.S. SW (import side).

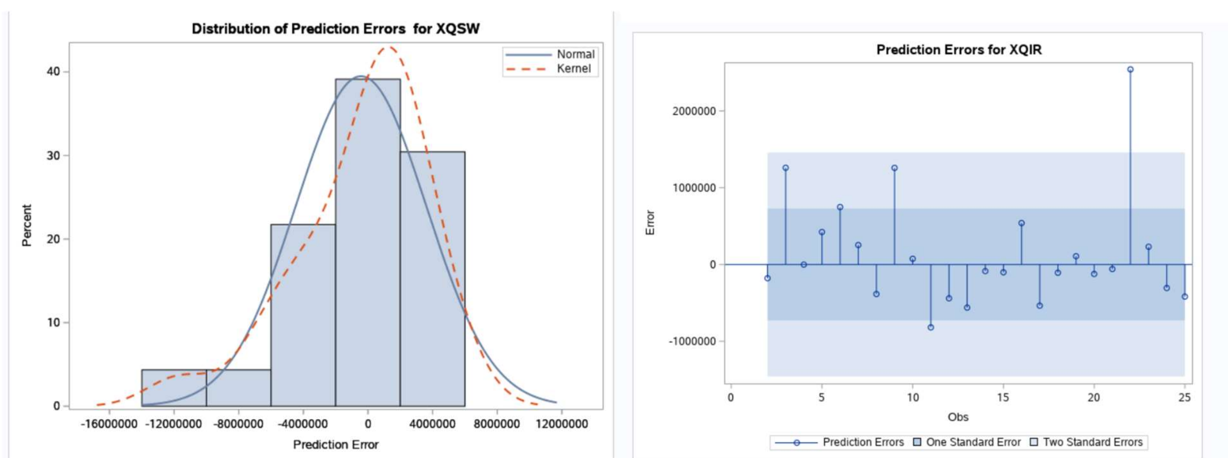

Figure A.11. Residual Plots for Canada to U.S. SW (export side).

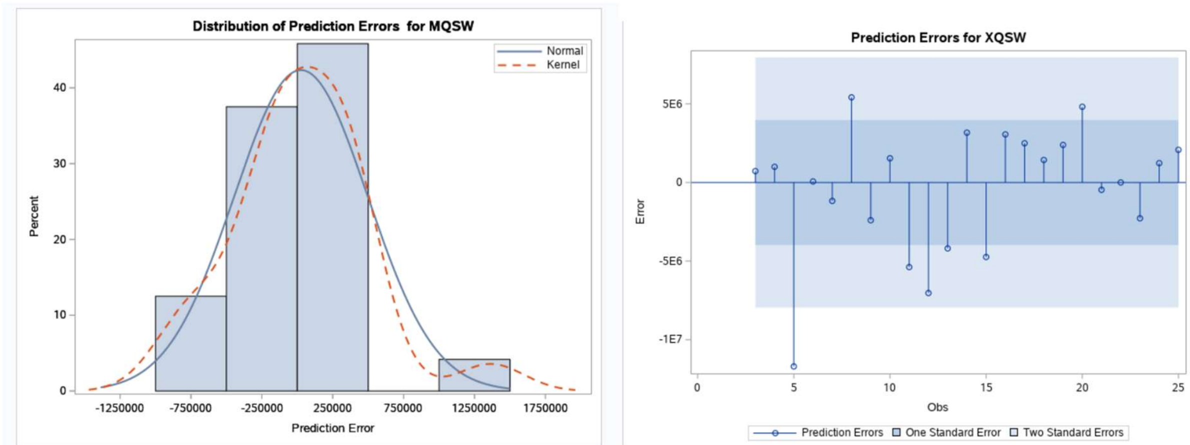

Figure A.12. Residual Plots for U.S. to China SW (import side).

