SPATIAL PATTERNS OF LANDSIDE TRADE IMPEDANCE IN CONTAINERIZED SOUTH AMERICAN EXPORTS

by

Kara Carroll Tiller

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Geography and Urban Regional Analysis

Charlotte

2015

Approved by:

____
Dr. Jean-Claude Thill

____
Dr. Tyrel Moore

____
Dr. Harrison Campbell, Jr.

____
Dr. John Daniels

ABSTRACT

KARA CARROLL TILLER. Spatial patterns of landside trade impedance in containerized South American exports. (Under the direction of DR. JEAN-CLAUDE THILL)

Scholars of Latin America's economic trajectory claim that higher domestic transportation costs stand as a significant barrier to trade and economic integration, and are primarily related to the region's inadequate transportation infrastructure. However, calls for infrastructure investment have not been preceded with the multi-scalar trade and transportation analyses necessary to fully investigate Latin American trade impedance. Although these steps have been broached by the Inter-American Development Bank (IDB), the present research is born of a unique conceptual approach and employs a dataset developed through an innovative verification process, addressing legacy data constraints in the field. Upon this foundation, this research establishes the existence, extent, location, and spatial distribution of trade impedance throughout South America. The vast majority of U.S.-bound export trade flows are found to have trade impedance proportionate to distance; for the total dataset and all segments, the share of flows by volume with disproportionately high expected distance (trade impedance) ranges from 1.3% to 11.2%. Global spatial autocorrelation analysis reveals that mean and median trade impedance ratios at origins are spatially clustered for the total dataset, but randomly distributed for all commodity-based segments. Mean and median trade impedance ratios at destinations are randomly distributed for the total dataset and all segments. Local spatial autocorrelation analysis reveals high trade impedance ratio mean and median clusters and outliers at origins and destinations, which vary by commodity segment.

DEDICATION

To my daughter, the most rare and precious gift.

.

ACKNOWLEDGEMENTS

I would like to thank my Ph.D. advisor and mentor, Dr. Jean-Claude Thill, for his model of intrepid visionary energy; a trail-blazer needn't suffer one trailing, yet he did. I wish to thank my Master's advisor and mentor, Dr. Tyrel Moore, for being a beacon for lost aviators, without whom I would not be a geographer. I would like to thank Dr. Harrison Campbell for the inspiration of his steadfast commitment to our field and vast knowledge of it. I wish to thank Dr. John Daniels for his kindness and patience in his service on this committee. I would like to thank my instructors for their teaching, especially Jamie Strickland, Dr. Wei-Ning Xiang, and Dr. William Graves. I am thankful to the staff of the Department of Geography & Earth Sciences and The Graduate School for their kind assistance and thoughtfulness, especially Stephen Bowman.

I am grateful for the assistance and support of the Department of Geography & Earth Sciences and The Graduate School at The University of North Carolina at Charlotte, through many years of the Graduate Assistant Support Plan. I wish to express my gratitude for the Lucille P. and Edward C. Giles Dissertation-Year Graduate Fellowship, which was granted at a pivotal time. I am thankful for the assistance provided by the Southeastern Transportation Center's Regional Outstanding Student Fellowship. I am thankful for support from alumni through the Kirk S. Ashton Memorial Scholarship.

I would like to thank my mother, father, and sister for standing by me in every way (since we know who we are, when we step out). I thank The Divas for kindness that bears all and friendship that shows up alongside. "To conclude this matter," I would like to thank Robert Carroll for being my coach and marching band. Thanks for pointing me toward the referee, who sees every play in every game - and knows the score.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMS	U.S. Customs and Border Protection Automated Manifest System
CONAIE	Confederation of Indigenous Nationalities of Ecuador
ETI	Enabling Trade Index
EU	European Union
FARC	Fuerzas Armadas Revolucionarias de Colombia
GDP	Gross domestic product
GNP	Gross national product
IDB	Inter-American Development Bank
IPIAI	Infrastructure Private Investment Attractiveness Index
IQGI	Infrastructure Quality Gap Index
LAC	Latin America and the Caribbean
O-D	Origin-destination
OLS	Ordinary least squares
TEU	Twenty-foot equivalent unit
TNC	Transnational corporation
U.S.	United States of America
UNCTAD	United Nations Conference on Trade and Development
UNECLAC	United Nations Economic Commission for Latin America and the Caribbean
VAT	Value added tax
WEF	World Economic Forum

INTRODUCTION

Latin America encompasses a spectrum of economic realities, both in terms of seemingly contradictory aspects of its economic performance as a world region, and also among its diverse national economies and their varied relationships with global trading partners. A pendulum-like regional history swings between economic and political ascension and downfall, cohesion and upheaval, and switchbacks along a path of economic development and integration (Hill, 2008, p. 3).

While Latin America as a region has had the slowest-growing productivity among world regions over the past 40 years (World Bank, 2010), it is also a region featuring nations best-positioned to become economic powerhouses in future decades, such as Brazil (Wu & Lin, 2008; O'Neill et al., 2005). As a region, Latin America boasts a reversal of its economic trajectory from decades past; the region has changed its global economic role from that of a large debtor to that of a creditor in recent years (World Bank, 2010). Its "silent revolution" (Green, 2003) of improvement of economic, financial, and associated governmental systems over the past 20 years (World Bank, 2010) led to trade liberalization and economic growth – yet it also led to failures, upheavals, disappointments, and false starts after its initial surge of success (Green, 2003). Considering Latin America's status as a developing region with great unfulfilled economic potential, economists note particular areas of concern for the region's economic integration, growth, and long-term stability.

Though the issues are many, the current focus remains upon creating comparative advantage in the region (Mesquita Moreira et al., 2013). Economists claim that an increasing concern in direct opposition to that goal is the region's high transportation costs. Lee notes that Latin American regional transport costs are now more detrimental to trade than are tariffs (Mesquita Moreira, Volpe, & Blyde, 2008; Lee, 2009, p. 2; Mesquita Moreira et al., 2013). Although the comparison is often made between the trade impediments of transport costs and tariffs, the IDB notes several characteristics of transport costs that make them unique trade barriers: their high variability over time; their differing unit costs based upon weight (Hummels, 2001; Inter-American Development Bank, 2010; Mesquita Moreira et al., 2013), value, and perishability; and the fact that highly complex policy actions are required to affect change in transportation costs, due to the varied nature of their influences. These influences include volume and composition of trade flows, degree of competition (between and within modes), and the varying quality of national infrastructures (Mesquita Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11; Mesquita Moreira et al., 2013). These factors, which span the gamut of trade and other national policy and economic/sectoral issues – as well as those more directly associated with transportation - can be summarized under the umbrella term *trade impedance* (Kockelman & Ruiz-Juri, 2003; Navajas et al., 2010).

According to economists, it is increasingly clear that for the most pressing economic issues of the region – economic development and inequality – change cannot be effectuated without first reducing higher transport costs within the region (Fay & Morrison, 2007, pp. 24-26; Mesquita Moreira et al., 2013). Closely related to this is the region's export composition, which has built the "beaten paths" (Plane, 1984, p. 244) currently utilized within the region (and while this has long been generally accepted in the literature, Mesquita Moreira et al. [2013] have recently acknowledged these beaten paths in a nod to the Inca road network).

Those who study the region claim that these disproportionate transport costs are primarily related to the region's inadequate transportation infrastructure. However, such calls for improvement are made without a preceding call for sound investigation regarding if, where, and/or how transportation infrastructure improvements could reduce trade impedance. Although some progress has recently been made in acknowledging this (Mesquita Moreira et al., 2013), no investigation yet has been appropriately sophisticated, multi-scalar, and complete in its geographic coverage in order to adequately and accurately address the intermediate steps that must be performed. Such steps are needed to ensure the most efficient and effective solutions are found, serving as a basis to address concerns ranging from the financial, to the socioeconomic, as well as the environmental.

As will be detailed in the following sections, history demonstrates that economic policy, processes, and regulation within Latin America pose significant challenges for financing and implementing infrastructure improvements (Green, 2003; Fay & Morrison, 2007). If indeed infrastructure improvement is markedly more challenging within Latin America (as compared to developed regions) in terms of finance and implementation, this fact alone suggests it is even more dire that any suggested infrastructure improvements be targeted for maximum effectiveness in reducing trade impedance.

The first step must be to determine whether regional trade impedance is indeed disproportionate - and, if so, where. In the present research, this is labeled Steps 1 (Existence Step) and 2 (Location Step). The spatial distribution of trade impedance must be investigated at global and local scales, leading to what is in this research Step 3

(Spatial Pattern Step). Step 4 (Association Step) - set out here as a future extension of this current analysis - would be used to determine which variables (such as transportation cost factors or commodity composition/sectoral factors) are most highly associated with disproportionate trade impedance. These four steps stand as crucial preliminary exploration into the need for transportation infrastructure improvement in South America, with a goal to expand the study to include all of Latin America. This study first lays the groundwork for focused further analysis of sub-continental regions in South America, delineated by trade flow corridors, as an initial launch of a larger research endeavor.

These initial steps enable the study to be narrowed to the geographical framework and economic context of the regional scale. If - or where - trade impedance is found to be disproportionate to actual distance, blanket statements regarding high transport costs within Latin America can be lifted, and the existence of disproportionate trade impedance can begin to be specified by location of occurrence and association with potential trade impedance factors. The Latin American sub-region of South America is analyzed, as the data reveals its trade context to be quite different than that of Central America, Caribbean island nations, and Mexico. Hornbeck (2010, p. 5) states that improving access to the U.S. market by lowering trade barriers could provide increased long-term trade and investment opportunities not only for the region as a whole, but for the subset of South American countries. Combined with its BRIC nation potential, this sub-region appears to be particularly poised to reap benefits from improved market access.

Therefore the objective of this dissertation is to specify further from the generalizations regarding high Latin American transport cost, for a South American subset of the data, determining: (1) whether evidence of disproportionately high trade impedance is found, and, if so, for which commodity categories and specific origin-destination [O-D] flows does disproportionately high trade impedance exist; thus establishing through visualization (2) where disproportionately high trade impedance is evidenced (identifying locations/regions); and (3) the spatial distribution of trade impedance, in terms of its pattern at both the global (in terms of the dataset) and local scales. These three steps set the stage for a proposed future extension, in which potentially trade impeding factors are investigated as to their association with disproportionate trade impedance. Thus the future extension will serve to separate transportation cost effects from the effects of other potentially trade impeding factors (such as trade policy effects).

1.1. Study Area: South America

The term "Latin America" originated in the 1850s. It commonly stands to represent those nations of South and Central America which share: (1) a legacy of colonial rule by Spain and Portugal; (2) the ensuing development patterns associated with that legacy; as well as (3) language; and (4) geography (Bulmer-Thomas, 2003, p. 1). A legacy of colonial rule has resulted in several characteristics related to development, including: development patterns centered on natural resource extraction; an export composition based upon natural resources and manufactured products of natural resources (primary commodities); inequality resulting from post-colonial land distribution; and continuing governmental struggles to structure regulation and provision of basic services upon independence (Bulmer-Thomas, 2003; Fay & Morrison, 2007). (Although the region shares both abundant common legacies as well as many diverse trajectories in terms of language and culture, these aspects will not be touched upon within this dissertation, except as they relate to the geographic, economic, and transportation-related inquiry herein.)

The term “Latin America” itself is subject to debate (Bulmer-Thomas, 2003, p. 1). For this reason, the definition used here hails from a trade-based perspective best suited to the purposes of this analysis. The definition of Latin America presented here is that used by The United Nations Conference on Trade and Development (UNCTAD, 2010). For simplicity, in this dissertation, the use of terms “Latin America” or “LAC” will both refer to the region which is often commonly termed “Latin America and the Caribbean.” (The exception to this reference will be when referring to other studies, as cited - when such references are made, the definition of Latin America is as defined by the cited study.) This definition is important to set out, as the larger research endeavor prosed here is focused upon Latin America as a world region – especially as it is the region in its entirety which is subject to concerns related to high transport cost.

However, for the purposes of the initial analysis to be performed for this dissertation, the study area encompasses only a subset of the region. Though a very extensive data verification process was conducted for the full Latin American dataset, it is apparent within the data that trade realities and geographies of shipments originating in Central America and the Caribbean differ significantly from those originating in South America. Therefore in the interest of controlling for some extraneous factor, the focus for this research is on South American export shipments (Table 1).

TABLE 1: South America: 2009 economic development level by nation and income groupings for developing economies

DEVELOPING ECONOMIES		
LOW-INCOME	MIDDLE-INCOME	HIGH-INCOME
2009 per capita current GDP below $1,000	2009 per capita current GDP between $1,000 and $4,600	2009 per capita current GDP above $4,600
Guyana [HIP]	Bolivia [LL, HIP] Brazil* Colombia Ecuador** Paraguay [LL] Peru Suriname	Argentina Chile Uruguay Venezuela**
TRANSITION ECONOMIES		
No South American countries designated as transition economies by UNCTAD in 2009.		
DEVELOPED ECONOMIES		
No South American countries designated as developed economies by UNCTAD in 2009.		
Note: (Source: UNCTAD, 2010.) * Indicates major manufactured goods exporter designation by UNCTAD in 2009 **Indicates major petroleum exporter designation by UNCTAD in 2009 LL: Landlocked developing country designation by UNCTAD in 2009 HIP: Heavily indebted poor country designation by UNCTAD in 2009		

As much of the scholarly writing regarding economic trajectory of this world region is in the context of the entire region as an entity, Latin America as a whole will remain the economic context for this discussion.

1.1.1. Latin American Economic Context

As can be seen in Table 1, at present, all South American countries fall into the category of developing economies (UNCTAD, 2010). Yet, as Bulmer-Thomas (2003) points out, the historical expectation was that for Latin America, the removal of colonial rule and its associated monopolies would result in great economic success (p. 4). In the wake of independence (achieved by most Latin American nations in the 1820s) this expectation simply has not been borne out, even today (Bulmer-Thomas, 2003, pp. 2-5). By the beginning of World War I, Latin America was considered "underindustrialized" due to the high cost and inefficiency of industry, caused by governmental neglect, insufficient energy provision, unusually high transport costs, and the necessity of importing machinery (Bulmer-Thomas, p. 396). The impetus for rapid industrialization (and thus rapid urbanization) was created by the Great Depression and World War II, and a program of import substitution began; by 1955, manufacturing replaced agriculture as the greatest contributor to real GDP (Bulmer-Thomas, p. 9).

Many researchers mention Latin America's policy of import substitution during these decades as harmful to its development and integration (Reyes, Schiavo, & Fagiolo, 2010, p. 219) as it resulted in "high-priced, low-quality goods" produced by firms that could not compete globally (Bulmer-Thomas, p. 9). The ruins of import substitution combined with the effects of the oil crises of the 1970s – as well as poor governmental policy-making in an intervention effort during the 1960s and 1970s - eventually led to an increase in external debt, followed by hyperinflation during the 1970s and early 1980s (Bulmer-Thomas, 2003, p. 9; Reyes, Schiavo, & Fagiolo, 2010, p. 218) These factors, combined with the economic crisis of the early 1980s, led to the adoption of policies that in the long term built debt, inflation, and a poor reputation for governmental institutions (Development Centre of the Organisation for Economic Co-operation and Development [OECD], 2009, p. 1).

During this time, the divergence between the world's economic core nations and Latin American nations (both periphery and semi-periphery) increased, as did divergence between the periphery (richer) and semi-periphery (poorer) nations within Latin America itself (Gwynne & Kay, 2004, p. 8). Within Latin America, the extreme inequality that existed after official independence from foreign colonizers was only intensified by the processes of globalization - placing the region among the worst in the world in terms of income distribution, as demonstrated by consistently high Gini coefficients (Bulmer-Thomas, 2003, p. 10). When investigating economic inequality in the region on the basis of individual nations, a wide diversity of economic realities are found, with year-2000 GNP per head ranging from $6,000 to $7,000 in some nations, and $500 in others (Bulmer-Thomas, 2003, p. 11).

However, over the past two decades, Latin America has "moved away from the false sense of security that protectionist policies provided" (Development Centre of the OECD, 2009, p. 1). Policy changes during this time period included privatization of companies previously owned by the state (in industries such as utilities, infrastructure, natural resources/mining, and the financial sector), which contributed to a huge increase in foreign direct investment in the region during the 1990s (Green, 2003, pp.104-110). This also led to reduction of inflation (Green, 2003, p. 113). Government bond sales raised billions (Green, 2003, pp. 94-95), and many other economic policy mechanisms created immediate, large positive effects. Many of these changes, part of the region's "silent revolution" (Green, 2003), led to improvements that later proved dubious.

Capital flows during this decade proved to be fickle (p. 92); privatization was subject to corruption, contract disputes, and eventually, outcry and lowered public opinion of the privatization (pp.104-110); total external debt increased (Green, 2003, p. 95); and foreign direct investment turned out to be overly concentrated in the wrong industries, never leading to large-scale job creation (Green, 2003, p. 95). Although there were success stories leading to longer-term stability, growth, and continued investment among the disappointments, the economic improvements, as always, were experienced unequally among nations of Latin America (Green, 2003, p. 114). Green (2003, p. 117) states: “What appears to be emerging in the early years of the new century is a Latin America of increasing fragmentation and differentiation.” However, as a region, trade openness and trade liberalization had been accomplished to a greater degree than ever before. The region now turned to the question of how to create comparative advantage (Green, 2003, p. 121; Gwynne & Kay, 2004, p. 19).

Many compare Latin America’s economic trajectory to that of the East Asian “Tiger” economies, in an effort to question why Latin America has not managed to achieve the same economic success. Reyes, Schiavo, & Fagiolo (2010) note that in the 1970s, Latin American countries (in their study: Argentina, Brazil, Chile, Mexico, and Venezuela) had a per capita GDP four times higher than “high-performing Asian economies” (China, Indonesia, South Korea, Malaysia, the Philippines, and Thailand) (p. 219). However, the high-performing Asian economies caught up to Latin American countries in the sample via growth rates that averaged 10% between 1970 and 2000, as compared to the latter’s average growth rates of approximately 4%. Reyes, Schiavo, & Fagiolo (2010) suggest through their network analysis that while the volume of trade has increased for both regions, the East Asian economies have increased their number of trading partners, whereas Latin American economies are very concentrated in terms of trading partners. The authors cite Kali et al. (2007), who made the case that the true economic benefits of increasing the number of trading partners through better connection to the world trade network are increased market, wider competition, and increased potential for technological spillovers to occur (p. 225).

In addition, Fay and Morrison (2007) state that the infrastructure gap between Latin America and the Caribbean and the seven East Asian "tiger" economies (Hong Kong, Indonesia, South Korea, Malaysia, Singapore, Taiwan, and Thailand) widened by a "huge margin" between the years 1980 and 1997. East Asian "tiger" economies showed an increase of 48% in advantage in terms of fixed phone lines, 91% for power generation capacity, and 53% for road length over Latin America; the unit of measure was infrastructure stocks per worker (p. 16). In addition, as will be seen in the following pages, economic policy, processes, and regulation *within* Latin American countries themselves pose significant challenges for financing basic infrastructure improvements (Fay & Morrison, 2007).

Indeed, many other substantial trade barriers have been noted. For example, a substantial trade barrier exists in what has been termed the "spaghetti bowl" (Bhagwati, 1995; Mesquita Moreira, Volpe, & Blyde, 2008) of bilateral/multilateral trade agreements between nations within the region and with global trading partners. Efforts in negotiating trade agreements encompassing the entire Western Hemisphere (such as the Free Trade Areas of the Americas agreement) have been unsuccessful (Hornbeck, 2010), and the regulatory and administrative environment caused by the current labyrinth of trade agreements is not conducive to trade growth for a variety of reasons (Mesquita Moreira, Volpe, & Blyde, 2008).

Yet, the region as a whole has been termed resilient (Development Centre of the OECD, 2009); in terms of recovery from the global financial recession of 2008, the Latin

American recovery is second only to that of Asia (World Bank, 2010). A major caveat regarding Latin America's recovery from the world economic recession, however, is the concern that this was accomplished by turning the clock backward on export composition, returning the region to an export model increasingly based upon natural resources ("Latin American," 2010). Certainly export composition is considered to be one of the major contributing factors to challenges in achieving economic integration and growth in the region.

Theories abound in an effort to appropriately frame the region's failure to achieve what is widely believed to be its economic potential. Most of these theories are part of larger theories of globalization and economic development which have been applied or interpreted in terms of this region. Yet as Bulmer-Thomas (2003) states, "No single theory will explain both the intermediate position occupied by Latin America on the scale of world income per head and the differences that have emerged among Latin American countries over time. Yet a theoretical framework is essential if economic history is to be more than mere description" (p. 14). Three of these globalization theories – hyperglobalist, sceptical, and transformationalist (Held et al., 1999; Gwynne & Kay, 2004) – will be presented, along with two associated theories of economic development (structuralism and dependency theory) which hail from time periods before the era of globalization (Gwynne & Kay, 2004).