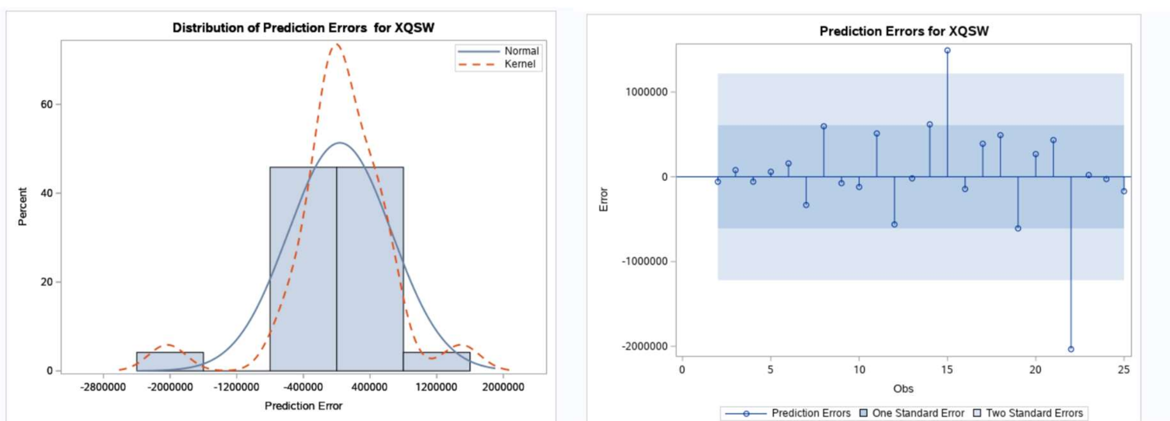

Figure A.13. Residual Plots for U.S. to China SW (export side).

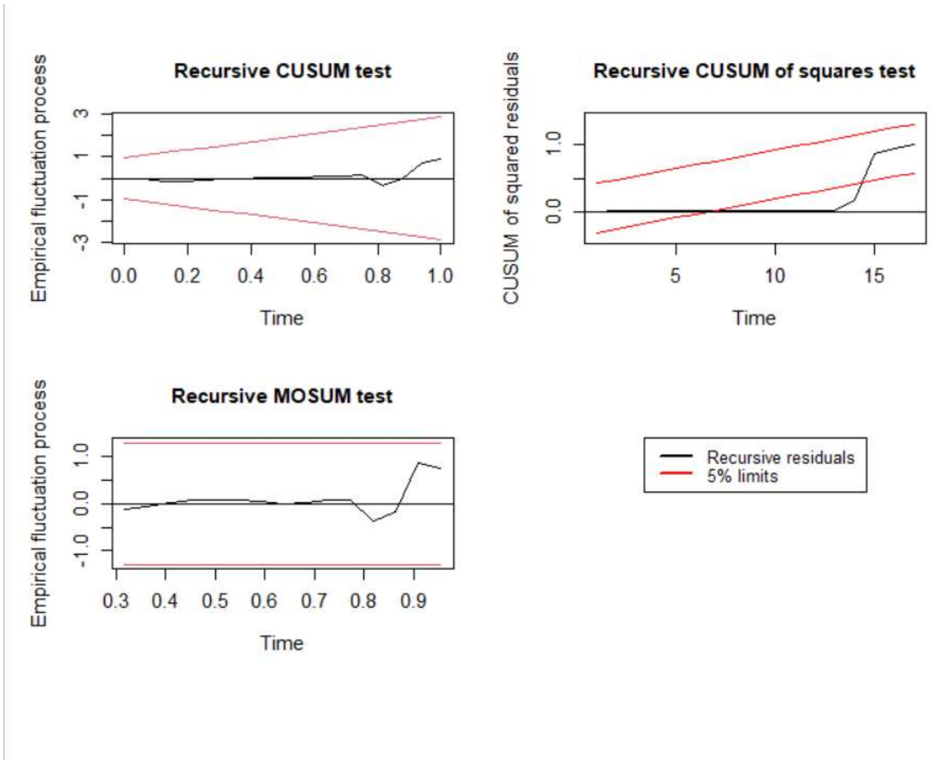

Figure A.14. CUMSUM and CUMSQ for Canada to U.S. IR.