Sceptical theory does not hold to the idea that the characteristics of the current era that we term "globalization" are unique – but rather that increased global interaction has occurred in other eras (such as the era of industrialization), and that the full integration into a globalized world has not yet arrived. Evidence used to support these elements of sceptical theory include the continuing relevance of nation-states as units/national economies, hence the need to create trading blocs; in addition, adherents to sceptical theory point to the immobile nature of labor (Gwynne & Kay, 2004, p. 6). Another perspective, the transformationalist theory, is that this era of globalization is in fact a completely new occurrence – one for which the outcome cannot be anticipated, the effects of which are unknown, and the new distinctions which it creates between economic entities cannot yet be grasped (Held et al., 1999; Gwynne & Kay, 2004, p. 7).

An economic development theory that existed pre-globalization (Bulmer-Thomas, 2003, p. 13), dependency theory characterizes nations/regions in the world economy as belonging to either the "core"/"center" (referring to developed countries with higher standards of living) or "periphery" (those nations with lower development levels and lower standards of living). This characterization is based upon the inequality of interaction or exchange between countries or regions (Bulmer-Thomas, 2003, p. 13). Street and James (1982, p. 674) note that dependency theory is an outgrowth of structuralism. While structuralism recognizes the economic system as "evolving" and human behavior to be based not just upon utilitarian response, but upon "habitual patterns resulting from cultural conditioning but also capable of intelligent response" (pp.673-4), it also recognizes power structures exhibiting institutional influence upon economic development. Dependency theory takes its own path within structuralism based upon what Street and James term its "exploitative conception of the growth process" (p. 674). Street and James point out, albeit not specifically making use of the term *agency*, that one criticism of this theory is its failure to account for the power and potential of actors (which the authors term "indigenous growth forces," p. 674) to affect such constraints.

Nations/regions of the periphery are theorized to become dependent upon the capital flows and technologies of the core, and are subject to dictation of terms of their internal policies by international economic bodies such as the World Bank and International Monetary Fund (IMF) (Gwynne & Kay, 2004, pp.11-14).

Reyes, Schiavo, and Fagiolo (2010) investigate both the Latin American and East Asian economies based upon this theoretical framework, but investigate both economic regions based upon indicators used to determine depth of integration a nation or region has achieved in the world trade network, conceptualized in terms of a continuum from core to periphery. The authors characterize nations or regions upon this basis not in terms of where productive activities are located but rather based upon the number of trade linkages (p. 218). They find that the group of high-performing Asian economies in their study has increased the number of - and intensity of relationships with - trading partners, thereby increasing its level of integration into the world trade network. Whereas the group of Latin American countries is shown to distribute increases in trade among fewer trading partners over the same time period, and thus has not increased its level of integration into the world trade network. Reyes, Schiavo, and Fagiolo (2010) also find evidence for the core-periphery structure of the world trade network.

Gwynne and Kay (2004) point out that all of the theories presented, with the exception of hyperglobalist/neoliberal theory presented below, tend to present evidence of increasing global divergence, or economic inequality – rather than global convergence (or reduction of inequality). Indeed, they demonstrate through the use of World Bank data from 1978, 1990, and 2001, a long-term trend of increasing divergence in any two-group ratio chosen from three groups of nations: 1. the six core economies (USA, Japan, Germany, UK, France, Italy); 2. the six Latin American economies with highest 2001 GNP per capita (Argentina, Uruguay, Chile, Brazil, Mexico, and Venezuela); and 3. the six Latin American economies with lowest recorded 2001 GNP per capita (Haiti, Honduras, Bolivia, Guatemala, Ecuador, and El Salvador) (p. 8). Whereas these theories point to inequality between and within nations, hyperglobalist/neoliberal theory focuses these issues through the lens of restructuring and differentiating roles within a global division of labor.

Hyperglobalist theories have as their foundation the notion that globalization is a force unlike any experienced previously, in that it causes the unit of the nation-state to become economically obsolete (Gwynne & Kay, 2004). Instead the transnational corporation (TNC) is the nation-less primary unit, operating globally without governmental constraint (Gwynne & Kay, 2004). Hyperglobalist theory encompasses both neoliberal ideology and a divergent neo-Marxist line of thought (Gwynne & Kay, 2004, p. 6).

Neoliberal theory is the basis of much of the literature already presented here surrounding Latin American economic development and integration. This theory as applied to Latin America views government policy/intervention as an obstacle to Latin American economic development, growth, and integration (Bulmer-Thomas, 2003, p. 14). There have been debates about the extent to which this is true, as many make the case that the East Asian “tiger” nations had intentionally put in place state policies which propelled them toward the economic success to which Latin America’s own economic trajectory is often compared; it is also recognized by some that these policies, while essential, must be “selective and temporary” (Gwynne & Kay, 2004, p. 12; Kay; 2002;

Gwynne, 1990). In general, neoliberal theory expounds market-driven over state-driven growth, as set forth by global economic regulating bodies such as the World Bank and the IMF (Gwynne & Kay, 2004). Gwynne and Kay (2004) explain that this theory focuses on five points of economic policy reform: "fiscal management, macro-economic stability, privatization of state firms, labour markets and trade liberalization" (p. 15). The authors refer to the state adoption of these economic policies during the 1990s as the "neoliberal turn" (p. 20) in Latin America, though they explain that this turn has occurred at different times and rates in individual nations. Though the "neoliberal turn" has usually been associated with a preceding turn to democratic governance, the economic policies can also be adhered to within nations under authoritarian rule (Gwynne & Kay, 2004).

One branch of this theory is focused upon "institutional and structural" characteristics of Latin America (Bulmer-Thomas, 2003, p. 13). It is interesting that Bulmer-Thomas refutes this theory (and frames it separately from neoliberal theory), stating, "The institutional and structural landscape inherited from the colonial period was not homogenous and has changed significantly over time" (Bulmer-Thomas, 2003, p. 13). Yet this author also claims that it was the inequitable land distribution, post-colonialism, which laid and cemented the pattern of inequality that still exists today (p. 10). Indeed, as is demonstrated in accompanying sections of this paper, many authors have pointed out that it is such institutional and structural characteristics that are a major constraint to Latin America's economic success and growth.

The differences between the two divergent principles which fall under hyperglobalist theory are eloquently summed up by Gwynne and Kay (2004, p. 6): "The neoliberals tend to see all countries (rich and poor) benefiting from globalization, although within each country significant restructuring will take place. In contrast, the neo-Marxists believe that global capitalism creates and reinforces structural patterns of inequality both between and within countries." The neo-Marxist perspective is in opposition to the goals and processes used by organizations such as the World Bank and International Monetary Fund, and neo-Marxism characterizes their neoliberal basis, policies, and processes differently.

As an example of this theory playing out "on the ground" politically and economically, Jameson (2010) discusses the success of the indigenous movement in Ecuador, which has in some respects successfully influenced economic policy and brought greater autonomy for the indigenous groups via the concept of a "plurinational state" (p. 64). The CONAIE (Confederation of Indigenous Nationalities of Ecuador) was formed in 1986 to unite indigenous groups from different geographic regions of Ecuador into this plurinational group. The group has weathered fragmentation and political chaos as well as built strategic victories for its anti-capitalist position against an "extractivist neoliberal model" (p. 70) of mining and oil operations, Washington-based free-trade agreements, and U.S. military presence. The group continues to have staying power and cohesion as a group of states separate from federal government, ostensibly fueled by an original guiding principle of an "ecologically planned communitarian economy" (CONAIE, 1994; Jameson, 2010, p. 66), though in recent years has been the subject of governmental attempts to thwart its operations and goals.

However, it is notable that a major challenge consistently faced by these smaller cooperatives and collectives is still *the lack of access to infrastructure* – particularly communication and transportation infrastructure (Calderón & Servén, 2002; U.N.

ECLAC, 2010). If actors both large (TNCs) and small (as in the cases of the preponderance of small firms operating in Latin America), and the agricultural cooperatives previously mentioned, experience the same transportation cost and accessibility constraints, then this also suggests that further investigation may be warranted. In fact, it appears that the effect of transport-related trade impedances may actually weigh more heavily upon smaller firms. For example, in terms of food exports from LAC, Schwartz et al. (2009, p. 9) state:

> "From the perspective of a firm, domestic logistics costs in LAC may be the largest single cost element of the final price of a good. While there are important variances by sub-region and type of firm, LAC's logistics costs are most sensitive to the size of the firm. For small mills, markets and retailers of foods in towns and secondary cities around Latin America, domestic logistics costs can total over 42 percent of the price of a firms' sales. By comparison, larger firms spend between 15 and 18 percent of sales on logistics. This is driven by such factors as lack [of] access to warehousing, storage and transfer facilities and the quality of the infrastructure and trucking services that link rural markets, smaller towns and secondary cities to large production and consumption centers."

Therefore, this dissertation represents a contribution toward investigating such trade impedances, and in particular the theory of "infrastructure gap" (Calderón & Servén, 2002; Fay & Morrison, 2007, p. 16). Chisholm and O'Sullivan (1973) explain that in order to accurately determine the necessity for, and appropriate location of, "large and continuing investment" (p. 1) in transportation infrastructure, traffic volume between nodes must be accurately forecasted, and appropriate models for doing so must be created and calibrated. In terms of examining theories regarding infrastructure gap, the first question in such an exploration would be to determine if there is sufficient evidence that such a gap exists in order to justify further study.

As it will be shown in the following sections, evidence has been accumulating suggesting that transportation infrastructure in Latin America may require improvement in order to reduce disproportionate freight transportation costs, both seaward and landside. Therefore this dissertation research carries out the first steps in an initial analysis investigating transportation infrastructure gap in Latin America, by determining whether disproportionately high trade impedance is observed for the domestic landside portion of U.S.-bound South American export commodity flows - and delineating its extent, locations of occurrence, and spatial pattern (for the total trade flow, and by export commodity composition). These first steps are undertaken with the intention of investigating, in a future extension of this analysis, the association of trade impedance with transportation, political-economic/trade policy, and sectoral/export composition factors. However, in order to investigate these commodity flows and their associated trade impedances, it must first be determined what comprises the commodity flows.

1.1.2. Latin American Global Export Trade

In terms of goods exports originating in LAC by value (Fig. 1), it is immediately clear that the U.S. has been and continues to be the most important trading partner for LAC as a region, as it receives the largest share of LAC exports (37.3% in 2009), as compared to other world regions. However, it should be noted that this is the lowest share of LAC exports by value claimed by the U.S. for all years displayed here. Although in absolute terms the monetary figure has increased approximately five-fold from 1990 ($47,765 mil.) to 2009 ($252,602 mil.), the next-lowest U.S. share of LAC goods exports by value was in 1990 (39.3%). LAC goods exports to the U.S. peaked in terms of value in 2008 at $355,366 mil.

A much greater decline in share was claimed by the EU, which fell from 24.1%, or nearly one-quarter, of LAC goods exports by value in 1990 ($29,953 mil.) to merely 11.8% ($80,174 mil.) in 2009. The EU is now the region accounting for the lowest share of LAC export goods by value.

The greatest increase in terms of share over the time period is by the group of nations termed here the "rest of the world," which grew from 11.6% ($14,103 mil.) share of LAC goods exports by value in 1990 to a 20.0% ($135,358 mil.) share in 2009. Another notable increase is that of the share of export goods trade with Asia, which increased from a share of 10.5% ($12,728 mil.) in 1990 to a share of 13.0% ($87,730 mil.) in 2009 – however, what is not shown in the figure is the rapid growth of the share of China alone (included in the total for Asia). China increased from an extremely low share of LAC export goods by value in 1990 at 0.7% ($875 mil.) to 6.7% ($45,229 mil.) in 2009.

In terms of LAC intra-regional exports, the share of LAC export goods by value has remained fairly stable, from 14.6% ($17,762 mil.) in 1990, to 17.9% ($121,161 mil.) in 2009, with some fluctuation over the time period.

The effect of the 2008 financial crisis is clear in Fig. 1, as the total of LAC goods exports by value grew each year until it reached its peak during this time period of $870,627 mil. in 2008 – after which, in 2009, the total of LAC export goods by value fell to pre-2007 levels.

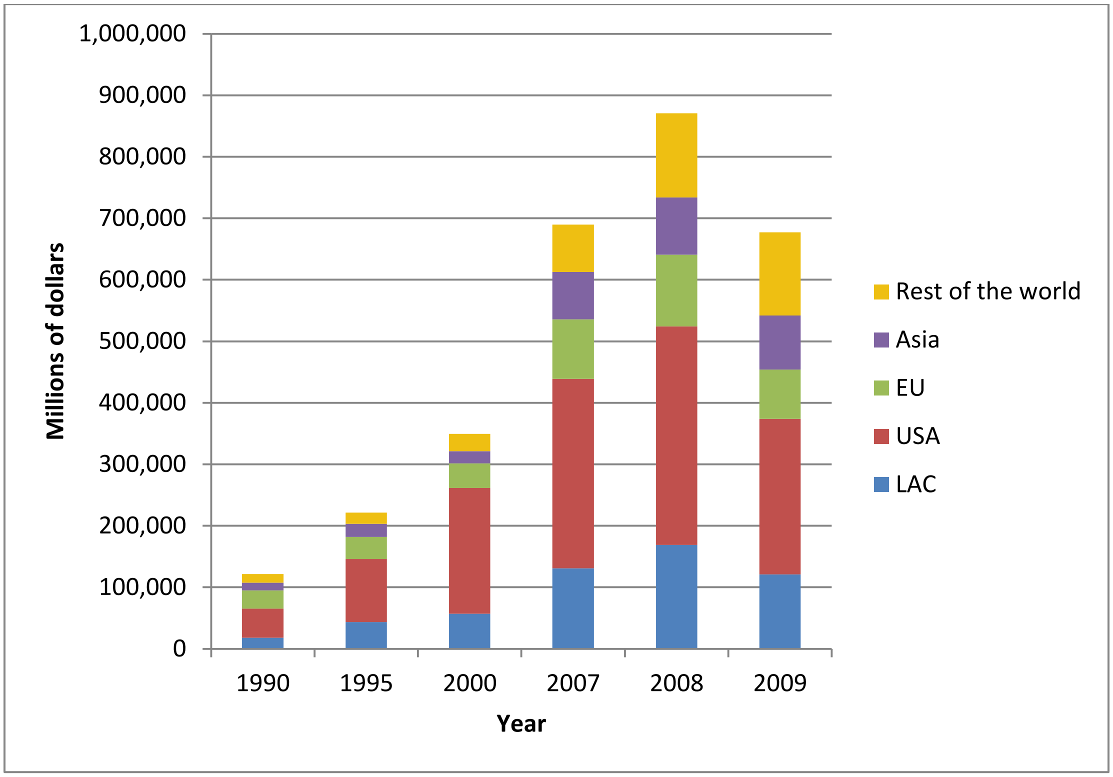


FIGURE 1. Latin America and the Caribbean (LAC): goods exports to selected markets - 1990, 1995, 2000, 2007, 2008, and 2009, in millions of dollars (Source: UNECLAC, 2010.)

In terms of export composition, the general statement remains that "primary commodities still provide the main link with the rest of the world" (Bulmer-Thomas, 2003, p. 8). Although their importance has been decreasing (particularly in Mexico), due to the phenomenon that many of Latin America's manufactured exports are still based upon natural resources, primary commodities still comprised two-thirds of Latin American exports in 2000 (Bulmer-Thomas, 2003, p. 8). Gwynne and Kay (2004, p. 9) state that since the 1950s, it is the smaller Latin American countries that generally still rely on primary commodities, whereas larger, industrialized nations rely on export of labor-intensive manufactured goods or inputs.

Table 2 displays country export trade structure by product group for South American countries. These countries can be typified by their largest percentage of export product categories by value. For example, South American countries with the majority of exports by value being comprised of fuels include Venezuela (92.6% of exports), Ecuador (59.2% of exports), Bolivia (48.8% of exports), and Columbia (38.4% of exports). Brazil is the only South American country with the largest share of exports by value being comprised of manufactured goods (49.7%). The majority of exports of a few South American countries is comprised of food items, including Paraguay (75.8%), Uruguay (55.2%), and Argentina (44.5%). Two countries have their largest percentage share of exports by value comprised of ores, metals, precious stones, and non-monetary gold - Chile (64.7%) and Peru (64.3%). Notable break-outs of manufactured goods type by value include the large share of Brazilian (24.2%) total exports comprised of machinery and transport equipment. This table also displays the absolute value for all exports for 2006, showing that Brazil, at $137.8 billion, far surpasses other South American countries. The lowest absolute value for all exports in 2006 was earned by Paraguay, with total value of exports of $1.9 billion.

Table 3 displays export concentration and diversification indices for South American countries. The two countries with the highest concentration indices, Venezuela (0.911) and Ecuador (0.532), are among those with dominant shares of exports in fuels, and are designated by UNCTAD as major petroleum exporters. The countries with the highest diversification indices are Venezuela (0.844), Peru (0.787), Chile (0.782), Paraguay (0.771), Bolivia (0.760), and Ecuador (0.736).

TABLE 2: Country export trade structure by product group for Latin American countries represented in dataset: 2006 economic development level by nation and income groupings for developing economies*

Country or territory	Year	Total value (millions of dollars)	By main SITC Revision 3 product group (percentage)							
			All food items	Agricultural raw materials	Fuels	Ores, metals, precious stones and non-monetary gold	Manufactured goods	Of which:		
								Chemical products	Machinery and transport equipment	Other manu-factured goods
Bolivia	2006	$ 4,223	14.2	1.6	48.8	25.1	10.2	1.2	3.1	6.0
Brazil	2006	$ 137,806	24.9	3.8	7.7	12.1	49.7	6.7	24.2	18.7
Colombia	2006	$ 24,391	15.8	4.4	38.4	6.2	35.2	8.3	6.2	20.7
Ecuador	2006	$ 12,728	27.0	4.0	59.3	1.0	8.7	1.4	3.7	3.6
Paraguay	2006	$ 1,906	75.8	7.2	0.0	1.1	15.9	3.1	1.2	11.7
Peru	2006	$ 23,765	14.7	1.4	8.0	64.3	11.7	2.0	0.5	9.2
Argentina	2006	$ 46,423	44.5	1.3	14.6	6.5	31.6	8.1	12.7	10.8
Chile	2006	$ 55,881	15.3	5.3	1.9	64.7	10.7	4.4	1.4	4.9
Uruguay	2006	$ 3,952	55.2	9.7	3.5	2.2	29.3	5.7	3.7	19.9
Venezuela	2006	$ 61,385	0.2	0.0	92.6	2.1	5.1	1.1	0.8	3.2

Note: (Source: UNCTAD Handbook of Statistics 2008 – Section 3.1. Country trade structure by product group – Exports.)
*The SITC codes for the product groups shown above are listed in Appendix E.

TABLE 3: Export concentration and diversification indices of Latin American countries represented in dataset: 2006 economic development level by nation and income groupings for developing economies

Region, country or territory	Exports - 2006		
	Number of products exported	Diversification index	Concentration index
WORLD	260	0.000	0.079
DEVELOPING ECONOMIES	260	0.234	0.142
ECONOMIES IN TRANSITION	255	0.584	0.314
DEVELOPED ECONOMIES	260	0.165	0.063
Bolivia	145	0.760	0.403
Brazil	248	0.475	0.092
Colombia	230	0.571	0.207
Ecuador	184	0.736	0.532
Paraguay	123	0.771	0.320
Peru	216	0.787	0.256
Argentina	243	0.561	0.132
Chile	228	0.782	0.402
Uruguay	185	0.671	0.228
Venezuela	174	0.844	0.911
(Source: UNCTAD Handbook of Statistics, 2008 – Section 4.1.1. Export and import concentration and diversification indices of countries and geographical regions.)			

1.1.3. Latin American Export Trade with the U.S.

Gwynne and Kay (2004, p. 10) state that for Latin America, "the key political and economic relationship is that with the United States, the dominant player in the global economic and political system of the early twenty-first century." The authors point out that within Latin America there are countries that may be more or less influenced by U.S. policy and have closer or more distant economic relationships. While it is clear that the U.S. is the most important trading partner for Latin America, it is interesting to investigate where Latin America ranks in terms of share of U.S. goods imports.

Latin America is becoming an increasingly important trading partner for the United States. In terms of total merchandise trade by value, Latin America is not the largest, but is the fastest-growing trading partner of the U.S. (Hornbeck, 2010, p. 13). In 1990, the LAC share of U.S. goods imports by value was 12.9% ($66,549 mil.), which increased to a share of 18.1% ($289,674 mil.) by 2009. While volumes overall increased, there were some slight changes in shares. In 2007, LAC actually tied the EU for share of U.S. goods imports by value at 17.4%, then surpassed the EU's contribution in 2008 and 2009. Both Asia overall and the rest of the world remained stable in their shares of U.S. input goods by value.