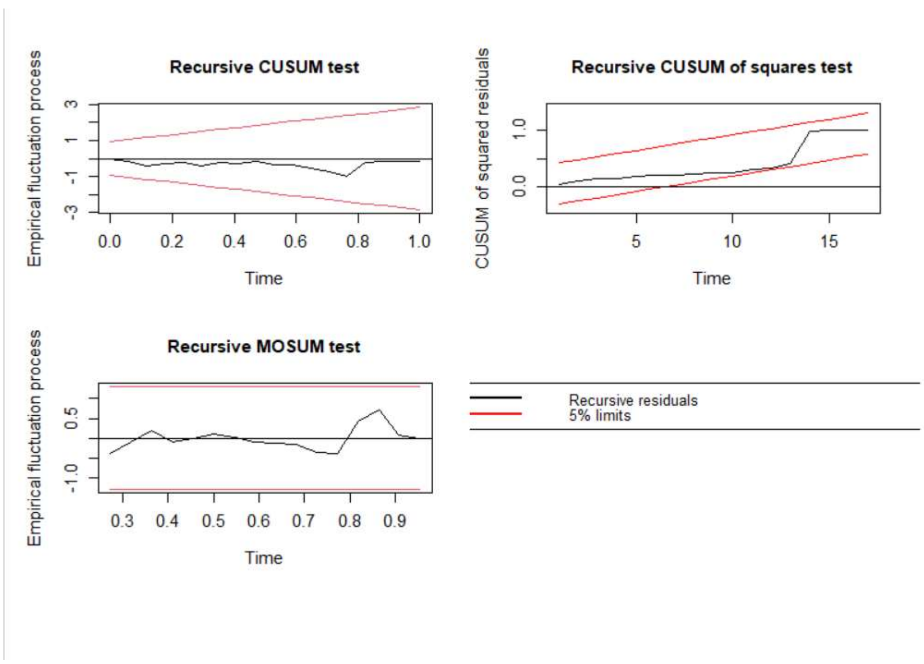

Figure A.15. CUMSUM and CUMSQ for U.S. to Japan IR.

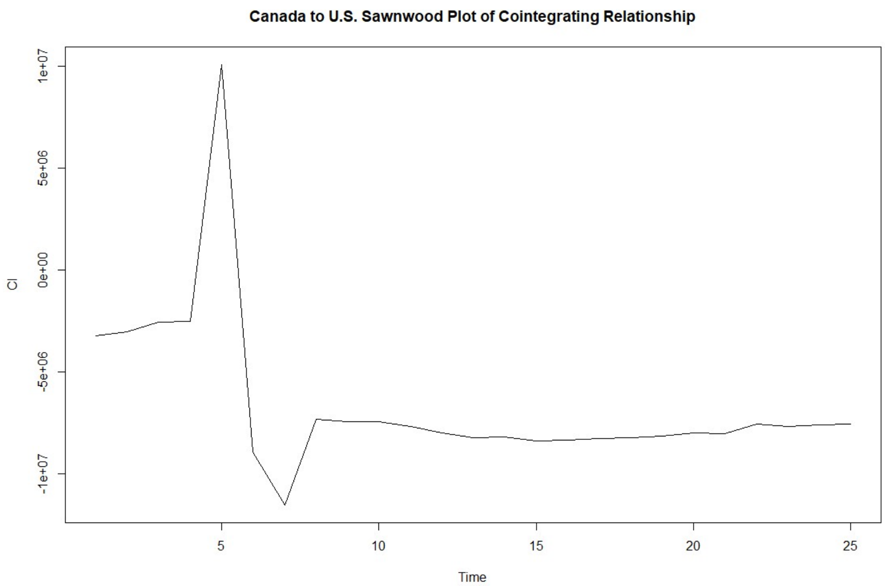

Figure A.16. Canada to U.S. Sawnwood Plot of Cointegrating Relationship.