Again, the effect of the 2008 financial crisis is clear in Fig. 2, as the total of U.S. goods imports by value grew each year until it reached its peak during this time period of $2,164,234 mil. in 2008 – after which, in 2009, the total of U.S. import goods by value fell to pre-2007 levels.

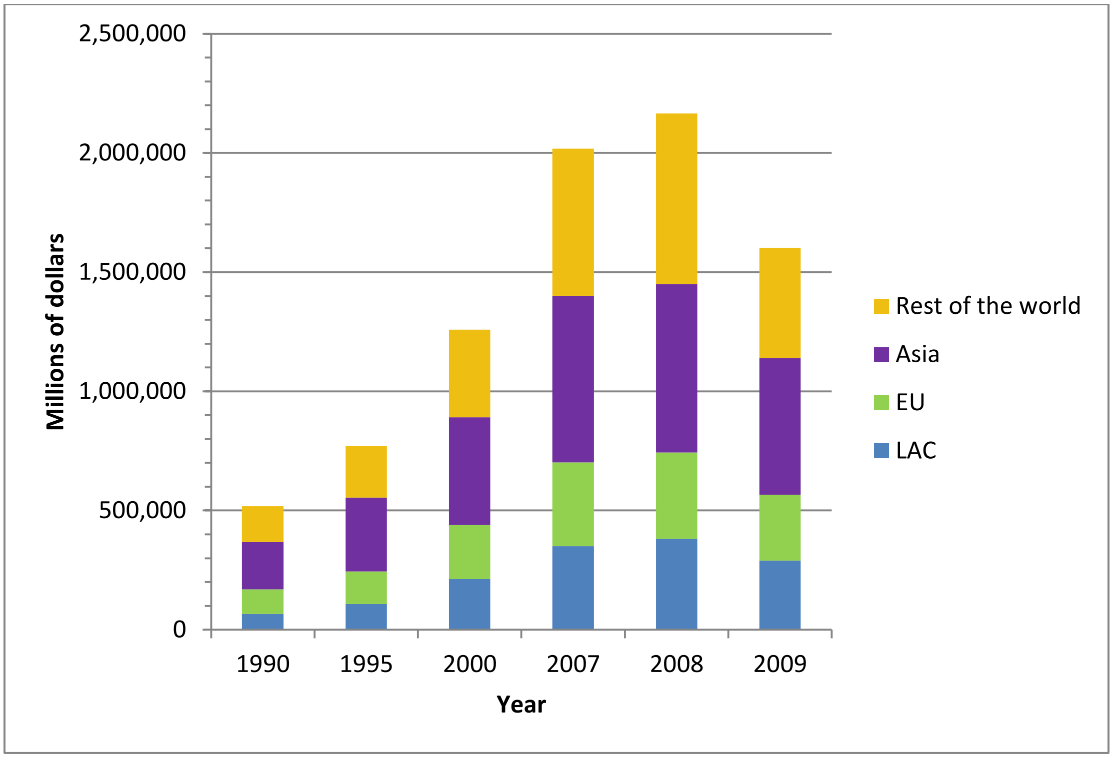


FIGURE 2: USA: Goods Imports from Selected Origins – 1990, 1995, 2000, 2007, 2008, and 2009, in millions of dollars (Source: UNECLAC, 2010.)

1.2. Factors Influencing Trade Impedance

It is clear in the literature that transport costs are highly influential upon trade flows; Limão and Venables (2001, p. 453) have shown that trade flows are highly elastic in terms of transport costs, demonstrating in their study that a 10% increase in transport costs reduces trade volumes by approximately 20%. Thus transport costs are a highly influential factor in trade impedance overall. However, there is extreme complexity in the number and nature of factors which may contribute, both directly and indirectly, to transport costs (Limão and Venables, 2001, p. 470). The impetus for this investigation – World Bank and IDB findings that transport costs are higher between and within Latin

America - was itself suggested to be caused by a collection of factors. Although transportation infrastructure was suggested as the primary transportation-related cause, the potential factors span the gamut from the physical (such as topography, as well as allegations of poorly-maintained roads) to the purely logistical (such as corporate structure and its relationship to individual firms' supply chains).

In addition to those factors of trade impedance which lie primarily within the realm of transportation concerns, there are also noted factors which lie within the realm of trade policy or which are related to the role (sector) of a particular economy; complete separation of these factors into mutually exclusive categories would be an abstraction, as there are many trade impeding factors which may overlap or interact. This is an excellent example as to why the delineation of trade space (relative space from the perspective of the O-D flows) should be explored, as it is in this research. As will be described in Chapter 4, this research suggests beginning with the distance-as-experienced (functional separation) (Plane, 1984) – or "nongeographic" distance (Thill, 2011, p.7) – which, for comparison, is expressed in geographic terms. As conceptualized for trade flows, this functional separation is specifically defined as trade impedance. Working backward from this functional separation to unravel contributors to what has already been experienced may better preserve and indicate the overlaps or interactions of trade impeding factors, than would pre-constructing suites of presumed transport cost variables in an attempt to get at the separation (cost) experienced by trade flows.

Regardless of their interactions or co-occurrences, the complexity of these potentially trade impeding factors must first be organized in such a way that provides structure to the investigation; therefore potentially contributing factors and possible interactions are set in a conceptual framework.

Factors related to transportation cost have been organized within a structure suggested by Notteboom (2001, p. 12) in his "Multi-Level Approach to the Issue of Land Accessibility" (Fig. 3). Notteboom (2001) developed a construct for conceptualizing land accessibility to seaports which is inclusive of the physical, infrastructural, transport-specific, and logistical aspects. This conceptual structure also integrates several ideas which are central to landside (and more generally, all) freight movement: spatial versus functional concerns; the timeframe in which the different levels can respond to change; and that demand for freight transportation is derived from demand for goods (Friedlaender & Spady, 1980). Although Notteboom (2001, p. 12) referred to this structure in terms of land accessibility to seaports in Europe at the turn of the twenty-first century, this structure is helpful in conceptualizing and organizing factors related to transport costs for landside freight movement in any world region (though region-specific versions of this conceptual structure could potentially differ). As well, this framework can be extended to other transaction costs, as will be demonstrated.

In this conceptual model, the author presents the elements of the model as follows: the locational layer refers to "the geographical location of a gateway," specifically in terms of places of economic significance as elements for accessibility of a seaport; the infrastructural layer is rather self-explanatory, including transport links and nodes; the transport layer "involves the physical aspects linked to transport chains" (p. 12) such as transshipment function at nodes; the logistical level involves organizing and integrating the transport chain into the logistic chain (p. 12). The upward arrow demonstrates that each layer gives value to the lower layers, while the downward arrow shows the direction of demand from the higher to the lower levels. The lower or more basic levels are characterized as having a lower responsiveness or adaptability (longer time lag) to changes in the market demand (p. 12). Inefficiencies at any level can reverberate throughout the port area and within the maritime or landside portions of the port-linked flow (p. 14).

Notteboom also contributes other concepts that go hand in hand with this conceptual model. The first is the concept of "revealed accessibility" (p. 11) which dovetails with the concept of functional (inferred) distance (Plane, 1984), to be introduced in Chapter 4. Notteboom defines revealed accessibility as "a particularly appropriate criterion for assessing the market's valuation and satisfaction as regards the quality of the land access to a sea port" (p. 11). This definition stands separate from a measure usually used to measure accessibility – the supply or capacity of transport infrastructure and related services; Notteboom instead terms this measure "intrinsic accessibility" and states that is it representative of "potential or opportunity" (p. 11). However, Notteboom's "revealed accessibility" is the accessibility that is experienced, and evidenced by actual freight flows. Notteboom emphasizes that what is being revealed by these actual flows is demand, valuation, and satisfaction with quality of land access to ports – not just intrinsic accessibility (p. 11).

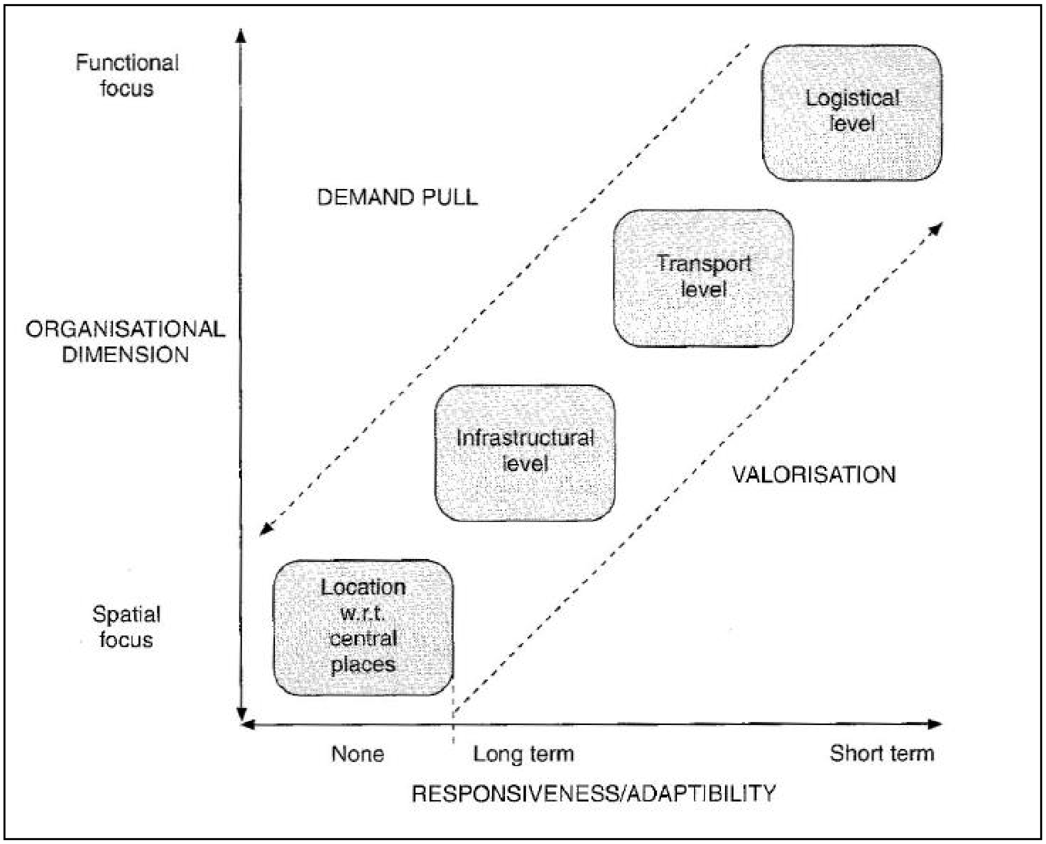


FIGURE 3: Notteboom's multi-level approach to the issue of land accessibility – 2001 (Source: Notteboom, 2001, p. 12.)

Notteboom's framework for investigation of factors potentially influential upon land accessibility to ports is particularly well-suited for structuring the landside transport costs that oppose accessibility. However, there are additional factors which contribute to trade impedance – some which may be defined as trade barriers, rather than transport costs, and some of which are a composite of transport costs and other trade barriers. For this reason, Notteboom's transport-cost specific model is expanded in the context of this dissertation to a broader conceptualization of *trade impedance*, inclusive of sectoral and political/economic as well as transportation-related dimensions. Although this dissertation is situated within a particular concern for transportation-related factors, the factors that comprise trade barriers have been examined more broadly – both in terms of a global scale, as well as a more inclusive context - than Notteboom's land-accessibility related study of Europe.

Whereas Notteboom demonstrates *conceptual* factors underlying trade flows (demand, valorization, temporal adaptability, and spatial versus functional concerns), the WEF illustrates a more detailed representation of *practical* trade facilitation factors and their complex interactions. To this end, Fig. 4 presents The World Economic Forum's (WEF) ten "pillars of enabling trade." It is this detailed look at trade facilitation (or, in the reverse, trade impedance) factors and their interactions which are the focus of this dissertation, while Notteboom's original organizing principles of landside accessibility provide an important framework to build upon.

The WEF ten "pillars of enabling trade" are placed within the four sub-indexes of potential trade barrier factors: formal border barriers (market access), border administration, transport and communications infrastructure, and regulatory and security measures (the business environment) (Lawrence et al., 2008, pp. 6-7). This organization of trade-barrier related variables is used to develop the WEF's Enabling Trade Index (or ETI), produced annually since 2008 in the WEF *Global Enabling Trade Report.* In Fig. 4, these pillars are conceptualized within a spatial and functional framework, in terms of location of occurrence and effect in terms of border crossings.

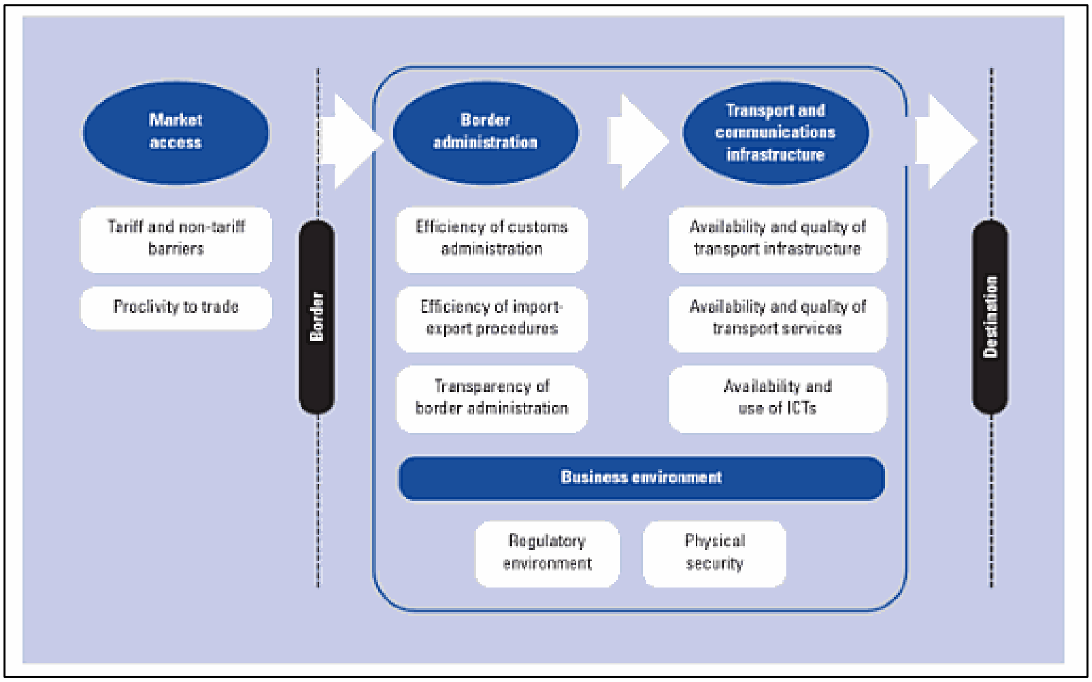


FIGURE 4: Composition of the four sub-indexes of the Enabling Trade Index (ETI) (Source: Lawrence et al., 2008.)

However, it is clear in the graphic that this concept, while including variables relevant to both directions of trade flow, is primarily geared toward the representation of import, not export, operations. Yet the relevant WEF report includes variables that appear to be non-specific as to import or export flow; the report context is "trade facilitation" and for the most part does not discriminate between imports and exports. In the report, Lawrence et al. (2008, p. 8) also state:

> "To keep this project at a feasible scope, we have explicitly focused on the flow of imports into a country in the ETI. Of course, it is important to note that most of these factors stimulate the flow of both exports and imports. Introducing elements explicitly enabling the flow of exports will be the subject of future research in this ongoing project studying obstacles to trade."

However, the 2009 Global Enabling Trade Report specifically addresses feedback that the index should be more explicit in including export-related measures by including country-specific variables related to cost, time, and required documentation to export, as well as tariffs and margin of preference in the destination market. Lawrence, Hanouz, and Moavenzadeh (2009, p. 9) state that, "The single most important change concerns the explicit introduction of the export dimension into the Index."

Although the trade barriers affecting import and export trade are usually discussed as influential without clear separation of directionality of trade flow, specificity in terms of direction of flow is essential - both in conceptualization and model implementation/ variable choice. While some trade impedance factors affect both imports and exports (such as transportation factors), directionality must be considered. For example, a country with a large trade imbalance may experience greater congestion/delay at ports and network links and nodes primarily in one direction of trade flow. Likewise, it can be expected that the directionality of congestion matters significantly, both at ports and within the landside; it is certainly a logistical consideration. In addition, many customs-related factors are related to the directionality of flow. Examples include import tariffs and other protectionist trade policies, and export flow factors, such as inspection or documentation delays for commodities eligible for value added tax (VAT) refund.

Clearly, some of these variables are best investigated during the routing portion of transportation analysis (for example, temporal differences experienced in directional travel over the same infrastructure links and nodes). Yet including directionality in other steps of analysis is necessary for proper identification of trade impedance factors and selection of relevant trade impedance variables. While this concept is simple, it is not always explicit - or even touched upon - in trade literature. While it may not be the intent to exclude the opposite direction of flow, often a discussion of trade impedance factors is presented in such a way that appears to be inclusive of both directions (as in a nation's "index score" apparently applicable to both directions of flow) – or, one direction of flow is not included to the extent of the other in either quantitative measures or related discussions. This may be due to a lack of readily available directional data and/or methodology for constructing trade impedance indexes or variables which are directionally specific.

This leaves a conceptual gap, as it is clear that: (1) in considering the overarching trade impedance factors, some are solely export- or import-specific; and (2) in considering trade impedance factors which apply bi-directionally (such as transportation infrastructure quality factors), the *sub-factors* may be vastly different for each direction of flow. For example, a factor such as customs operations is applicable to both directions of flow – however, a sub-factor such as "Irregular payments in exports and imports" for a particular country (a perception index from the World Economic Forum Executive Opinion Survey included in the Global Enabling Trade Index [Lawrence et al., 2008]) may provide a measure of customs corruption. Yet in attempting to apply this sub-factor to directional trade flow, one might ask: Can this rating truly be applied *equally* to import or export flows, or would that be a misapplication? Is this corruption index primarily related to payments for import inspection/clearance/release (perhaps a more likely scenario) - or rather for overlooking the reporting of fictitious/fraudulent exports to claim drawback? Yet such indices, non-specific as to directionality, are often the best available approximation - or perhaps the only available data - in many cases where such sub-factors must be applied, in order to include the political/economic effects of trade impedance.

However, as trade is an interaction, to conceptually account for both sending and receiving sides of the equation is necessary for accurate and targeted analysis (as well as more accurate and targeted *interpretation* for policy-making informed by trade impedance modeling). Figure 5 presents a framework for considering the export- and import-specific elements of a trade flow, as well as the ways that factors shared by both landsides may have unique manifestations due to the direction of flow. (It should be noted that the particular practical manifestations [displayed in white ovals] are examples only, and not inclusive of all potential manifestations of directionally-specific trade impedance). This figure stands as a directional conceptualization of trade impedance - including broad-based functional trade barriers indicated by the World Economic Forum, within the context of landside accessibility as introduced by Notteboom.

In Fig. 5, Notteboom's original conceptual model, which displays multiple levels of transport-related factors affecting land accessibility to ports, has been re-conceptualized as a multiple-level approach to broader landside trade impedance factors. Three overarching factors contribute to trade impedance in this conceptualization: political-economic factors, sectoral (export composition) factors, and transportation factors. This trade impedance approach, based upon these three dimensions, expands Notteboom's four levels of land accessibility to seven levels of trade impedance, highlighting the levels at which the contributing factors may overlap or interact.

Some trade impeding effects originate solely from these three factors of trade impedance, individually. For example, a direct cause of trade impedance may be that some sectors have inherent trade impedances due to export composition (such as heavy commodities); this trade impedance factor becomes apparent at the export composition/commodity level. Likewise, trade impedance may be purely related to political/economic factors, such as embargoes; this trade impedance factor becomes apparent at the trade policy level. And still other trade impedance factors originate solely from transportation factors, such as congestion; this trade impedance factor becomes apparent at the transportation level.

Yet it is clear that interactions of the three overarching trade impedance factors (political-economic, sectoral, and transportation factors) combine to create trade impedance effects as well; these trade impedances become apparent at different levels. For example, transportation infrastructure, which is obviously among transportation factors, also has a strong relationship with the economic sector of a region. This could be conceptualized both in terms of the world-region (as in sectoral role of a world region in the global economy) or in terms of the sub-national or sub–continental region (in terms of a producing region with well-established pathways from natural resources to point of export, for example). The sectoral role of the economy affects not only freight movement, in terms of vehicles and modes used, but also affects commuting, transit, and other transportation flows, as well as land use - altogether creating the need for specific types and organizations of transportation infrastructure. The interaction of sectoral and transportation trade impedance factors therefore becomes apparent at the transportation infrastructure level.