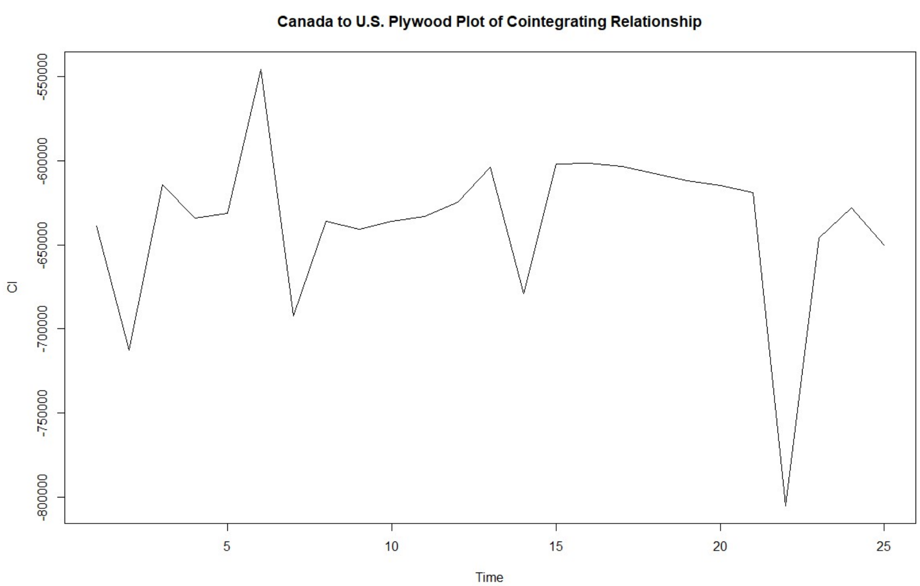

Figure A.17. Canada to U.S. Plywood Plot of Cointegrating Relationship.

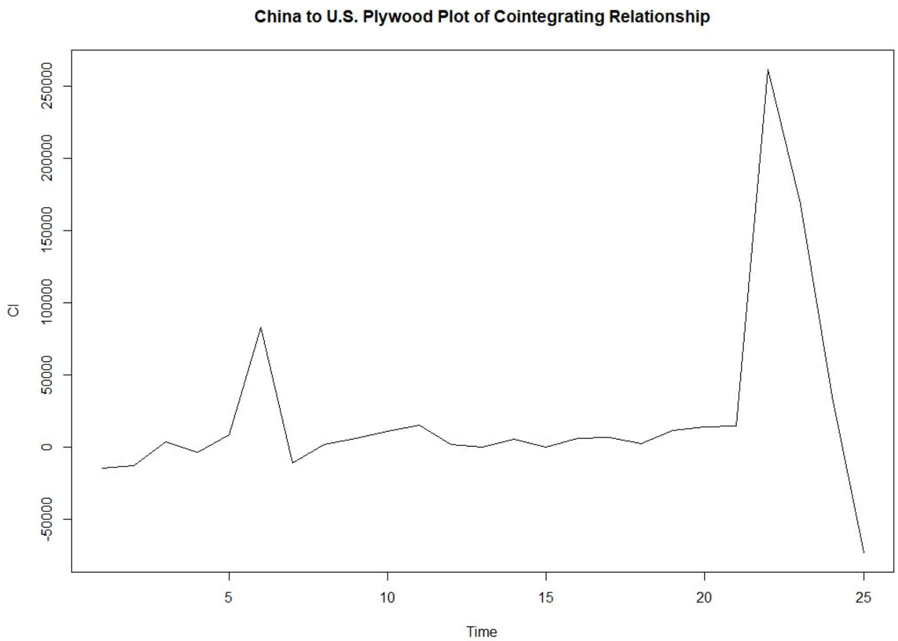

Figure A.18. U.S. to China Plywood Plot of Cointegrating Relationship.

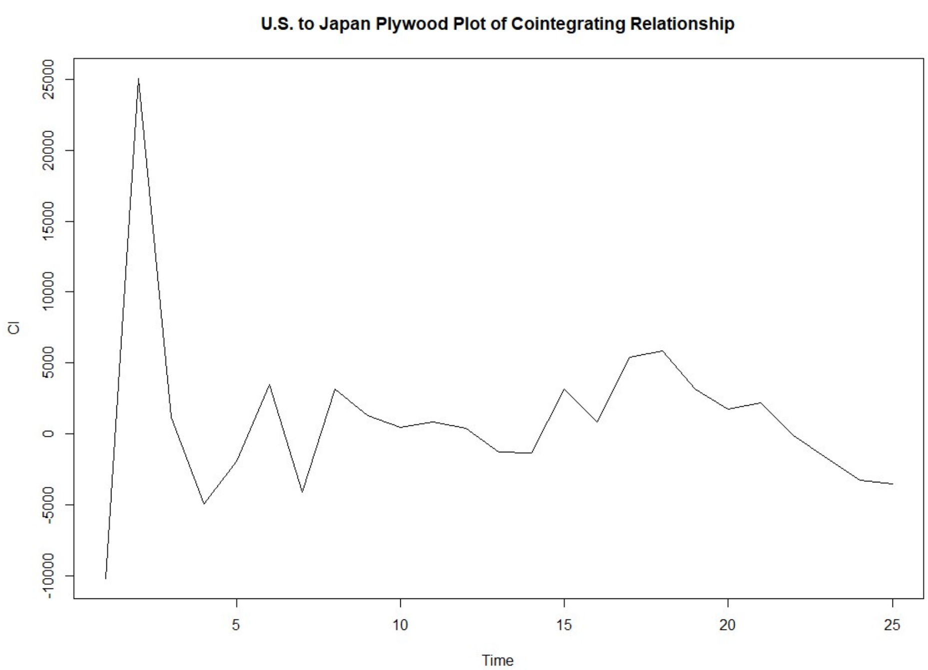

Figure A.19. Japan to the U.S. Plywood Plot of Cointegrating Relationship.

References

Ahad, M., 2017. Impact of financial development on trade balance: An ARDL cointegration and causality approach for Pakistan. *Global Business Review, 18(5),* pp.1199-1214.

Ali, D.A., Johari, F., Alias, M.H., 2019. Quality assessment of trade data in Malaysia. *Malaysian Journal of Economic Studies*, *56*(1), pp.23-42.

Bisgaard, S., Kulahci, M., 2011. *Time series analysis and forecasting by example*. John Wiley & Sons.

Brown, R.L., Durbin, J. and Evans, J.M., 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological),* 37(2), pp.149-163.

Buongiorno, J., 2018. On the accuracy of international forest product statistics. *Forestry: An International Journal of Forest Research*, 91(5), pp.541-551.

Buongiorno, J., Johnston, C., 2018. Effects of parameter and data uncertainty on long-term projections in a model of the global forest sector. *Forest Policy and Economics*, 93, pp.10-17.

Cafiero, C., Melgar-Quinonez, H. R., Ballard, T. J., Kepple, A. W., 2014. Validity and reliability of food security measures. *Annals of the New York Academy of Sciences*, 1331(1), pp. 230-248.

Carfagna, E., Pratesi, M., Carfagna, A., 2013. Methodological developments for improving the reliability and cost-effectiveness of agricultural statistics in developing countries. *In Proceedings of the 59th World Statistics Congress of the International Statistical Institute*. *International Statistical Institute, The Hague, the Netherlands*.

Chatham House, 2018. Resource Trade Earth – Data. *Chatham House*. Available from URL: https://resourcetrade.earth/ [Accessed 03 January 2022]

Circlaeys, S., Kanitkar, C. and Kumazawa, D., 2017. Bilateral trade flow prediction. *Unpublished manuscript*, available Available from URL: at http://cs229. stanford. edu/proj2017/final-reports/5240224. pdf.

Dal Bianco, A., Boatto, V. L., Caracciolo, F., Santeramo, F. G., 2016. Tariffs and non-tariff frictions in the world wine trade. *European Review of Agricultural Economics*, 43(1), pp.31-57.

Development, U. N. C. on T., 2021. Key Statistics and Trends in International Trade 2020. *United Nations.* https://www.un-ilibrary.org/content/books/9789210056502

Enders, W. (2009). *Applied Econometric Time Series*. Wiley. ISBN 10: 0470505397 / ISBN 13: 9780470505397.

FAO. Forestry Production and Trade. License: CC BY-NC-SA 3.0 IGO. Available from URL: https://www.fao.org/faostat/en/#data/FO.[Accessed 1 July 2022]

FAO. Forestry Trade Flows. License: CC BY-NC-SA 3.0 IGO. Available from URL: https://www.fao.org/faostat/en/#data/FT .[Accessed 23 June 2022]

Ferrantino, M. J., Liu, X., Wang, Z., 2012. Evasion behaviors of exporters and importers: Evidence from the US–China trade data discrepancy. *Journal of International Economics*, 86(1), pp.141-157.

Ferrantino, M. J., Wang, Z., 2008. Accounting for discrepancies in bilateral trade: The case of China, Hong Kong, and the United States. *China Economic Review*, 19(3), pp.502-520.

Ferrier, P. M., 2021. Detecting origin fraud with trade data: the case of US honey imports. *Australian Journal of Agricultural and Resource Economics*, 65(1), pp.222-245.

Food and Agriculture Organization of the United Nations (FAO), 2020. FAO Yearbook of Forest Products 2018. *Food and Agriculture Organization of the United Nations Publisher*, Rome.

Franses, P.H. and Haldrup, N., 1994. The effects of additive outliers on tests for unit roots and cointegration. *Journal of Business & Economic Statistics*, 12(4), pp.471-478.

Gopinath, M., Batarseh, F.A., Beckman, J., Kulkarni, A. and Jeong, S., 2021. International agricultural trade forecasting using machine learning. *Data & Policy*, 3.