Likewise, the interaction of transportation and political/economic trade impedance factors becomes apparent at the location level. As location must be considered both in absolute terms (as in distance) and relative terms (as in global political-economic arrangements), this interaction is clear. In absolute terms, distance is overcome by transportation. However, relative location in terms of the global economy is overcome by trade policy. It is clear how the location level can then be conceptualized as the point at which the interaction of transportation and political-economic trade impedance factors becomes apparent.

Customs-related trade impedance has a relationship with both political-economic and sectoral factors. For example, the interaction of political-economic and sectoral factors becomes apparent at the customs level. A particular manifestation may be bribery related to customs commodity classification of an export, enabling the shipper to receive governmental refunds based upon such a classification (drawback).

The logistical level – according to Notteboom, one that deals with supply chain management by operators seeking to move goods through the transport system – encompasses factors related to regulatory (political-economic) and sectoral (or commodity-related) aspects, as well transport system factors. This middle space represents the reality in which firms operate; trade impedance is comprised of the overlap of these three factors. The logistical level is also the confluence between the export trade flow and the import trade flow.

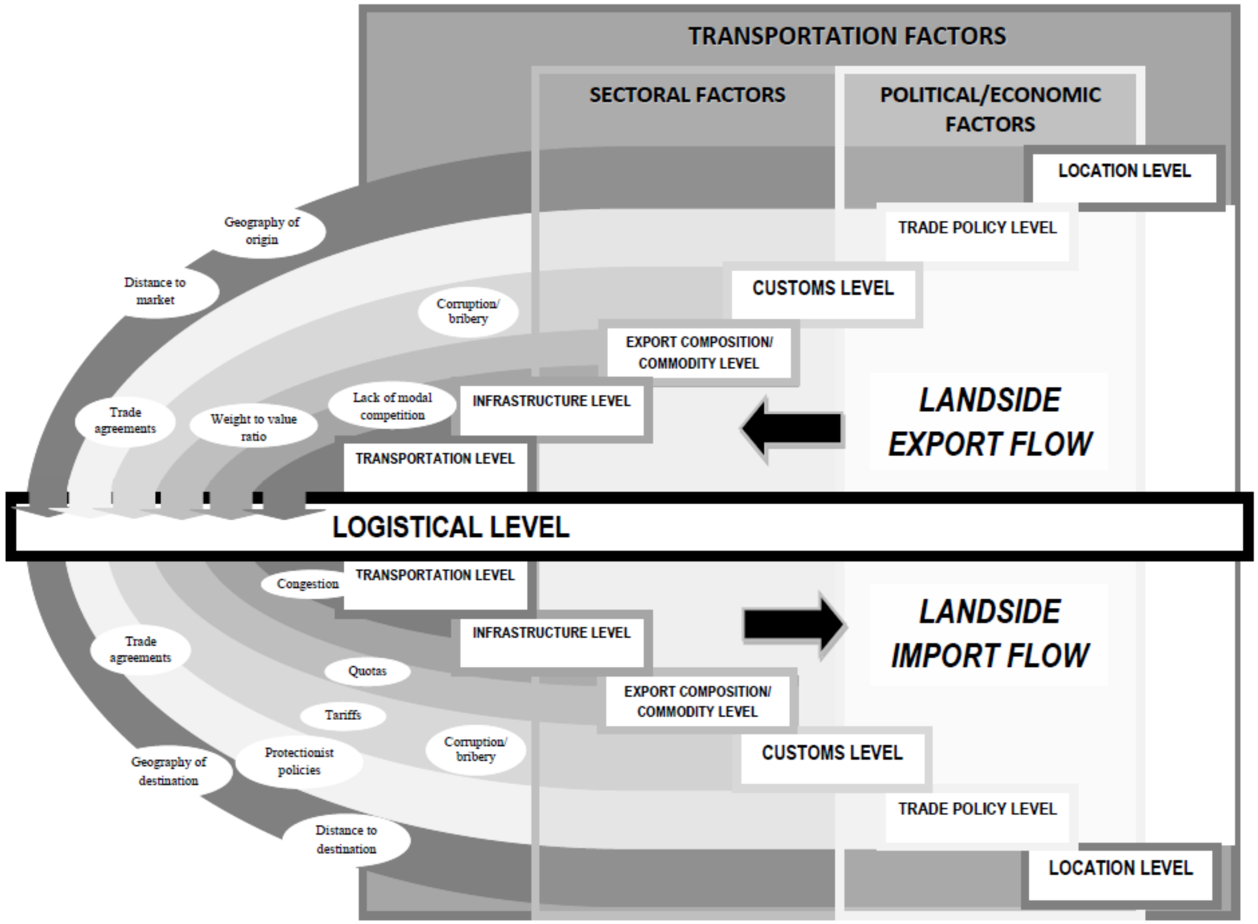


FIGURE 5: Multi-level approach to landside trade impedance (including Notteboom's [2001] Levels of Approach for Landside Accessibility) (Source: Notteboom, 2001, p. 12; Mesquita Moreira, Volpe, & Blyde, 2008; "Latin American," 2010; Lawrence et al., 2008; Graphic by K. Tiller.)

1.3. Trade Impedance In Latin America

The economic history of Latin America is a rich demonstration of the complexity in the number and nature of factors which may contribute, both directly and indirectly, to trade impedance. Transport costs, political/economic factors (trade policy), and sectoral factors (export composition) intertwine to create a complex web of trade impedance factors. Examples are as follows.

Trade policy is inherently linked to export composition when shipments transit borders. Here, trade facilitation efforts such as trade organization membership or trade agreements may also change the border-crossing equation, potentially lowering trade impedance (Lawrence et al., 2008, p. 5; Limão and Venables, 2001, p. 453). The border-crossing equation may also include the relative size of economies participating in the cross-border trade interaction (Stocker, 2011).

Sectoral issues such as export composition are also related to mode and vehicle choice, and therefore are linked to transport infrastructure as the underlying structure of the economy dictates physical structures of land use and transportation.

Even such factors as vendors colluding to raise prices for rendering transportation infrastructure services under government contract - and even the absence of fiscal mechanisms such as consistent, transparent request-for-proposal processes - have been linked to troubled transportation infrastructure, and thereby effects upon transport costs, in the past (Green, 2003). The assessment of world economic regulatory organizations is that Latin American openness to trade growth is impeded by "deficiencies in the quality of infrastructure together with rigid regulatory frameworks and weak freight logistics" (Guerrero, Lucenti, and Galarza, 2009, p. 22).

1.3.1. Transportation Factors in Latin American Context

Lawrence et al. (2008) state that, "There is overwhelming evidence that international trade is very sensitive to distance. The most obvious cause is transportation costs" (p. 4; Disdier and Head, 2008; Hummels, 2007). The degree to which transportation costs are noted as an impediment to Latin American economic growth and integration is unequivocal. In terms of Latin American export trade with the United States, ocean freight rates alone have been found to be much higher than freight rates from far less proximate regions, such as Europe and Asia (Mesquita Moreira, Volpe, & Blyde, 2008; Inter-American Development Bank, 2009, 2010). The Inter-American Development Bank (IDB) states that this in itself is "alarming."

Although differences in ocean freight rates are clearly cause for concern, the reasons that port-to-port freight traffic analyses prevail is not necessarily because these costs are of greater concern than those landside; rather, the IDB states that this is due to the availability of data and the expediency of such studies (Inter-American Development Bank, 2010, p. 24). However, the IDB notes that "counting transport costs only between borders is clearly arbitrary" (Inter-American Development Bank, 2010, p. 24). It is not just seaport and airport efficiency that stands as a trade impediment. Latin American domestic transportation infrastructure across all modes is characterized as embodying this trade barrier for a variety of reasons.

The authors point out that for Latin America, the relationship between transport costs, trade, and productivity is of particular importance - as the region both relies on "transport-intensive" exports (in terms of weight [Hummels, 2001], perishability, time sensitivity, etc.) and that the region "suffers from well-known deficiencies in its infrastructure" (Inter-American Development Bank, 2010, p. 11). The IDB states clearly that transportation infrastructure is a constraint to Latin American trade growth.

For U.S. imports, insurance and freight costs are twice as costly as tariffs (Lawrence et al., 2008, p.4). In addition, distance is associated with other generalized (non-monetary) transportation costs; according to Grossman (1998; Lawrence et al., 2008, p.4), these include: search, preparation and scheduling time; movement; communication; loss; quality deterioration; inventory management; and risk of theft.

The Latin American and Caribbean Economic System (SELA, 2011) provides a different perspective of the connectivity issues faced by the region, by beginning with the basics: first, that simple geography creates connectivity challenges. For example, the Darién Plug at the border of Panama and Colombia prevents southward expansion of the Pan-American Highway; the mountainous, riverine, and jungle lands of South America "fragment" and prevent connections throughout the continent; and practical challenges of geography separate the Caribbean from the rest of Latin America. Additionally, as Limao and Venables (2001, p. 453) state: "Poor infrastructure accounts for 40 percent of predicted transport costs for coastal countries and up to 60 percent for landlocked countries." However, SELA acknowledges the need for infrastructure improvement, while making the case that several important infrastructure project initiatives are already addressing these needs.

Figure 6 displays the basic transportation infrastructure of Latin America at a small scale. Brief discussions regarding each mode are detailed in the sections to follow; however, intensive focus upon describing each factor occurring in the study area will be reserved for the Association Step (Step 4) in a future extension of this analysis.

FIGURE 6: Latin American transportation infrastructure

1.3.1.1. Landside Transportation Factors in Latin America: Roads

In a review of the region's transportation infrastructure, SELA (2011, p. 4) notes Latin America's particular mode-specific challenges. The percentage of paved roads in Latin America is 16%, which is far below the world average of 57% paved roads. In addition, road coverage in Latin America (156 km per 1,000 km^2) is far less than the global average (241 km per 1,000 km^2) (SELA, 2011, p. 4). As well, road maintenance in Latin America is known to be deficient; distance and road conditions are said to combine to increase transport costs by 8% to 19% (SELA, 2011, p. 4).

Fay and Morrison (2007) state plainly: "The quality of Latin America's roads are generally poor" (p. 86). They demonstrate through World Bank data for selected Latin American countries that the percentage of national or regional roads in good condition (for which such assessment data were available) appears to lie in the range of 30% or less. In the South American subset, one country is the exception - Argentina (along with the Central American country of Guatemala) (p. 87). The quality of roads assessment is displayed in Fig. 7. The percentage of both classes of Argentina's roads considered to be "in good condition" lies at or near 80%, whereas just 45% of its regional roads merit this label (p. 87). No date is given for this assessment data, however, based upon context, it is believed to be 2003 (based upon the World Banks's 2005 World Development Indicators); there is no explanation as to what is considered to be "good" condition.

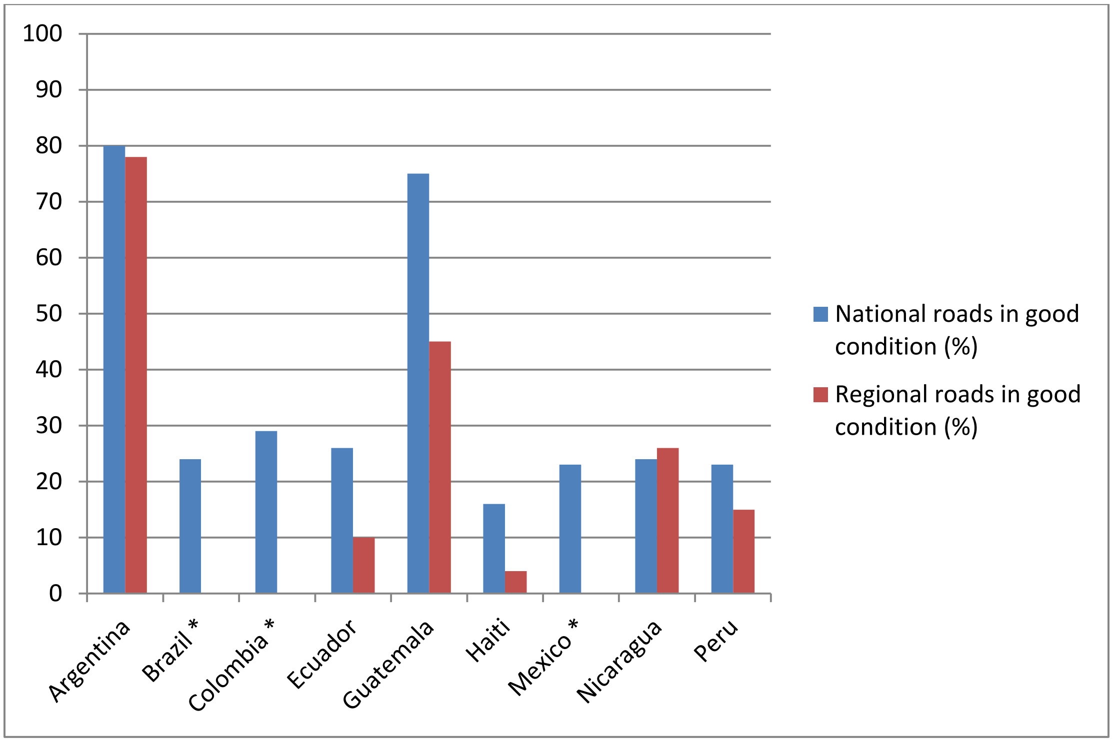


FIGURE 7: Selected Latin American and Caribbean countries: quality of roads – based on government assessments (Year not provided) (Source: Fay & Morrison, 2007; World Bank data.) (Note: *Regional road quality not available for countries indicated.)

In addition, Fay and Morrison note that road density barely increased between the years 1985 and 2001 for Latin America and the Caribbean, whereas for middle-income countries[1] (and also for Korea), road density increased during that time period (2007, p. 85). The authors also cite Calderón and Servén (2004) in finding that for 2000/2001 (the latest year of data available), Latin America and the Caribbean have much lower road densities than do middle-income countries, both when measured by kilometers of road

[1] Regarding the definition for middle-income countries, Calderón and Servén (2004) make backward references to data sources used in their previous work; the original data reference appears to be in Calderón and Servén (2002), which references the International Road Federation (IRF) World Road Statistics as the data source. For the IRF definition, data compilation including the years referenced is available at cost from the following: http://www.irfnet.org/statistics.php.

per 1,000 square kilometers (36 for LAC; 82 for middle-income countries) and per 10,000 workers (30 for LAC; 50 for middle-income countries) (Fay & Morrison, 2007, p. 86). For these reasons, road density is considered to be an important factor to investigate for its influence upon trade impedance in the region.

In their report on private infrastructure investment in Latin America, Mia, Estrada, & Geiger (2007) detail their development of an Infrastructure Private Investment Attractiveness Index, or IPIAI, which encompasses many aspects of the economic, business, and regulatory climate for private investment in transportation infrastructure – as well as the current (to date of publication) state of transportation infrastructure. The background and perception that the "Infrastructure Quality Gap Index" (IQGI) scores provide is useful in developing country-specific context, in conjunction with other sources. The IQGI is based upon hard data and perception scores, with individual scores on 12 variables including: ground (road), port, air transport, and supporting infrastructures. The transportation-related variables contributing to the IQGI generally rely upon perception data from the World Economic Forum Executive Opinion Survey, 2006 – with the exception of the "paved roads" variable (collected from the CIA World Factbook, February 2007) and the "customs clearance" variable (collected from the World Bank, February 2002). (Other variables included in the IQGI, regarding air transport and power infrastructure, hailed from these same sources - in addition to World Bank World Development Indicators, 2006, as well as private firm Booz Allen Hamilton Inc.)

The "gap" in the IQGI refers to the difference between each country assessed and Germany, which the authors identify as being the top-ranked country in terms of infrastructure for the Global Competitiveness Index (GCI) for 2006-2007 (p. 14; Lopez-Claros et al., 2006). Gap scores range from zero to seven, with a score of zero indicating no need for additional infrastructure investment and "world-class levels of infrastructure development" (Mia, Estrada, & Geiger [2007], p. 14). The scores have been re-ranked from the original here, in that a country with the smallest quality gap is ranked first. Although all Latin American countries included in the IQGI are presented here, for brevity, only those for the research subset - South America - will be discussed.

In terms of the 12 Latin American countries included in the IQGI scoring, comparisons are made to Chile, which had the best performance – a gap of 1.4 - on the IQGI index (inclusive of variables related to transportation infrastructure). The authors state that Chile has "the most developed and best quality infrastructure network" among the 12 countries in the study (with scores nearly matching Germany's in terms of port infrastructure - a gap of 0.1) (p. 22). However, road infrastructure quality (a gap of 3.2) was still assessed as needing improvement (p. 22). There exists a large difference between the IQGI scores of Chile and all other nations in Latin America.

Among 12 countries of the region included in the IQGI index, only three fared better than Argentina, as it was among the lowest overall "infrastructure quality gap" scores in the region (3.8) due to its "already fairly developed infrastructure network" (Mia, Estrada, & Geiger, 2007, p. 18). The authors stated that Argentina may fare better than other Latin American nations in terms of transportation infrastructure, due to the timing of its transportation infrastructure investments - which occurred prior to the global economic crisis (however it was felt that the economic crisis would soon catch up to these gains in the form of need for infrastructure investment) (p. 18). It is interesting to note that this primarily perception-based index of road quality - which grants Argentina a fairly high score (3.9 on a scale from one to seven) - does not seem to agree with that presented by Fay and Morrison above (Fig. 7) which claims that national and regional roads approximate 80% in good condition. This may be an artifact of the timing of data, as the Fay and Morrison (2007) data is assumed to hail from 2003, whereas perception data for the IQGI hails from 2006. It may be that the predicted infrastructure investment needs became more apparent from 2003 through 2006, driving down perceived road quality, though the question still remains.

Uruguay ranks sixth in terms of overall infrastructure quality gap among the countries in the sample; its overall IQGI score is 4.1. The road infrastructure quality gap for Uruguay is rather high at 4.1.

Brazil ranks eighth among the 12 Latin American countries in terms of IQGI, with an overall infrastructure quality score of 4.4. The authors state that this “betrays the relatively poor state of its infrastructure and the need for further investment”; they discuss the Brazilian transportation infrastructure by noting special concerns regarding “shortcomings” in road infrastructure quality (Mia, Estrada, & Geiger, 2007, p. 21). Indeed, Brazil’s road quality gap is 5.1. An IRAP (International Road Assessment Programme) (2011) case study of Brazil’s increasing use of the roundabout as a traffic control device in São Paulo says of Brazil’s road infrastructure (Fuller, 2011, p. 2): “Many of its (often unpaved) roads are in poor condition, heavily congested, and also badly maintained.” The author notes as well that the roads are “well used” due to the lack of modal competition with rail. Fuller also notes that freight is hauled by drivers who are not legally obligated to take breaks (p. 2). In addition, Fuller notes that annual road casualties are higher than in the U.S. – though the U.S. has ten times the number of vehicles on the roads (p. 2).

Among the countries in the sample, Venezuela has a poor overall score in terms of IQGI (4.5) and ranks ninth among the 12 countries. In what the authors term a "significant quality gap with Germany" (p. 32), Venezuela has a large quality gap in terms of roads (4.2).

Colombia, ranking tenth among the countries in the sample, has a high infrastructure quality gap, with an IQGI score of 4.9. In particular, the quality of Colombia's roads demonstrates an opportunity for improvement, with a large quality gap of 4.7 (Mia, Estrada, & Geiger, 2007, p. 23).

Peru scores the second worst in the sample on the IQGI with a 5.5 overall infrastructure quality gap. Road infrastructure in Peru demonstrates a large gap (4.7) (pp. 28-29).

Bolivia has the highest infrastructure quality gap in the region, with a much higher overall infrastructure quality gap score than any other country in the sample. Bolivia's overall infrastructure quality gap score is 6.7 on a scale of one to seven. Bolivia also has the highest road quality gap, with a score of 5.3 (p. 20). Mia, Estrada, & Geiger (2007, p. 20), explain that the economic conditions related to debt and corruption (tax evasion), as well as other issues, have made it difficult for the government to afford the infrastructure investments to improve the situation.

Fig. 8 demonstrates visually the road and port scores on the IQGI, to enable comparison of the scores of the 12 Latin American countries in the sample.