GOV.U.K, 2022. UK Integrated Online Tariff. Available from URL: https://www.trade-tariff.service.gov.uk/find_commodity?country=US [Accessed 11 July 2022]

Granger, C. W., Newbold, P., 1974. Spurious regressions in econometrics. *Journal of Econometrics*, 2(2), pp.111-120.

Guru-Gharana, K.K., 2012. Econometric investigation of relationships among export, FDI and growth in India: An application of Toda-Yamamoto-Dolado-Lutkephol Granger causality test. *The Journal of Developing Areas*, pp.231-247.

Hamanaka, S., 2012. Whose trade statistics are correct? Multiple mirror comparison techniques: A test case of Cambodia. *Journal of Economic Policy Reform*, 15(1), pp.33-56.

Hill, R. C., Griffiths, W. E., Lim, G. C., 2018. *Principles of econometrics*. John Wiley & Sons.

Hung-Pin L. Renewable energy consumption and economic growth in nine OECD countries: bounds test approach and causality analysis. *ScientificWorldJournal*. 2014 Jan 14;2014:919167. doi: 10.1155/2014/919167. PMID: 24558343; PMCID: PMC3914386.

Iheanacho, E., 2017. ARDL approach to trade libralisation and economic growth in the developing country: Evidence from Nigeria. *African Research Review, 11(2),* pp.138-159.

International Trade Centre (ITC). 2022. ITC Trade Map. Available from URL: https://www.trademap.org/. [Accessed April 2022]

International Trade Centre (ITC). 2022. Market Access Map. Available from URL: https://www.macmap.org/. [Accessed July 2022]

Jayasinghe, S., Beghin, J. C., Moschini, G., 2010. Determinants of world demand for US corn seeds: the role of trade costs. *American Journal of Agricultural Economics*, 92(4), pp.999-1010.

Johansen, S. and Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and statistics, 52(2),* pp.169-210.

Kallio, A. M. I., Solberg, B., 2018. On the reliability of international forest sector statistics: problems and needs for improvements. *Forests, 9(7),* pp.407.

Kazunobu, H., 2020. Departure months and discrepancy in mirror trade data (No. 795). *IDE Discussion Paper.*

Kee, H.L. and Nicita, A., 2016, November. Trade frauds, trade elasticities and non-tariff measures. *In 5th IMF-World Bank-WTO Trade Research Workshop*, Washington, DC, November (Vol. 30).

Kottou, E.M., Grubelich, T.A. and Wang, X., 2020, December. Bilateral Trade Flow Prediction Models Enhanced By Wavelet and Machine Learning Algorithms. *In 2020 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 1510-1516). IEEE.

Kripfganz, S. and Schneider, D.C., 2018, September. ardl: Estimating autoregressive distributed lag and equilibrium correction models. *In Proceedings of the 2018 London Stata conference* (p. 59).

Latta, G. S., Sjølie, H. K., Solberg, B., 2013. A review of recent developments and applications of partial equilibrium models of the forest sector. *Journal of Forest Economics*, 19(4), pp.350-360.

Lawal, E.O. and Ezeuchenne, K., 2017. International trade and economic growth in Nigeria. *Journal of Humanities and Social Science,* 22(6), pp.35-43.

Lee Rodgers, J., Nicewander, W. A., 1988. Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1), pp.59-66.

Lütkepohl, H., 1993. Introduction to multiple time series analysis. Springer Science & Business Media.

Markowicz, I., Baran, P., 2020. A new method for calculating mirror data asymmetry in international trade. *Oeconomia Copernicana*, 11(4), pp.637-656.

Md Reza, S., Fan, H., Wang, B., Bhuiyan, M.A. and KM Mehedi, A., 2019. Trade (exports) as an opportunity for Bangladesh: A VECM analysis. *The International Trade Journal,* 33(1), pp.95-110.

Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives,* 31(2), pp.87-106.

P.K Narayan, The Saving and Investment Nexus for China: Evidence from Cointegration Tests, *Applied Economics*, 37, (2005), 1979–1990.

Nkoro, E. and Uko, A.K., 2016. Autoregressive Distributed Lag (ARDL) cointegration technique: application and interpretation. *Journal of Statistical and Econometric methods*, 5(4), pp.63-91.