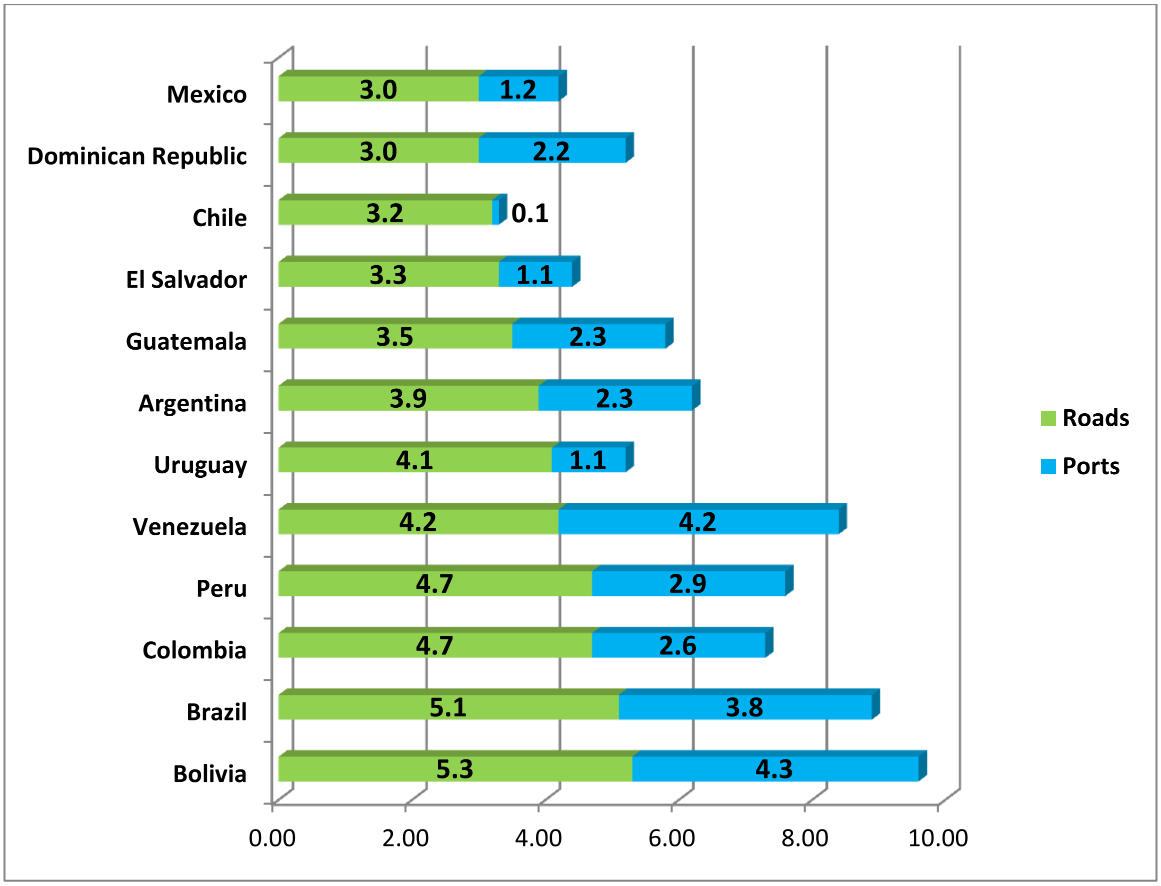


FIGURE 8: Infrastructure Quality Gap Index (IQGI) scores of 12 Latin American countries: roads and ports. Index scores for each mode range from zero (0) to seven (7). (Based upon data from 2006/2007 [Roads], 2002/2006 [Ports]. Chart by author. Source: Mia, Estrada, & Geiger, 2007, p. 34. Data sources used by Mia, Estrada, & Geiger, 2007: World Economic Forum, 2006; CIA World Factbook, 2007; Clark, Dollar, & Micco, 2002/World Bank, 2002.)

1.3.1.2. Landside Transportation Factors in Latin America: Rail and Inland Waterways

Lack of modal competition in Latin America must also be investigated for its effect in terms of transport costs. SELA states that multimodalism needs more promotion in the region (p. 4). World and hemispheric economic authorities agree: “This limited road network with its relatively poor quality is clearly inadequate to meet LAC’s growing demand for cargo transportation, a problem exacerbated by the lack of domestic intermodal competition" (IDB/World Bank/UNECLAC, 2010, p. 12).

While road quality continues to be a challenge, this challenge is all the more dire in the context of Latin American rail. As SELA states: "There is a glaring shortage of railways – only 0.2% of South America's intra-regional trade is carried out by train, compared to 40% by land" (p. 4). Since the era of privatization, rail service has decreased as companies shut down routes associated with loss (Fay & Morrison, p. 16). Fuller (2011, p. 2) states that Brazil's rail network comprises just 29,295 km of track, compared to the road network's almost two million kilometers of highway. Lack of modal competition appears to be an issue for inland waterway options, as well. In Brazil, inland waterways are underdeveloped and underused compared to their potential (Mia, Estrada, & Geiger, 2007, p. 21). A regulatory agency for Brazil's inland waterways (ANTAQ) was not developed until 2002 (p. 21).

1.3.1.3. Landside Transportation Factors in Latin America: Ports

In addition to issues of road quality, capacity, and availability, and lack of modal competition (in terms of rail/waterways) in areas where such modes should otherwise be available, port efficiency and other port issues (such as liner competition) are also discussed in the literature as issues potentially contributing to trade impedance as well. The Inter-American Development Bank (IDB) states that ports in the region of Latin America and the Caribbean are among the least efficient globally (Mesquita Moreira, Volpe, & Blyde, 2008; Inter-American Development Bank, 2009, 2010). Port congestion (Notteboom, 2001, p. 14) and urban congestion (Inter-American Development Bank, 2010, p. 26; p. 31; OECD & ECMT, 2007, p. 37, pp. 48-49) may also be contributors to these efficiency issues. Inadequate port investment, regulatory issues, poor road connectivity, and lack of regional cabotage all contribute to regional traffic issues at sea, in conjunction with the growth of global trade (SELA, 2011, p. 4). In addition, SELA notes other seaborne challenges such as inadequate sea transport between island nations and the Latin American landmass (SELA, 2011, p. 19). Also related to competition effects upon transport cost – and encompassed within Notteboom's "Location (with respect to central places) Level" of land access to ports - is number of maritime lines at port of export. Greater competition among carriers reduces both monopoly power and transport costs (Clark, Micco, & Dollar, 2004, p. 423; Mesquita Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11).

As demonstrated in Fig. 8 above, Mia, Estrada, & Geiger (2007) state that Chile has "the most developed and best quality infrastructure network" among the 12 countries in the study, with scores nearly matching Germany's in terms of port infrastructure (a gap of 0.1) (p. 22). The authors also note the especially low infrastructure quality gap scores in terms of the ports of Uruguay (1.1), and somewhat low infrastructure quality gap scores are earned by the ports of Argentina (2.3). While Colombia (2.6) and Peru (2.9) have mid-range scores on the port infrastructure quality gap, showing opportunities for improvement, Brazil (3.8), Venezuela (4.2), and Bolivia (4.3) have higher port infrastructure quality gap scores. Mia, Estrada, & Geiger (2007, p. 21) also discuss Brazil's port capacity issues.

Figure 9 demonstrates the Latin American ports which figure most prominently in terms of total TEU throughput for the year of the dataset (2006). This chart displays the

TEU throughput for the top 25 Latin American ports in terms of exports to, and imports from, all nations. This chart also provides global context via the *world* port rank of each Latin American port displayed (in terms of total TEU for 2006). In Fig. 9, for example, the port of Santos, Brazil, while clearly ranked first in terms of 2006 TEU throughput among Latin American ports, is listed as "Santos, Brazil – 37," indicating that it is ranked 37th in the world by total TEU throughput for 2006. Also featuring prominently is Buenos Aires, Argentina (fourth in the region/61st in the world) at 1,624,077 TEU. This gives a sense of the throughput level of ports of the Latin American region as compared to ports globally. Many Latin American ports in Central America and the Caribbean serve an important transshipment function, which informs the rankings here.

It is notable that sectoral aspects play a role here as well. In a previous study of port specialization and concentration by Thill, Tiller, and Kashiha (2010) using the Brazilian portion of the dataset, Santos was found to be the least specialized among 14 Brazilian container ports, in terms of the proportions of commodities exported to the U.S. for the time period of the dataset. Upon the use of hierarchical cluster analysis (Ward's criterion) to identify port clusters based upon their export commodity specialization, Santos formed its own single-port cluster typified by the export of manufactured and processed products.

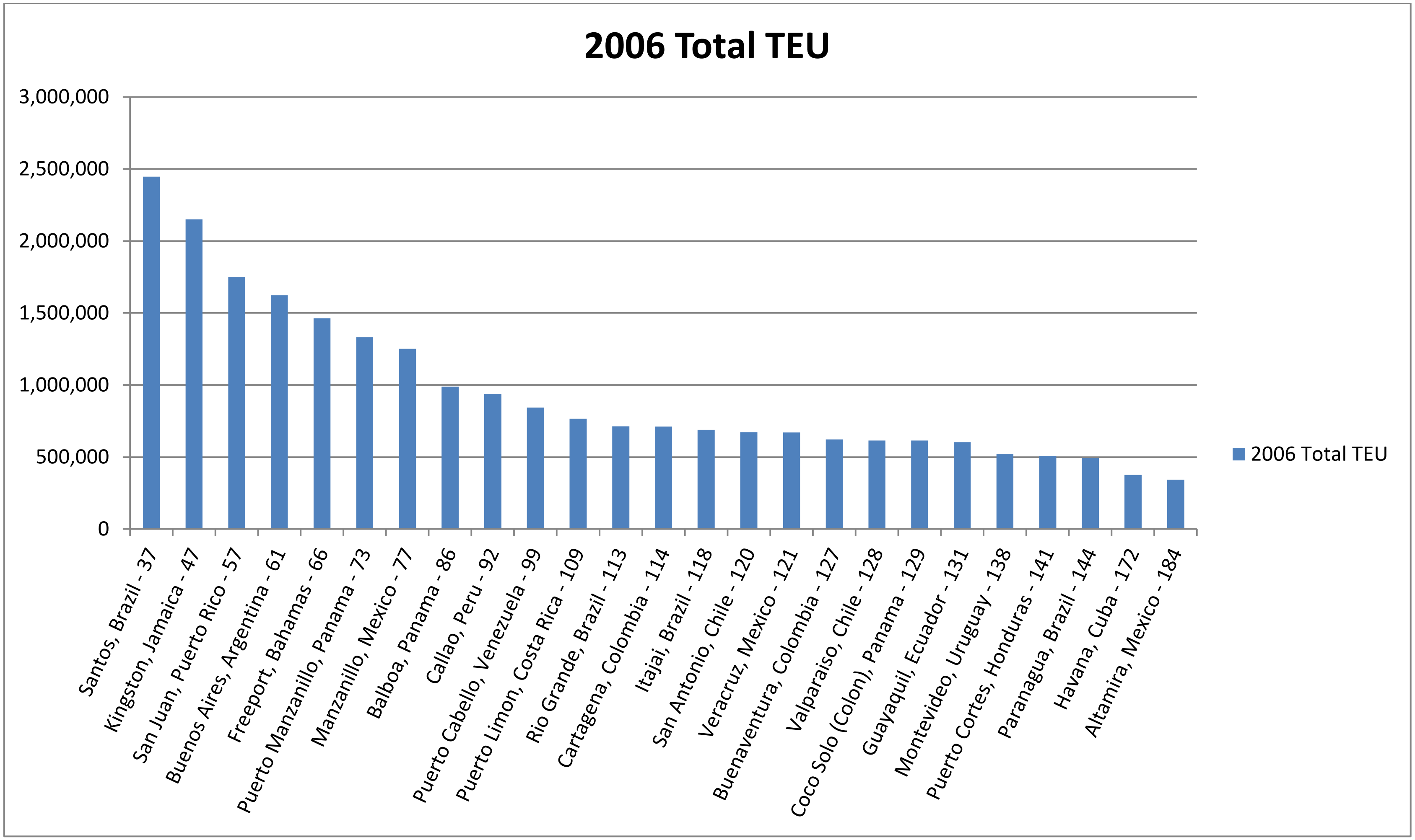


FIGURE 9: Top 25 Latin American ports of export by total shipment volume (TEU) for 2006 – by port name/2006 world port rank by total volume (TEU) (Source: Containerisation International, 2009.)

1.3.2. Political/Economic Factors in Latin America

Another potential trade impedance factor is presented by World Bank and IDB literature regarding the trade barriers inherent in the "spaghetti bowl" (Bhagwati, 1995; Mesquita Moreira, Volpe, & Blyde, 2008) of bilateral/multilateral trade policies and the finding by Limao and Venables (2001) of the importance of shared borders. Limao and Venables (2001, p. 455) suggest that trade transiting through a country other than that where it originated is subject to border delays; transport coordination problems; uncertainty; delays creating higher insurance costs; direct charges by the transit country; trans-shipment costs (p. 460); and transportation network integration problems (p. 460). Thus, clearly, the issues with "logistical costs" are not purely transport-related; SELA notes that customs delays increase transport costs between 4% and 12%, and that border crossings are associated with restrictions related to issuing documents, exchanging and managing information, and monitoring (p. 4). SELA takes a clear position that both the "hardware" or physical infrastructure (including supporting infrastructure such as for electricity and communications), as well as the "software" of regional trade agreements – and the "collective actions" of trade-related regulation and operations such as customs - must all be improved (p. 17).

Customs clearance delays in Latin America have been shown to be some of the worst in the world, behind only the regions of East and South Africa and West Africa (Clark, Micco, & Dollar, 2004, p. 425). Latin America also has some of the highest handling charges at the ports (Clark, Micco, & Dollar, 2004, p. 425). When port issues interact with sectoral and trade policy/customs issues at the logistical level, the effect upon trade facilitation can be daunting - but the returns for improvement may be large.

For example, according to the IDB/World Bank/UNECLAC (2010, p. 18), "for the Port of Santos…reduction in customs processing time by four days would have the effect of as much as a 16 percent reduction in the total logistics cost…" As Fig. 9 above demonstrates, this example is an important one, as the port of Santos claims the largest percentage of Latin American total TEU throughout for 2006. The confluence of this IDB/World Bank/UNECLAC statistic and the identification of customs-related factors as a top threat to trade growth in the region (named so by World Bank/IDB) is a demonstration of how significant a factor these "software" issues may be – as such, or in conjunction with transportation infrastructure and trade policy issues.

Ferreira, Engelschalk, and Mayville (2006, p. 369) state that there are three roles of customs: to collect revenue, efficiently facilitate cross-border trade, and to prevent the trafficking of contraband. However, clearly there is also a fourth purpose of customs, and this is to carry out the practicalities of implementing trade policy, demonstrated in operations such as classification, enforcement of quotas and specific policies, etc. In terms of trade policy, trade barriers that manifest at borders may include the implementation of protectionist trade policies or tariffs and human/physical infrastructure (Lawrence et al., 2008, p. 3). These border-crossing trade costs are detrimental precisely "because they limit trade across borders" (p. 3).

An important quote from the Global Enabling Trade Report 2008 (Lawrence et al., 2008) summarizes the many customs issues that may impede trade (p. 45):

> "Today, it is recognized that clearance processes by customs and other agencies are among the most important and problematic links in the global supply chain. High costs and administrative difficulties associated with outdated and excessively bureaucratic border clearance processes are now cited as more serious barriers to trade than tariffs. Inefficient border

> processing systems, procedures, and infrastructure result in high transaction costs; cause long delays in the clearance of imports, exports, and transit goods; and present significant opportunities for administrative corruption."

Corruption, which exists in all countries by different means and to different degrees (Thachuk, 2005, p. 149), is often viewed as being mitigated by neoliberal reform in Latin American countries. But Manzetti and Blake (1996) beg to differ, stating that "instead, market reforms will engender changes in the modalities of corruption" (p. 670). Thachuk (2005) suggests that political and economic changeovers to capitalism and democracy bring less, not greater, order which gives the corruption more "opportunity and maneuverability" (p. 144). This is echoed by Berríos (2010). Thachuk terms the specific type of corruption that is manifested by bribery payouts to customs officials to be "clienteleism," in which the individual officials who are paid to operate within the *framework* of governmental system make *themselves* the government by adopting clients to garner their favorable decisions.

These new modalities of corruption affect not just customs operations, but many other aspects of trade impedance as well – including seemingly innocuous activities such as road maintenance. One example provided by Manzetti and Blake (1996, p. 679) is that of highway privatization in Argentina, in which contracts were awarded for 10,000 km without Congress receiving any information about content, bidders, and awardees; later it came to light that payment for road repairs was 2.5 times what had been communicated, that the payment was transferred before repairs were ever made - and that the government officials in charge of the privatization bids were on the payroll of the contract awardees. This eventually led to the appointment of a new minister of the economy in 1991 and an "overhaul of the entire privatization framework" by the new minister and World Bank, with renegotiated contracts for highway and airline privatization. (This highlights how the refrain of "transportation infrastructure issues" in Latin America may actually be due to the interaction of policy and other issues, which may become *apparent* at the transportation infrastructure level.)

Thachuk writes, "Criminal and terrorist groups depend on unimpeded cross-border movements, and so border guards, customs officers, and immigration personnel are notable targets of corruption" (p. 147). At first glance, corrupting customs officials with bribery in order to traffick contraband may not appear to have a lot to do with the movement of legitimate goods. However, the constant presence of this influence (Thachuk specifically references the FARC, Fuerzas Armadas Revolucionarias de Colombia) reflects one concept that Manzetti and Blake forward, which is that high corruption is more likely to occur in countries in which corruption occurs so widely that that is "accepted and tolerated" (p. 666). Though this may be a harsh assessment – perhaps the corruption is more so unsurprising or expected than it is truly tolerated – it is clear that if customs officials are regularly corrupted in the movement of contraband, customs corruption and bribery is already present in the same system which moves legitimate goods.

This is an immense challenge. Ferreira, Engelschalk, and Mayville (2006, p. 368) explain that "the potential for corruption cannot be overstated" as customs holds a monopoly on the movement of goods, and is usually one of the largest governmental agencies for a country, due to the office network and staffing required (Lane, 1998). An overall lack of transparency (Ferreira, Engelschalk, and Mayville, 2006, p. 374) and a situation in which lax human resources management policies and procedures allow applicants to “buy” their positions with the express intent of taking bribes (p. 374) can both affect customs operations overall. In addition, customs work is often conducted in remote areas with a low number of staff, and the potential for collusion and intimidation among customs staff is high (Ferreira, Engelschalk, and Mayville, 2006, p. 369).

However, according to Ferreira, Engelschalk, and Mayville, UNCTAD and the World Bank (1996) take the position that it is neither the omnipresence of corruption in customs nor the payments required which impact trade facilitation the most, but rather the delays caused by “unnecessarily complicated procedures,” causing shippers to have to use bribes to speed the movement of their goods (2006, p. 372). These payments may become a regular cost of business for shippers in certain time-sensitive industries (Ferreira, Engelschalk, and Mayville, 2006, p. 369; p. 371).

Thus corruption creates unpredictability and inefficiency in operations for the shipper and “raises the costs of cross-border trade” (p. 371). Some typical methods of corruption that may affect exports include reduction of service hours for clearance or processing (p. 372), unnecessary administrative steps (p. 374), misclassification or the threat of misclassification (p. 373; McLinden, 2005, p. 69), creation of administrative barriers, arbitrary embargo, and soliciting payment for expediting goods or documents (p. 374; McLinden, 2005, p. 69). For imports, these are applicable as well - and in addition imports may face import clearance delays and longer physical inspections (p. 371) as well as higher risk of corrupt activities occurring due to greater direct contact between import agents and customs officials. Also, in terms of exports subject to VAT (value-added tax) and other similar suspense/exemption/drawback schemes, fraudulent input coefficients may be claimed related to production of inputs later benefitting from these schemes, or drawbacks may be claimed for fraudulent or fictitious exports (p. 374; McLinden, 2005, p. 69).

Another important example of the ways in which trade impedance factors interact is discussed by Ferreira, Engelschalk, and Mayville (2006, p. 369): “The diversion of imports to inefficient transportation routes – especially when corruption resides in the nearest customs clearance posts – is yet another example of efficiency loss resulting from corruption” (p. 372). This avoidance of particular customs posts, causing the use of inefficient transportation routes, could also be a factor for exports.

1.3.3. Sectoral Factors in Latin America

Both literature regarding trade issues in the study region, as well as studies of the proposed model to be used (Sect. 4), recognize the importance of commodity composition in terms of trade distance and transport cost. Though this has already been discussed in this document - in terms of commodity composition effects upon, and interactions with, trade impedance in general – specific to the application to this study area, one recent study is particularly notable. The IDB, in a study released in 2013, used broad product categories to investigate transport costs from municipal origin to customs of exit (Mesquita Moreira et al., 2013, pp. 12-14). This is similar, in some respects, to the model used in this dissertation research (as proposed in 2011), in which landside origin point to port of exit O-D flows are segmented by commodity category. In the recent IDB study, the product dimension and transportability were considered (the authors cite Hummels, 2001) due to the likely differences in impact of transport costs by broad product category and country; manufacturing, agriculture, and mining product categories were defined (p. 13). The differences in transport costs were investigated for impact upon level of exports, and displayed by these broad product categories and countries. Differences in impacts are clear. One of the countries more with varied transport costs by product category was Brazil, for which a 1.0% ad valorem reduction in transport costs can increase exports by approximately 1.0% in mining exports – but exports would be increased by nearly 4.0% in in manufacturing, and over 5.0% in agriculture. Although there are potential methodological criticisms which can be made here (Sect. 8), this example demonstrates the importance that the IDB places upon investigation of transport costs by sectors/product groups/commodity categories, defined by transportability.

Empirical evidence found in support of commodity segmentation for trade analysis in Latin America is an important element motivating the approach adopted for the model proposed for this research (Sect. 4). In the use of spatial interaction models for trade analysis, distance may be used as a proxy for the measurement of trade costs and their portion due to transport costs. However, van Bergeijk and Brakman (2010, p. 14) state that even if distance can be directly measured, some studies show it may be inadequate to represent transport costs (here Limao and Venables, 2001, and Combes and Lafourcade 2005 are cited), as transport costs differ between commodities. Though there may also be differences in the geography of production, this is a reference to the logistics cost differences based upon intrinsic characteristics of commodities. As noted above, more recently in Mesquita Moreira et al. (2013), these differences are couched in terms of *transportability* of commodities (Hummels, 2001).

Thus together, the trade, study area, and methodological literature suggest classifying trade by commodity type in the context of transportability.

JUSTIFICATION FOR FURTHER STUDY

2.1 . Problem Statement

Mesquita Moreira, Volpe, and Blyde (2008, p. 138) claim that transport costs greatly influence Latin American U.S.-bound export trade volumes and diversification – far more so than tariffs. Mesquita Moreira, Volpe, and Blyde (2008, p. 139) also give country- and commodity-specific examples for the region, demonstrating that "an inefficient transport network hurts a country's trade." The IDB lists three reasons for higher transport costs in Latin America: lack of modal competition, urban congestion, and most especially, poor road quality and condition (Inter-American Development Bank, 2010, p. 26; p. 31). Trade costs in customs delays are also mentioned (Mesquita Moreira, Volpe, & Blyde, 2008; World Bank, 2009; Inter-American Development Bank, 2010, p. 30). As stated by Mia, Estrada, and Geiger of the World Economic Forum (2007, p. 5):

> "… the participants at the World Economic Forum on Latin America in São Paulo in April 2006 identified poor infrastructure as the major economic hindrance for the region's ability to compete globally and as one of the priority areas in which the Forum needed to explore alternatives and catalyse actions to overcome the current shortcomings."

While the efficiency of Latin American transportation infrastructure is widely criticized as a trade barrier, the criticism is vague in that it has not been subjected to the types of multi-scalar trade and transportation analyses necessary to fully investigate Latin American trade impedance, which may definitively lead to particular infrastructure improvements or policy recommendations. Rather, blanket statements are used to indicate that infrastructure improvements are needed across all modes and throughout the region.

Therefore this research lays the groundwork for focused further analysis of sub-continental regions in Latin America, as an initial launch of a larger research endeavor. This dissertation aims to narrow the overly general notion that transport costs are higher in Latin America, a criticism which is not treatable by either policy or infrastructure improvements in its current amorphous form. Evidence for the existence and levels of disproportionate trade impedance, a more specific reference for the location of its occurrence, and a profile of its spatial pattern, can provide a foundation for the future investigation of the associated region-specific trade impedance factors. These steps, which outline a larger overall research program encompassing multiple scales of analysis, can begin to clarify and specify this overarching claim.

As shown in Fig. 1, in terms of goods by value, the U.S. has been and continues to be the most important trading partner for Latin America and the Caribbean as a region. Due to the primary importance of the U.S. as a trading partner, and specific reports of disproportionate transport costs in terms of U.S.-bound export shipments from the region, accordingly this initial research will be conducted in terms of U.S.-bound export shipments. For this first foray into the use of a unique methodology (Sect. 4), conceptual approach to the study area (Sect. 4), and dataset (Sect. 5), the extent will be narrowed to the sub-region of South America. (This will set aside some of the specific questions related to the trade flows of Central America, the Caribbean, and Mexico, as identified in Sect. 1, to be examined in greater detail after the research methodology is further established via analysis in this initial sub-region.) A focus upon this sub-region is supported by the literature. As Hornbeck (2010, p. 5) states:

> “Still, many economists believe that lowering barriers to U.S. exports and guaranteeing market access may generate long-term trade and investment opportunities, which in turn could lead to higher growth in productivity and output, with both producer and consumer benefits. Similarly, the prospect for even greater access [to] the large U.S. market presents attractive opportunities for South American countries, as well.”

In order to recommend transportation infrastructure, trade policy, and other related policy improvements to better facilitate Latin American trade growth and economic integration, first, basic but valuable facts must be established regarding the trade impedance widely held to be disproportionate to the distance of the shipments traveling upon the Latin American domestic transportation infrastructure. Therefore, this research aims to investigate the existence, extent, location, and spatial distribution of disproportionate trade impedance within the landside of the South American sub-region (between shipment origin point and port of export), for one month of U.S.-bound waterborne containerized export shipments. These crucial first steps precede future discrimination of factors highly associated with export trade impedance, in a proposed extension to this study.

RESEARCH QUESTIONS AND HYPOTHESES

This dissertation seeks to answer the following research questions in terms of containerized U.S.-bound South American export trade. Each stated research question below is followed by the respective hypothesis formulated in response.

1. For the total dataset, how are the observed O-D flows by volume associated with actual distance between origins and destinations in the domestic/landside portion of the shipment path within the South American region? For the total dataset, are expected/inferred distance values disproportionately high as compared to actual physical distance in the domestic/landside portion of the shipment path within the South American region?

   Null Hypothesis (H_0): Expected/inferred distances are not disproportionately high as compared to actual physical distance for O-D flows in the domestic/landside portion of the shipment path within South America.

   Alternative Hypothesis (H_a): Expected/inferred distances are disproportionately high as compared to actual physical distance for O-D flows in the domestic/landside portion of the shipment path within South America.

2. For each commodity category, how are the observed O-D flows by volume associated with actual physical distance between origins and destinations in the domestic/landside portion of the shipment path within the South American region? For each commodity category, are expected/inferred distance values

disproportionately high as compared to actual physical distance for O-D flows in the domestic/landside portion of the shipment path within the South American region?

Null Hypothesis (H_0): Expected/inferred distance values are not disproportionately high as compared to actual physical distance for specific commodity categories in the domestic/landside portion of the shipment path within South America.

Alternative Hypothesis (H_a): Expected/inferred distance values are disproportionately high as compared to actual physical distance for specific commodity categories in the domestic/landside portion of the shipment path within South America.

Specifically, due to the higher transport costs for heavy commodities (Hummels, 2001) and higher costs of insurance for higher-value goods (Hummels, 2009, p. 19) – at least in part, a function of travel time – it is anticipated that expected/inferred distance values are disproportionately high as compared to actual physical distance for the following commodity categories:

1) Heavy commodities/minimally processed natural resources
2) High-value manufactured products

It is anticipated that expected/inferred distance values are proportionate to actual physical distance for Perishable products, Non-perishable time-sensitive products, Manufactured products, and Miscellaneous/unable to be determined commodity categories.

3. For both the set of shipment origin points and the set of destination (port of export) points, what is the spatial pattern in the distribution of expected-to-actual distance (EtA) ratios?

   Null Hypothesis (H_0): The spatial pattern of the distribution of expected-to-actual distance (EtA) ratios, both by origins and destinations, is random.

   Alternative Hypothesis (H_a): The spatial pattern of the distribution of expected-to-actual distance (EtA) ratios, both by origins and destinations, is not random.

METHODOLOGY AND DESIGN

4.1. Background: Spatial Interaction Modeling

In searching for methods to investigate the relationship of trade and economic development/economic integration, it is natural to turn to spatial interaction modeling. Fotheringham and O'Kelly (1989, pp. 1-2) explain that a crucial characteristic of a society with a higher level of economic development is increased interaction in terms of trade and communication. This is due to increased specialization in economic activity, which leads society to become more dependent upon other cities, towns, regions, or nations to provide the wide variety of products and services for which the society was once self-sufficient. In other words, societies require less external contact and trade when economic activity is focused upon provision of an entire array of general needs, as opposed to when economic activity focuses time and material resources instead upon specialized product(s) or service(s) for external trade (p. 2). The authors go on to explain that this is why more economically advanced societies are marked by the provision of infrastructure, such as roads and communication networks; these are conduits, enabling a society to communicate and trade its specialized products or services externally (Fotheringham and O'Kelly, 1989, p. 2). Ullman (1980a, p. 15) states this eloquently in saying, "In order to have interaction between two areas there must be a demand in one and a supply in the other…Specific complementarity is required before interchange takes place. Complementarity is thus the first factor in an interaction system; because it makes possible the establishment of transport routes."

Chisholm and O'Sullivan (1973) point out that the investment necessary to build transportation infrastructure is extremely far-reaching, and involves far more investment than just that of the huge outlay that must come from the transportation sector alone (p. 5). The authors state this in order to point out the crucial nature of such studies of trade flow and transportation cost, as these are often used as preliminary investigations for infrastructure improvement needs assessments (p. 1).

The backbone of spatial interaction modeling is the concept of spatial interaction, a generic term for exchange processes which occur over space (Haynes & Fotheringham, 1984, p. 10). Haynes and Fotheringham (1984) state that spatial interaction can include any of the following examples of types of exchange: "…shopping, migrating, commuting, distributing, collecting, vacationing, and communicating…" (p. 10). Spatial interaction models are characterized as mathematical models used to "analyze and forecast" these exchanges or spatial interactions (p. 9). The genesis of this type of model within the discipline, and the origin of the phrase "spatial interaction," is generally attributed to geographer Edward Ullman in 1954 (Ullman, 1980a, p. 13; Johnston et al., 2000, p. 776-777). However, it is often noted that it was French geographers in the early 1900s who conceptualized the construct initially, and the term "géographie de circulation" is attributed to these unnamed geographers (Johnston et al., 2000, p. 776-777). Ullman himself made reference to both the term "circulation" as well as its French origin; he states that it encompasses the broader concept of the study of transportation (Ullman, 1980b, p. 29). In fact, Ullman noted at the time that the ideas he presented were not new but that his aim was to make them explicit, stating further that the idea of spatial interaction applied to many different types of movements (Ullman, 1980b).

The study of the spatial distribution of freight traffic - and its relationship to economic development - draws from the main traditions of microeconomic perspective (partial equilibrium location theory) and macroeconomic perspective (general equilibrium theory) (Chisholm & O'Sullivan, 1973, p. 6). Partial equilibrium theory approaches the issues of freight distribution from a firm-level perspective, and generally these are approached as firm profit maximization problems; some methods include optimization methods utilized in location theory and potential accessibility models (pp. 6-9). General equilibrium theory, which focuses on the relationships between a wider set of economic entities and variables, encompasses the methods of linear programming, input-output modeling, and gravity models, among others (pp. 12-18).

Chisholm and O'Sullivan (1973) indicate that gravity/spatial interaction models are borrowed from those used in the physical sciences, referring to gravity models as a method of "social physics" (p. 14, 1973). Indeed, many within geography and related disciplines cite that the origins of the model lie in Newtonian physics (Roy & Thill, 2004; Haynes & Fotheringham, 1984). However, Tobler (1981), in his description of building a vector field (a method borrowed from hydrodynamics for use in a migration model), makes a statement that could be more generally applied to the whole of spatial interaction modeling in geography: "…here there is no borrowing of physical concepts for use in social sciences" (p. 6). In this, he affirms that it is the basic mathematical models and visualization tools, not concepts, which are borrowed from physical sciences (p. 6). In other words, such models, and the particularities of the models, are used not so much as a conceptual analogy, but rather as a mathematical description of movement or interaction which happens to be useful in different systems and within different disciplines.

Indeed, Wilson stated that the term gravity model was "something of a misnomer" when he purported to offer a new framework and "more fruitful analogies" through the lens of statistical mechanics (1971, p. 1). As Roy and Thill point out (2004, p. 339), physics moved on to the theories of Einstein, and the field continues to forge new theories only possible due to the newest technological advances in testing and experimentation. Therefore, in our hearkening back to the Newtonian model of gravity, we borrow a static snapshot of one particular moment in the mathematics used in physics - not the entire field itself or even the concept itself. Haynes and Fotheringham (1984) acknowledged the dual role of this borrowing by stating that, "The Newtonian physics approach to social science analysis continues to influence and cloud the application and interpretation of the model" (p. 17).

Thus the overarching analogy guiding the development of the gravity model was the subject of debate, in terms of the formulation best used and how the different model methodologies are described mathematically. A major division in this debate concerns the use of a form that is multiplicative focused upon an entropy-based objective hailing from information science (Wilson, 1971; Batty, 1976; Batty & Sidkar, 1982; Fotheringham & O'Kelly, 1989) versus the additive form used by Tobler focused upon a quadratic objective (1981, 1983; Dorigo & Tobler, 1983; Fotheringham & O'Kelly, 1989). Ledent (1985) presents "a more functional form" (p. 261) of the doubly constrained model which can then be modified for either objective; (Fotheringham and O'Kelly (1989, p. 39) characterize Ledent's explanation as providing a continuum of models. As Ledent states the superiority of either form "is simply an empirical matter" (p. 261).

Wilson's often misconstrued (1971, p. 7) explanation of the basis for a more useful gravity model framework employs principles of statistical mechanics. He explains the new framework thusly: "This method assumes, in effect, that we are calculating the statistical averages of behavior, possibly subject to various constraints, of all the individuals in the system" (1971, p. 7). Wilson effectively lays the Newtonian physics analogy to rest in favor of his statistical averaging methodology by citing Carrothers (1956), who pointed out that while the behavior of individual molecules is not usually predictable, en masse their behavior is subject to prediction upon the basis of mathematical probability – and likewise the behavior of individual actors may not be mathematically predictable, but conceivably the behavior of groups may be (1971, p. 8). Wilson also notes that there is no inherent validity to an entropy maximizing model in comparison to a model hailing from another framework, and that models of phenomena amenable to an entropy maximizing framework could also be discovered via other methods or frameworks (p. 8).

The original concept of spatial interaction was comprised of three characteristics: complementarity ("a function of areal differentiation promoting spatial interaction"), intervening opportunities between places, and distance measured in real terms (Ullman, 1980a, p. 18). Fotheringham and O'Kelly (1989) touch upon these in describing the concept of spatial interaction as being, at its core, a "trade-off" between the attraction of the destination, or the potential advantage gained by making the interaction, and the cost of traversing the separation between origin and destination (p. 1). In more general terms of measures of separation, this is conceptualized as impedance (Roy & Thill, 2004).

Haynes and Fotheringham (1984) explain that gravity models separate these notions into two interrelated concepts: scale impacts and distance impacts (p. 11). Scale impacts are associated with Ullman's concept of "site" (Ullman, 1980a, p. 13). Scale impacts are those that affect the level of interaction between origins and destinations via the level of attractivity of origins and/or destinations, usually indicated by a "mass term" (Wilson, 1970, p. 2) or variable such as population (hence the use of the term scale) (p. 11). Mass terms are usually presented as some measure of attraction or an interaction volume such as inflow or outflow (Wilson, 1971, p. 13). Distance impacts are those that affect the level of interaction between origins and destinations via the amount of separation between origins and destinations; this separation is usually indicated by a variable such as distance (p. 11). Distance impacts are associated with Ullman's concept of "situation," or "horizontal relationship" (Ullman, 1980a, p. 13). However, there are other means of defining both scale and distance elements depending upon the variables and parameters included within a particular gravity model formulation, as will be detailed in this passage.

Gravity models are the most often-used type of spatial interaction model (Wilson, 1971, p. 1). Ullman described gravity models as empirical formulas describing a relationship in which "interaction between two places is directly proportionate to the product of the populations or some other measures of volume of two places and inversely proportionate to the distance (or distance to some exponent) apart of the two areas" (1980a, p. 18). The gravity model is widely used as an appropriate method for investigating interactions such as commodity flows (Wilson, 1971; Hay, 1979; Ullman, 1980a) and migration (Tobler, 1981; Plane, 1984). Haynes and Fotheringham (1984) state that its applications include but are not limited to “transportation, marketing, retailing, and urban analysis” (p. 11) and also encompass spatial interactions of communication, information, and population (in the form of migrants or travelers, for example).

Thus, the basic gravity model given by Haynes and Fotheringham (1984, p. 16) is:

$$T_{ij} = k\, P_i^{\lambda}\, P_j^{\alpha}\, d_{ij}^{\beta} \qquad (1)$$

In which:

(1) T_{ij} represents interaction between any pair of origins (i) and destinations (j);

(2) P represents the scale of each origin or destination (in the authors’ example, population);

(3) The exponents λ and α represent the allowance for other variables in addition to the mass variable of P which influence the amount of interaction (positive exponents indicate that as the mass variable increases, interaction increases – while larger exponents represent a greater effect of the mass variable upon the amount of interaction);

(4) d_{ij} represents the measure of separation between origin and destination (usually - and in the authors’ example - distance);

(5) The exponent β represents the proportionality of the distance impact. As Haynes and Fotheringham state: “Even though distance will always have a negative influence on interaction, in some cases it may be more negative than in others. An exponent on the distance variable, d_{ij}, allows us to represent this variability” (p. 12). In this formulation of the model, β is expected to be negative. This term

could also be included in the model formulation by the function exponential (βd_{ij}); and

(6) The constant k represents a scale parameter.

Haynes and Fotheringham (1984) then present an expanded general model, based upon the inclusion of the concepts of interaction potential (estimation of a set of flows in terms of "single centered subtotal of flows…referred to as a center's interaction potential" [p. 18]), separation of origin and destination attributes, inclusions of separation measure impacts other than those of distance, and inclusion of the spatial structure concept (pp. 18-19):

$$T_{ij} = f(V_i, W_j, S_{ij}) \qquad (2)$$

In which:

(1) T_{ij} represents the estimated interaction between any pair of origins (i) and destinations (j);

(2) V_i represents a vector of origin attributes;

(3) W_j represents a vector of destination attributes; and

(4) S_{ij} represents a vector of separation attributes.

Gravity models are characterized by their purposes and capabilities as being either explanatory or predictive, depending upon the type of model used (Fotheringham and O'Kelly, 1989; Chisholm & O'Sullivan, 1973). Fotheringham and O'Kelly (1989) note that the purpose of the attraction-constrained, production-constrained, and unconstrained models is explanatory, whereas the purpose of the doubly-constrained model is predictive of the "pattern of flows" (p. 3), giving no indication as to the propulsive elements of origins and or attractive elements of destinations for a particular type of flow (p. 3). The production-constrained gravity model form is one that includes a known total interaction volume, or outflow, leaving each origin, and estimates the inflow totals into each destination (Haynes & Fotheringham, 1984, pp. 22-23). The attraction-constrained gravity model form is the opposite: the model form includes a known inflow (total interaction volume) into each destination, and estimates the outflow totals leaving each origin (Haynes & Fotheringham, 1984, pp. 23-25). Wilson explains that the general formulation of the gravity model is always “interaction = balancing factors x mass term x mass term x distance function” - pointing out that what really differentiates members of the family of models is the nature of their constraints and balancing factors (1971, p. 3).

Of particular interest to the subject of this dissertation research is the doubly-constrained model (also termed the production-attraction constrained gravity model [Haynes & Fotheringham, 1984, p. 25]). With the doubly-constrained gravity model, both the outflow totals leaving each origin and the inflow totals into each destination are known (either through observation or accurate estimation) (Haynes & Fotheringham, 1984, p. 24-25). The doubly-constrained model provides a unique capability to use observed flow data to forecast, by generating estimated flows between origins and destinations (Haynes & Fotheringham, 1984, p. 24-25).

Haynes and Fotheringham (1984, p. 25) give the general formulation of the doubly-constrained model as:

$$\hat{T}_{ij} = A_i O_i B_j D_j d_{ij}^{\beta} \qquad (3)$$

where

$$A_i = [\sum_j B_j D_j d_{ij}^{\beta}]^{-1} \quad (4)$$

and

$$B_j = [\sum_i A_i O_i d_{ij}^{\beta}]^{-1} \quad (5)$$

In which:

(1) $\hat{T}_{ij}$ is the estimated interaction between origin i and destination j;

(2) Ai is a balancing factor imposing that the predicted total interaction volume leaving each origin should equal the known value, O_i;

(3) O_i is the known value of interaction volume (outflow) leaving each origin;

(4) B_j is a balancing factor imposing that the predicted total interaction volume leaving each destination should equal the known value, D_j;

(5) D_j is the known value of interaction volume (inflow) into each destination;

(6) d_{ij} represents the measure of separation between origin and destination (usually - and in the authors' example - distance);

(7) The exponent β represents the proportionality of the distance impact – see Equation 1.