Pardo-Herrera, C., 2021. The International Links of Peruvian Illegal Timber: A Trade Discrepancy Analysis. *Targeting Natural Resource Corruption.*

Pesaran, M.H., Smith, R.J. and Shin, Y., 1997. Structural analysis of vector error correction models with exogenous I (1) variables (first version) (No. 7). *Edinburgh School of Economics, University of Edinburgh.*

Pesaran, M.H., Shin, Y. and Smith, R.J., 2001. Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics,* 16(3), pp.289-326.

Pyzhev, A. I., Gordeev, R. V., Vaganov, E. A., 2021. Reliability and Integrity of Forest Sector Statistics—A Major Constraint to Effective Forest Policy in Russia. *Sustainability*, 13(1), 86.

Riviere, M., Caurla, S., 2020. Representations of the Forest Sector in Economic Models. *Œconomia. History, Methodology, Philosophy*, (10-3), pp.521-553.

Sen, R., 2000. Analyzing International Trade Data in a Small Open Economy: The Case of Singapore. *ASEAN Economic Bulletin*, 17(1), pp.23-35.

Skjerstad, S. H., Kallio, A. M. I., Bergland, O., Solberg, B., 2021. New elasticities and projections of global demand for coniferous sawnwood. *Forest Policy and Economics*, 122, 102336.

Stamer, V., 2022. Thinking Outside the Container: A Sparse Partial Least Squares Approach to Forecasting Trade Flows.

Steele, A., Lebedys, A., Arkhipova, I., 2021, February 9-11. [Workshop Presentation]. *ECE and FAO Forest Product Statistics Workshop, online meeting*. https://unece.org/sites/default/files/2021-02/steel.pdf [Accessed November 2021]

Torres-Rojo, J. M., 2021. Illegal Logging and the Productivity Trap of Timber Production in Mexico. *Forests*, 12(7), pp.838.

Vincent, J. R., 2004. Detecting illegal trade practices by analyzing discrepancies in forest products trade statistics: an application to Europe, with a focus on Romania. *World Bank Policy Research Working Paper*, (3261).

World Bank national accounts data, and OECD National Accounts data files, 2020. Trade (% of GDP). Available from URL: https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS [Accessed 07 January 2022]

Worldwide Governance Indicators, 2018. Control of Corruption Statistics, *World Bank*. Available from URL: https://info.worldbank.org/governance/wgi/#home [Accessed 04 January 2022]

WTO (World Trade Organization), 2012. Detailed Query, WTO Integrated Trade Intelligence Portal. Available from URL: http://i-tip.wto.org/goods/Forms/TableView.aspx [Accessed 22 October 2022]

WTO (World Trade Organization), 2018. Tariff Download Faciliy,WTO Statiastics Library. Available from URL: http://tariffdata.wto.org/Default.aspx?culture=en-US [Accessed 11 July 2022]

WWF. We need to Safeguard our Forests, 2019. Available from URL: https://wwf.panda.org/discover/our_focus/forests_practice/importance_forests/? [Accessed 04 January 2022]

WWF. Forest Sector Transformation, 2020. Available from URL: https://wwf.panda.org/discover/our_focus/forests_practice/forest_sector_transformation_updated/ [Accessed 04 January 2022]

Yurik, S., Pushkin, N., Yurik, V., Halík, J., Smutka, L., 2020. Analysis of Czech agricultural exports to Russia using mirror statistics. *Entrepreneurial Business and Economics Review,* 8(2), pp.27-46.

Zapata, H.O. and Rambaldi, A.N., 1997. Monte Carlo Evidence on Cointegration and Causation. *Oxford Bulletin of Economics and statistics*, 59(2), pp.285-298.

Vita

Morgan Daigle hails from the small rural community Branch, Louisiana just outside of Church Point. She received her Bachelor of Science in Business Administration with a major in Economics in December of 2019 from the University of Louisiana at Lafayette. She gained professional experience in the financial, medical, and customer service industries during her undergraduate studies, and the year leading up to graduate school. In the Fall of 2020, she began her graduate studies in the Department of Agricultural Economics and Agribusiness at Louisiana State University. In the Fall of 2021, she accepted a graduate research assistantship under the guidance of Dr. Jinggang Guo. She anticipates receiving her Master of Science degree in Agricultural Economics in the Spring of 2023. Morgan wishes to pursue a career as an Economist or Data Analyst for a government or non-profit organization.