Haynes and Fotheringham (1984, p. 25) note that the balancing factors (A_i and B_j) are functions of each other, solved by initially setting all B_j's equal to 1.0 and deriving estimates of A_i and B_j iteratively until no changes are observed. These balancing factors are also referred to as "spatial competition terms" (Plane, 1984, p. 245), and Plane explains that the balancing factors encompass the competing opportunity and spatial structure effects (p. 247), previously mentioned here as expansions of the basic gravity model (Haynes & Fotheringham, 1984, pp. 18-19). They are used as a means of cross-accounting to ensure that the estimated inflow and outflow totals are equal to the known inflow and outflow totals (Haynes & Fotheringham, 1984, pp. 22-24; Plane, 1984, p. 245).

In his exploration of U.S. interstate migration, Plane (1984) used a version of the doubly constrained gravity model in which observed physical migration distances are an input used to derive “revealed or inferred distances” (p. 244) via a unique calibration of the model. The inferred distances that Plane derived are differentiated from functional distances yielded by Markov chain modeling, in that the latter outputs combined summary measures of both attributal (related to site or origin/destination mass variables) and associational (related to situation, or separation) properties of the origin and destination set (p. 244-5; Ullman, 1980a, p. 13). However, Plane’s use of the doubly constrained gravity model can be calibrated in such a way to provide a summary measure that distinguishes the associational properties of the origin and destination set – which Plane later refers to as a measure of functional separation, “inferred distances,” or “generalized migration ‘cost’” (p. 246).

Plane (1984, p. 246) states of “inferred distances”:

> “These distances, when derived from a doubly constrained model, are composite measures of the associational properties of the regions in a migration system, because attributal properties have been represented by the actual inflows and outflows that are used as the model’s mass variables. Any variable that affects the general attractiveness of a region should influence total inflow and outflow, while any variable measuring

> the specific attractiveness of a region relative to another region should influence an inferred distance."

With the use of inflow and outflow totals as mass variables firmly supported, Plane (1984) moves forward to present the model form for derivation of inferred distances. Although Plane presents both the negative exponential form and the negative power function form, only the latter will be presented here. Fotheringham and O'Kelly (1989) state that it is generally accepted that the power function form of the model is considered to be more appropriate for long-distance interactions. In addition, the power function form provides scale-independent parameter estimates (p. 11). Fotheringham and O'Kelly (1989) explain that the inverse power function is most appropriate when the cases (or, in the case of migration, trip-makers) are heterogeneous and the distance-decay parameters for them are distributed according to a gamma distribution; meanwhile, the exponential function provides scale-dependent parameter estimates, and is best used for movement within an urban area, with the movements originating with a homogenous group of actors/cases. One caveat is that the power function form of the model overestimates both short- and long-distance movements (p. 11).

For his study of interstate migration, Plane presented the negative power function form of the doubly constrained model as:

$$m_{ij} = A_i O_i B_j I_j d_{ij}^{-\gamma} \epsilon_{ij} \quad (i,j = 1,\ldots, r;\ i \neq j) \qquad (6)$$

In which:

(1) m_{ij} are observed interregional flow levels;

(2) A_i and B_j are spatial competition terms (or balancing factors);

(3) O_i is the total number of out-migrants from region i;

(4) I_j is the total number of in-migrants to region j;

(5) d_{ij} is the distance from region i to region j;

(6) γ is a distance-decay parameter; and

(7) ϵ_{ij} is a multiplicative stochastic error term.

Equation 6 is then solved for d_{ij} as:

$$\ln(d_{ij}) = 1/\gamma \ln A_i + 1/\gamma \ln B_j + 1/\gamma \ln(O_i I_j/m_{ij}) + \epsilon'_{ij} \quad (7)$$

In which the new error term $\epsilon'_{ij} = (1/\beta)\ln \epsilon_{ij}$, encompasses specification error, measurement error, and pure random effects.

The inferred distance estimate for the negative power function model is then given as:

$$\hat{d}_{ij} = \exp((1/\gamma)\ln A_i + (1/\gamma)\ln B_j + (1/\gamma)\ln(O_i I_j/m_{ij}))$$

$$(i,j = 1,\ldots r;\ i \neq j) \quad (8)$$

Plane notes that while inclusion of balancing factors is still essential, calibration must be approached via a method other than the use of a determinate system of equations. He presents a method of obtaining inferred distances via a standard linear regression, using dummy variables to represent origin and destination balancing factors. For the negative power function, Plane defines the model as:

$$\ln(\hat{d}_{ij}) = a_0 + \sum_{k=1}^{r-1} a_k U_k + \sum_{\ell=1}^{r-1} b_\ell V_\ell + c \ln(O_i I_j/m_{ij}) \quad (9)$$

In which dummy variables are defined as:

$U_k = \{1 \text{ if } k = i, 0 \text{ otherwise}$ (10)

$V_\ell = \{1 \text{ if } \ell = j, 0 \text{ otherwise}$ (11)

Plane states that the balancing factors and distance-decay parameters may be found from the regression coefficients a_k, b_ℓ, and c, as follows:

$A_i = \exp((a_0 + a_i)/c)$ $(i = 1,\ldots,r - 1)$ (12)

$A_i = \exp(a_0/c)$ (13)

$B_j = \exp(b_j/c)$ $(j = 1,\ldots,r - 1)$ (14)

$\gamma = c^{-1}$ (15)

Plane states that B_j is fixed at unity as the balancing factors of the doubly constrained model are unique only up to a multiplicative constant (p. 247).

This model was then used to derive inferred distances, or functional separations, which were later compared to the observed physical distances. This was facilitated by creating a U.S. map based upon population centroids from the U.S. census which represented "aggregate locations of both potential migrants and of potential migration destinations" between which the observed and inferred distances are represented as vectors (pp. 248-249). Ratios of inferred to physical distance were calculated, then physical distance was multiplied against the ratio and a vector of the resulting distance was redrawn between the two centroids. When the vectors between one origin centroid and all destination centroids are redrawn in this fashion, the result is a "warping" of the original map of the U.S. to show the "migration space" for a particular origin centroid.

Overall, the methodology demonstrates how observed inflows and outflows can be used to derive the functional separation between population centroids, which may be greater or lesser than the physical distance between those centroids (pp. 248-249). Others had previously reversed the calibration of the doubly-constrained model, as Plane points out. Tobler, Mielke, and Detwyler (1970) used a similar reversed method with the unconstrained spatial interaction model to obtain "geobotanical distances" between islands in the New Zealand region, based upon numbers of common plant species between islands and the size of islands. The inferred geobotanical distances obtained were then compared to actual great-circle distances, facilitating a comparison between the geobotanical and geographical maps of the region to demonstrate "floristic relations" (p. 537). A decade later, Tobler (1981) used a method to obtain estimated migration flows between origins and destination when only a change in attractivity units (mass term) at an origin or destination is known. However, as Plane (1984) notes, never previously was the calibration of the doubly constrained model reversed in such a way as to attempt to obtain estimated distances based upon actual flows.

4.2. Model Overview: A Method for Investigating Trade Impedance

In his investigation of "migration space," Plane reversed the usual script for the doubly-constrained spatial interaction model. Plane asked: "In place of the usual practice of employing a matrix of interregional separations to obtain a matrix of interregional flows, suppose one asks what matrix of interregional separations used in the distribution model would exactly reproduce the observed matrix of flows?" (1984, p. 246).

In Plane's reconfiguration and reverse calibration of the doubly constrained spatial interaction model into a linear regression function, cases are the O-D flows; the dependent variable is the (natural log of) actual distance between origin and destination; the independent variables are comprised of a flow term (natural log of the product of total origin outflow and total destination inflow volume, over the O-D flow volume), and a vectorised matrix of origins and destinations (each represented as a dummy variable).

The result of interest is this expected value of the dependent variable (distance) for each case (O-D flow), which according to Plane can be conceptualized as generalized transaction cost (also referred to variously as inferred or expected distance) between origins and destinations (1984, p. 246). These expected values can then be used in a ratio of expected to actual distance in order to represent if/where this expected distance (generalized transaction cost) is disproportionate to physical (or absolute) distance. For Plane, this cost was specified as a generalized migration "cost" [p. 246]; Plane then used the ratios of predicted to actual distances with cartograms to determine the degree of warping of the U.S. outline that would occur in "migration space," in terms of selected state origins or destinations.

Likewise, for this dissertation research, if the output of predicted distances was to differ substantially from actual distances, such a disparity would be conceptualized as an "experienced" or expected-versus-actual distance mismatch. In this research, this is termed the expected-to-actual distance, or EtA, ratio. The expected distance is as-evidenced-by observed trade flows that comprise the independent variable matrix. In the literature regarding the study area, transportation cost is often identified as the leading contributor to lower (or less than potential) export flows in the region. However, when using a method which outputs the generalized transaction cost – in other words, the distance *experienced* by trade flows – this "experienced cost" or "experienced separation" may certainly be comprised of impedances other than transportation cost alone.

Political-economic, sectoral, and transportation factors have all been suggested in this literature as Latin American trade impediments, and their relative contributions have not yet been examined via an analysis based upon separation or generalized cost as experienced by trade flows. Therefore, the generalized transaction cost that this expected distance output represents has been conceptualized as "trade impedance" (Kockelman & Ruiz-Juri, 2003; Navajas et al., 2010). As it is used here, the term "trade impedance" (Kockelman & Ruiz-Juri, 2003; Navajas et al., 2010; see also Sect. 5.2) is meant to encompass the various terms which have been used to describe distance-as-experienced by the feature in a relative conceptualization of space - termed by Plane (1984) variously as generalized transaction cost; functional separation; or predicted, inferred or expected distance, and by Thill (2011) as nongeographic or relative distance. Here, the concept is applied to trade flows. Thus the initial regression analysis (a reverse calibration of the doubly constrained spatial interaction model) is used to obtain expected values of distance (trade impedance).

The modeling framework of this dissertation is depicted graphically in Fig. 10. This figure in its entirety represents the larger research endeavor. However, items indicated in red represent steps completed in this initial analysis performed as dissertation research. These are the first foundational steps of establishing the existence, location, and spatial pattern of trade impedance.

The ratio of predicted to actual distance for O-D flows represents the proportionality of trade impedance to actual distance, providing an initial measure indicating where in Latin America trade impedance may be higher or lower than what would be expected, based upon actual physical distances and observed O-D trade flows. Due to the focus of the literature upon trade impedance – and specifically the portion of it related to transportation cost – the flows of special interest are those which are revealed to have disproportionately high trade impedance (as compared to actual distance).

Upon obtaining the expected-to-actual distance (or EtA) ratio of trade impedance to distance for the total trade flow dataset, the regression analysis is repeated based upon a disaggregation of the total observed trade flow dataset by commodity categories, in order to obtain commodity category-specific O-D flows. As shown in Appendix A, six commodity categories have been devised based upon aggregation of two-digit Harmonized Tariff Codes (United States International Trade Commission, 2007). As discussed in the trade literature, these were aggregated into categories based upon transportability. This repetition of the initial analysis using the dataset disaggregated by commodity category segments examines trade impedance disproportionality to distance in terms of export composition (or sectoral) factors, as an initial exploration of these effects. This comprises the existence step of the research.

For flows with trade impedance found to be disproportionate to actual distance the location of occurrence begins to be specified by mapping the trade impedance proportionality (EtA ratios) for the total trade flow (total dataset) and for trade flows disaggregated by commodity category. This comprises the location step of the research.

In the spatial pattern step, the distribution of expected-to-actual distance (EtA) ratios for O-D flows is examined first using Global Moran's I analysis as an indicator of the existence and degree of global spatial autocorrelation in each data segment (for the total dataset and each of six commodity category segments). The spatial pattern of the distribution of expected to actual distance (EtA) ratios for O-D flows is examined first by assigning mean and median EtA ratios for O-D flows to the (aggregated) shipment origin points, then again by assigning the mean and median EtA ratios for O-D flows to the destination points (ports of export). (The analysis is performed this way as the number of flows is larger, for any data segment, than the number of origins or destinations.) The Spatial Autocorrelation/Global Moran's I tool in ArcMap 10.2 is used for both sets of points to determine the distribution of the mean and median EtA ratios among shipment origin points, then among ports of export. The distribution of the mean and median EtA ratios for flows among their associated shipment origin points and ports of export is based upon the assumption of randomization.

Following this, local spatial autocorrelation analysis is used to indicate significant clusters of low and high values and outliers, again using mean and median of the EtA ratios for flows at each associated origin and destination. The local analysis is performed using the Cluster and Outlier Analysis/Anselin Local Moran's I tool in ArcGIS 10.2. Both global and local analyses are performed using an inverse distance conceptualization of spatial relationships, Euclidean distance, and a threshold distance of zero, with row standardization. The clusters and outliers are visualized/mapped, and the analysis repeated for both the total flows dataset and the dataset disaggregated by commodity category.

As illustrated in Fig. 10, the analysis performed here for the purposes of dissertation research completes a portion of the overall research endeavor which may be performed as an extension to this analysis. Future extensions based upon this research design are discussed in Sect. 9.

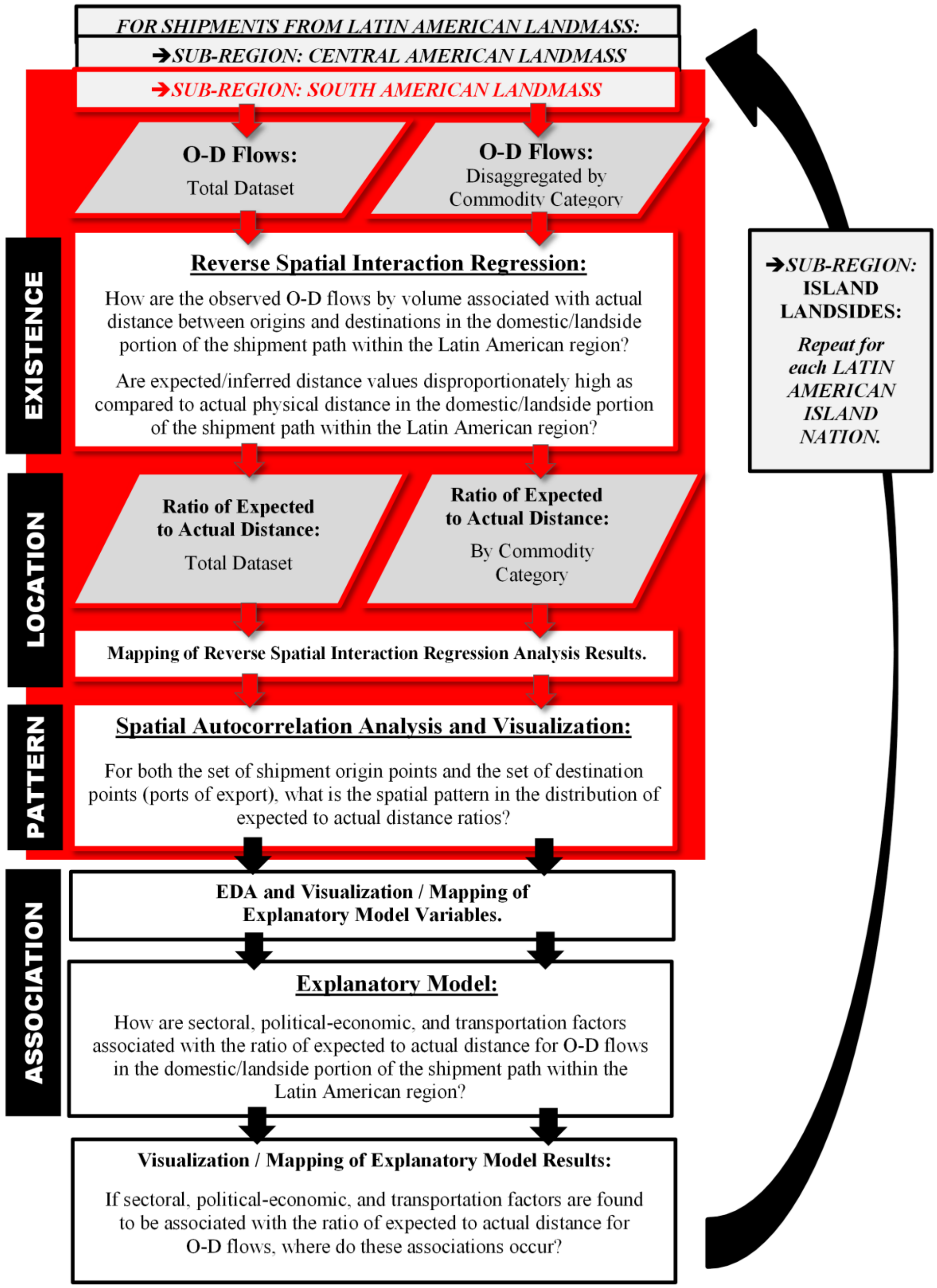


FIGURE 10: Model overview (Graphic by: K. Tiller)

DATA

5.1. Data Source and Verification Process

Shipment flows from shipment origins to final destinations (consignee locations) are comprised of a dataset of U.S.-bound waterborne containerized export shipments with South American origins. The full dataset covers a period of one year for all waterborne containerized U.S. imports; however, the subset of data used in this dissertation will cover one month only (October 2006). The data is comprised of information submitted through both the U.S. Customs and Border Protection Automated Manifest System (AMS) and manifests submitted at the ports; these are corrected for mistakes in entry/missing records and packaged with proprietary data field additions by data vendor PIERS, which then provides the raw data product for commercial sale (PIERS/UBM Global/JOC Group, Inc., 2007). Since 2006, development of the database and associated verification methodologies has been a long-term project by a research team under the leadership of Dr. Jean-Claude Thill in support of varied geographic and multi-disciplinary research applications extending to those currently underway as of the date of this dissertation (Thill, Tiller, & Tang, 2008; Thill, Tiller, & Kashiha, 2010; Tiller & Thill, 2011a, 2001b; Kashiha, Tiller, & Thill, 2013).

This database will stand as an extremely valuable resource for verified door-to-door shipment paths of every containerized U.S. waterborne import shipment for the period of one year. The verification of actual locations - from production origin to final shipment destination - addresses what has been termed “a legacy problem” of unverified or missing physical waypoints in shipment path data, heretofore preventing shipping manifest data from being useful for determining physical movements of goods/actual shipment paths. This immense challenge has gone unresolved since such data have been available for analysis - since at least the 1970s. Certainly the spatial extent, duration, and volume of a locationally verified shipment database has never previously been developed for academic research purposes. The time period of the dataset places it before the global financial crisis of 2008; therefore it stands as an excellent base-year dataset.

The development of this unparalleled spatial database is briefly detailed here in order to trace the development from the original collection of raw shipment records from the data vendor (of which the Latin American export records are one subset) to the verified spatial data subset used in this dissertation analysis, which is one research trajectory of the larger overall project. The original collection of raw shipment records from the data vendor include all containerized waterborne imports to the U.S. for the period of one year (July 1, 2006 through June 30, 2007), for a total of approximately 9.5 million individual shipment records. To develop the process for tackling the substantial data cleaning and verification steps necessary to locationally verify this large dataset, a sample of one month (October 2006) of U.S. import shipment data was used. Creation of the verified database required development of machine-learning algorithms and manual methods for locational verification of data items (such as shipment origins by country and locality), verifying coordinates of shipment origin localities, and categorization of shipment movement typologies, as well as other improvements, as detailed below.

The unique data cleaning process employed for this dataset includes verification of physical origins (production locations) of shipments, and assessing/verifying origin and waypoint locations where locations were mistaken or missing in the raw data. First, locality-of-origin names were queried to determine their match to files of known localities. Two locality databases were merged to perform this matching function – one originating from the National Geospatial Intelligence Agency (2006) and the second originating from Europa Technologies' Global Insight Plus (2009.2). In order to use this process to mine useful locational information from several fields within the records, a SQL database (Microsoft Corp., 2005, 2008) was developed. This application's proprietary tools, such as SQL Studio Business Intelligent Development data mining/exploration tools, were utilized to complete many of the automated tasks. In addition, customized SQL querying and coding was authored by members of the research team.

Market intelligence is the crux of this process, as the logistical realities of global commodity chains were connected with production locations for companies with multiple facilities. Industry, product, and geographical realities were examined together to determine the most likely paths, origins, and waypoints for shipments with missing or mistaken locational information. An example of a decision point encountered in the manual process is the determination of whether the given address represents the physical production origin or corporate/legal address of a corporation, when locational information appears to be incorrect – and, if so, the determination of the correct production origin locality (often via learning the geography of production, as necessary). Facility locations were determined from corporate websites and from evidence within shipment records - such as in the case of a "precarrier location" data field match with a particular production facility (precarrier location being the point at which a carrier took possession of the cargo).

This manual process, while more labor intensive, was undertaken for the purpose of salvaging those records which might otherwise have been eliminated from the "clean"/verified dataset, as physical origin points could not be verified using the automated processing of raw shipment record data for this portion of the records. Additional refinement of the data cleaning for purposes requiring it may be possible. However, the market intelligence method used here enabled accurate total shipment flows to be preserved – with verified physical origins - for applications such as the one used for this dissertation. Through the preservation of these records, a verified, accurate, and complete spatial dataset of U.S. import flows has been achieved for this purpose.

Basemap data for analysis, mapping, and presentation was acquired from Global Insight Plus, version 2009.2 (Europa Technologies, 2009), for country and state polygons, water features, and transportation layers (including seaports). All other data sources for variables are as displayed in Table 4, which presents the variables for the Reverse Spatial Interaction Regression Analysis.

5.2. Definition of Terms

Several terms referencing data and variables in the analysis require definition. The first term requiring definition is *shipment*. The disaggregate or "atomic" unit (Tagashira and Okabe, 2002) of the data is a shipment - cargo travelling from shipper (at the locality of origin) to consignee (at the final destination). The definition of a record as a shipment comes from its entry on a single shipping manifest filed with the U.S. Customs and Border Protection Agency (PIERS/UBM Global, 2007). This definition is primarily an administrative one, as a shipment is defined by what is entered on the shipping manifest (record/filing); therefore there are no uniform parameters as to physical content of a "shipment." There are definite reasons why no additional parameters or definitions are included here, and this is because of the characteristics of the "shipment" as an atomic unit in this sense. To demonstrate - each shipment in this dataset:

1. May or may not comprise a full container;
2. May or may not be comprised of more than one container;
3. Therefore, may or may not be carried by more than one vehicle (truck, vessel, etc.);
4. May or may not be comprised of multiple commodities, for which proportions by value, weight, and/or volume cannot be verified;
5. Is not directly comparable to other shipments other than by a standard measurement of volume (twenty-foot equivalent units, or TEU) or weight;
6. Cannot be associated with a verifiable value based upon its native/original attributes;
7. Has only, in terms of this dissertation, certain verified or reliable attributes:
    a. Physical location of origin, including:
        i. Verified country of origin;
        ii. Verified locality within the country of origin;
    b. Port of export information: port name and port code (U.S. Army Corps of Engineers Schedule K code) where shipment is placed on an oceangoing vessel for eventual import to the U.S.;
    c. Volume of shipment in TEUs (twenty-foot equivalent units);

d. Harmonized tariff code (6-digit) (later generalized to commodity category, which in a small percentage of cases is coded as miscellaneous/unknown);

(Appendix D contains a full listing of data fields present in the original dataset and any improvements made to the dataset.)

Unlike the data used in similar or related transportation and trade flow analyses, each "shipment," as defined by its shipping manifest filing, represents the atomic unit of data; in other words, the shipment is as-defined on its manifest, and therefore does not represent typical units used in transportation modeling (such as a container load or vehicle trip). Nor can the shipment be defined by a "value," as its monetary value is assigned by a proprietary formula used by the data vendor. Therefore, there are few entities to whom the disaggregate units of "shipment" would have meaning or use. For this reason, this analysis defines shipment flows by volume measured in TEUs, or twenty-foot equivalent units.

The "landside" portion of the shipment path used in this dissertation is the domestic portion of the U.S.-bound South American export journey - the O-D flow from departure at a verified locality of origin within South America (aggregated and assigned to "shipment origin aggregation points" as defined below), to arrival at a South American port of export (as the shipment is en route to a U.S. port of import/U.S. final destination). (It should be noted that in cases in which a third-party logistics service has been used to consolidate and pack multiple items or commodities together to constitute one "shipment," it cannot be verified as to whether the locality of origin verified in the raw shipment data represents a production or consolidation locality; nevertheless, these consolidated shipments constitute a very small percentage of the total dataset.)

In the original dataset, the initial shipment origin point for each shipment is identified by country, locality, and associated coordinates. In this dissertation analysis, however, "shipment origin point" refers to aggregated points of shipment origin (also termed "weighted mean centers of 100-km grid cells"). For this dissertation, the original shipment origin points shapefile was overlaid with a 100-kilometer cell size vector grid (created using ET GeoWizards, Version 11.0 for ArcGIS 10.1 [Tchoukanski, 2013]). The weighted mean center of each vector grid cell was determined using volume (TEU) as the weight. The weighted mean center of shipments originating from each 100-kilometer vector grid cell then functions as the representative shipment origin point for these aggregated flows. (Total outflow from each weighted mean center/shipment origin point is then calculated by summating the TEUs of the aggregate outflow for each grid cell.). As these aggregated outflows from shipment origin points and inflows to "destinations" (ports of export) are included in the construction of variables in the Reverse Spatial Interaction Regression Analysis, the aggregation scheme will be further detailed in Section 5.4. Different commodity categories have differing numbers of shipment origin points, as not all original 424 grid cells of the total dataset export all commodities.

"Destination," "shipment destination," or "seaport," in this dissertation, all refer to the South American port of export which terminates the South American domestic/landside portion of the shipment path. There are 34 South American seaports which have been identified as domestic/landside shipment destinations for this dataset.

Several ports which originally were included in the dataset (and their associated cases) were dropped from the analysis; these are ports with dataset TEU inflow volumes which total less than 63 TEU. This threshold was not a decision point for dropping these ports, but instead is mentioned here to illustrate the very low volume of such ports which differentiates them from other ports in the dataset. The lowest volume of these ports have throughput and trade realities which are very dissimilar to other ports within the dataset. Additionally, 10 ports which were determined to be within close proximity of each other (by the same measure of proximity used to aggregate origins), and which were found to have direct-call carriers, significant hinterland overlap, and/or a vast difference in annual TEU throughput (according to Containerisation International), were merged into "pseudo ports" representing each proximate port pair. The 5 pseudo ports merging 10 proximate ports are: Pecem-Fortaleza (Brazil), Rio de Janeiro-Sepetiba Bay (Brazil), San Vicente-Lirquen (Chile), Guayaquil-Puerto Bolivar (Ecuador), and Barranquilla-Santa Marta (Colombia), for a total of 34 port of export destinations for the total dataset. Different commodity categories have differing numbers of port of export destinations, as not all categories have shipments exporting from all 34 ports.

For this dissertation, ports will be used as destinations, instead of individual terminals within ports. The main reason for this is that within the dataset, the raw data is reported primarily in terms of ports, not disaggregated to the terminal level; therefore, it cannot be verified which terminal was used (except in cases which it is obvious due to the existence of only one appropriate container terminal).

"Physical," "actual," or "absolute distance" here refers to distance as measured between shipment origin points and destinations/ports of export, using ellipsoidal distance based upon the South American Datum of 1969. The coordinate system used is GCS South America 1969, and the projection used is South America Equidistant Conic (ESRI, 2008). Actual distance between origin and destination (natural log transformation) served as the dependent variable, determined via the Calculate Geometry function within ArcMap 10.1 to determine the length of OD flow polylines, generated via the XY to Line function. Distance is measured in meters.

"Trade impedance" is a term originally used by Kockelman and Ruiz-Juri (2003) in an input-output analysis simulating trade and labor flows between the 254 counties of Texas based upon domestic and foreign demand; in this study, the authors use travel cost on a bi-modal transportation network to represent trade impedance. Navajas et al. (2010) also used the term in a project funded by the European Commission Directorate General for Energy and Transport. The project aimed to estimate the carbon footprint of freight transportation based upon the drivers of the logistics system; in the trade model portion of the project, using a doubly-constrained spatial interaction model, the authors estimated intra-country trade. Navajas et al. (2010) conceptualized trade impedance as "border barrier effect," which in combination with a trade association membership variable was found to be captured by a variable titled "country-pair specific border barrier." This variable was representative of trade resistance factors, the most important of which was geographic distance (pp. 62-63). It is clear that the term "trade impedance" fits here, in an analysis that seeks to include both transportation cost and border-related trade barriers (as well as commodity-specific effects). In more general terms of measures of separation, Roy and Thill (2004) have conceptualized this as impedance. Previous discussion related to the relative distance concept underlying the use of this term here (though not specifically the term "trade impedance" itself) by Plane (1984), Notteboom (2001), and Thill (2011) can be found in Chapter 4 of this document.

"Commodity categories" are defined as presented in Appendix A. The six commodity categories represent aggregations of the Harmonized Tariff Schedule of the United States (United States International Trade Commission, 2007), termed "HS codes" in this dissertation. Commodity categories are comprised of groups of similar commodities defined by ranges of two-digit HS codes. The commodity categories are defined as perishable products, non-perishable time-sensitive products, heavy commodities/minimally processed natural resources, manufactured products, high-value manufactured products, and miscellaneous/unable to be determined. A more detailed review of the composition of each commodity category is available in Appendix A. (Additionally, where the term "segments" or "data segments" is used, this refers to one of the subsets of the total dataset which is parsed by commodity category - Commodity Categories 1 to 6. The complete dataset is termed the "total," "total flow," or "all shipments" dataset.)

Commodity categories are grouped in reference to the previously presented hypotheses regarding commodity-specific transportability. Within the original shipment dataset, some shipments contained multiple commodities not in the same aggregated commodity category. For these shipment records, each record was split by commodity category to create single-commodity category shipment records, as described further in Sect. 5.5. When commodity categories were unable to be determined based upon the HS codes or the brief commodity descriptions given in the shipment record, such shipments were designated as Commodity Category 6 (Miscellaneous/unable to be determined).

Commodity values were determined in the following manner (for determination of Commodity Category 5 (High-value manufactured products). The raw data vendor (PIERS/UBM Global/JOC Group Inc.) had included in their data additions/improvements a field titled "Estimated Value," which used a proprietary formula to determine an estimated value for each shipment. Per PIERS/UBM Global/JOC Group Inc. (2007), the estimated value for each shipment in the raw data was calculated thusly:

> "The value is generated by dividing the value by the weight in the table for each U.S. coast, country, and HS code [six-digit Harmonized Tariff Schedule code] combination to obtain a value per ton factor. That factor is then applied to the PIERS database for each U.S. coast, country, and HS code combination to derive an estimated cargo value for each PIERS transaction. Taxes and duties are not included in the calculation. This method is purely "estimation." The estimated value should not be used for trending." (PIERS/UBM Global/JOC Group Inc.), 2007)

For this dissertation, a commodity type's relative value was calculated as a ratio of shipment value (as described above) to shipment weight in kilograms, or shipment value per kilogram, for a sample of the data. Shipments' six-digit HS codes were then truncated to two-digit HS codes (corresponding with each chapter of the Harmonized Tariff schedule) and an average value per kilogram for each two-digit HS code was calculated, as an average of the ratios for each shipment. The quartiles of these average value ratios were then determined. The "high-value manufactured goods" category is comprised of those commodity types (two-digit HS codes) which have average value-to-weight ratios (value per kilogram) in the 75th percentile, among average per-kilogram values for each commodity (2-digit HS code). Commodity types with an average value-to-weight ratio of 10.42 and above are in the 75th percentile, and are therefore considered to be "high value." (It should be noted that the commodity category "high-value manufactured goods" also includes some commodity types that would have been included in the "non-perishable time sensitive goods" category if not for their high value – such as apparel items types identified as high-value.)

It should also be noted that shipments containing multiple commodities (multiple items from differing two-digit HS codes) were placed in a separate category (Multiple Commodities) for valuation. While these records will be parsed by commodity category for the purposes of analysis, there is no method that could be used to parse the mixed-commodity shipment value field with any reasonable degree of accuracy. (This is consistent with the U.S. Census Commodity Flow Survey, which identifies some shipments as "mixed freight" for valuation) (U.S. Department of Transportation [DOT] Bureau of Transportation Statistics [BTS]/U.S. Census Bureau, 2007).

Whenever "ratio" is used, this refers to EtA ratio (expected-to-actual distance ratio), unless otherwise noted. Where "higher-ratio flows" is used, this is meant to be inclusive of both the high- and very high-EtA ratio flow subsets.

5.3. Design Limitations

One major limitation in this research design should be noted in terms of spatial autocorrelation analysis. LeSage and Pace (2008, p. 942) state, "The notion that use of distance functions in conventional spatial interaction models effectively capture spatial dependence in interregional flows has long been challenged." LeSage and Pace go on to make the example that for flows which create spatial spillover effects, origins have neighboring regions, destinations have neighboring regions, and O-D flows create a link between these neighbors (p. 947). Both Fischer and Griffith (2008) and LeSage and Pace

(2008) present that O-D flows have unique spatial characteristics which make both non-spatial statistical methods (with the assumption of independence) as well as traditional spatial statistical methods (with methods developed for point data, not O-D pair data) inadequate. Fischer and Griffith (2008) used a module titled "spdep" in R spatial statistical software (Gentleman & Ihaka, 2011) in order to investigate spatial autocorrelation within O-D flows.

However, this concept and its associated methodology, while acknowledged, cannot be used here. The shipment records represent only the origin point, destination point, and volume of flow – not the route which was taken between the origin and destination, which hearkens to the routing step of the transportation problem. Obviously, it is a limitation that shipment routes and modes on the transportation network are unknown. Estimating these is an eventual goal of the extended research. The fact that the two points of origin and destination have been verified should not be misconstrued to suggest that the straight-line distance between the points is an actual land-side distance traveled; rather, it is an acceptable *proxy* for actual land-side distance traveled.

Indeed, straight-line distance is used in this analysis as an approximation, and trade impedance ratios are visualized using straight lines between origins and destination, for display purposes. However, assuming a straight-line *route* as an approximation between the known origin and destination points for the purpose of spatial autocorrelation analysis would be a further abstraction and a means of introducing error. For this reason, the separate analysis of spatial autocorrelation for the *known* origin and destination points is used instead.

A more ideal method of exploring the spatial pattern of O-D flows is paired-point cluster correspondence (as in Lu and Thill, 2003). (Conceptually, this multiple-point set/multi-location event methodology also accords well with an analysis of spatial autocorrelation for known origin and destination points, for which the actual route between points is unknown - and for which the assumption of such would be further abstraction.) However, this is a significantly resource- and skill-intensive research endeavor of its own merit, which stands as a potential future extension of this analysis. Baseline results obtained as described above could later be compared with those garnered from a more intensive focus on vector autocorrelation analysis.

On a related note, O-D flows cannot necessarily be construed to capture the attributes of the spaces traversed between origins and destinations at this point in the analysis. This limitation is related to the stage of modeling; as actual routes (and indeed, even modal choice) of O-D flows are unknown at this point, it is not possible to definitively identify which specific areas between origins and destinations can be associated with the routes taken, nor is it possible to include these attributes in the current stage of modeling. Methods for approaching this challenge are discussed further in Sect. 9. Attempted routing of O-D flows along the network can be completed in a future extension of this analysis. However, currently, the purpose of this dissertation is to narrow actual locations of disproportionate trade impedance evidenced by trade flows – giving a general indication of area of occurrence and spatial pattern.

This is another indication that handling the aggregation of the disaggregate shipment origins through the use of a fine resolution vector grid cell overlay may present a level of exactitude that actually is not present natively in the raw data.

Another major limitation involving data is the difficult interpretation and application of the disaggregate unit of the independent variable (shipments, which constitute O-D flows by volume) for the Reverse Spatial Interaction Regression Analysis. This makes the use of a disaggregate analysis impractical, as the interpretation of results related to that data unit would suffer from difficulty in interpretation and application as well. This is discussed further in the following section.

A related limitation involves units reported in the original data as well. The shipment volume in TEU for the original multiple-commodity records will be split proportionally based upon a field given in the original data ("QTY" field) which gives unit counts by commodity. As this was the only field in the original data giving a commodity-specific numerical indication of shipment composition, this field was the only available option for splitting records by commodity. As a future extension of the analysis, a more rigorous and extensive analysis could attempt to use coding to proportion volume when splitting multi-commodity records based upon HS code and unit type. However, this method may also be quite limited, as the original data often reports only one unit type when it is clear from other descriptive fields that other, very different unit types comprise the shipment as well. This limitation is an unfortunate artifact of the original data reporting.

Also, in this dissertation research it is important to note that only containerized shipments are included. For a complete picture of U.S.-bound export trade originating in Latin America, non-containerized shipments should be included. Additionally, the dataset is comprised of U.S.-bound flows only.

It should be added that the use of one month of shipment data is not directly generalizable to the entire year; extensions of this analysis should include additional seasons and/or months.

A limitation that is inherent in the use of third-party basemap data is that the analyses based upon these data sources inherit any error in the original data.

5.4. Aggregation Method and Data Cube

In terms of spatial interaction modeling, Roy and Thill (2004) caution the modeler to use judgment in reference to the advantages and disadvantages of particular methods of aggregation. Among the possible reasons for aggregation of flows as given by Roy and Thill is that although there are discrete operators at the micro level (in this case, individual shipments from origin locality to destination port of export), the number of individual shipment flows is great. The need for aggregation stems not only from the need to reduce the number of individual flows for convenience in handling of data and computation - but more so from the need to draw meaning both from the units of flow as well as to draw both meaning and usability from the results. In this regard, the modeler must ask: What would such an analysis mean for each individual flow? How would the results be interpreted? Of course, the interpretation and eventual application of the results should be linked back to the original purpose of the analysis. For instance, if one purpose of the analysis, as stated, is to inform policy by determining the need for further assessment of transportation infrastructure quality, then the analysis should be carried out at the aggregate level of the state/province, or some other sub-national administrative division relevant to policy-making.

In addition, Roy and Thill (p. 344) note that flow modeling which will be used to predict demand can be expected to be estimated at the aggregate level. Holguín-Veras and Patil (2007, p. 60) use term “policy-sensitive” model in discussing the need for their demand models to have meaningful interpretation to reveal the impact of “freight-specific policies.” Holguín-Veras and Patil (2007, p. 60) also state that freight models are currently limited by their major modeling platforms – commodity-based and trip-based movement - neither of which can accurately or completely recreate freight movement within the model.

These concerns are relevant to the dataset used for this dissertation in that, as detailed in the preceding section, shipments (as disaggregate units of data) are not defined as container loads, vehicle trips, or monetary values. There are few entities to whom the disaggregate units of “shipment” would have meaning or use, save individual firms or corporations between which the shipment records represent interaction, or governmental entities involved in assessment and collection of tariffs and specific trade regulation (such as import quotas). This indicates that for the bulk of academic, public sector, and corporate entities potentially interested in this research, the disaggregate analysis would carry little meaning. This analysis defines the aggregated shipments as *shipment O-D flows* by volume, measured in TEUs, or twenty-foot equivalent units, a measure which is readily interpretable. In the Reverse Spatial Interaction Regression Analysis, individual shipments are aggregated into shipment outflows from shipment origin points and shipments inflows to domestic destinations/seaports, in order to obtain trade impedances.

But what of aggregation bias? Some previous studies using gravity or spatial interaction models appear to have used what is most interpretable or convenient as aggregation schemes, such as Chisholm and O'Sullivan's (1973) use of 78 zones, each represented by the "most important city" contained within (p. 34). Others, such as Tagashira and Okabe (2002) entered into a detailed, thorough, systematic investigation of the modifiable areal unit problem (including scale and aggregation issues [p.2; Openshaw, 1984]) and potential ecological fallacy implications of the aggregation scheme chosen. In a regression analysis in which the dependent variable is an attribute of the disaggregate spatial data unit, and one independent variable is the distance from each disaggregate spatial data unit to a central point, Tagashira and Okabe (2002) found those aggregation schemes and methods which would enable the data placed within the aggregation scheme to approach the performance of disaggregate data by comparing the relative efficiency of variance of the estimator for the slope coefficient in terms of the number of zones. In order to find the zoning system with the minimum variance, the authors used different shapes and numbers of zones, finding that the best-performing aggregation system was equal-width concentric zones or square grid zones - with higher numbers of zones leading the variance of the slope coefficient estimator to approach that of the disaggregate model (p. 19).

While the finding by Tagashira and Okabe is useful (though the finding regarding higher numbers of zones is hardly surprising), the reason that the authors entered into the investigation was due to the assumed unavailability of disaggregate data. Indeed, often these decisions are simply constrained by data availability. However, for this dissertation, data disaggregated to the shipment level is available. Yet interpretability of the unit of flow requires an aggregation scheme for the shipment records dataset.

As an alternative, a focus upon inherent interpretability of spatial units might suggest the use of sub-national administrative units (hereinafter referred to as states) as the units of aggregation. This aggregation method would especially lend itself to the purpose of application and interpretation in terms of policy. If used in this dissertation, approximately 500 states would be used as spatial units of aggregation; representative aggregation points would be comprised of weighted mean centers of the shipment origin points falling within each state, with weighting by volume of shipment in TEUs. However, this aggregation method has a significant drawback, as larger spatial units would be associated with greater error in calculating distances between origin and destination. Likewise, grid cell aggregation has the drawback of potentially producing populated grid cells which cross borders, thus causing interpretation to be difficult in terms of origins.

As a compromise providing less error in calculating distances than would the aggregation to state spatial units, shipment origin point data is aggregated to 100-kilometer grid cells of a vector grid overlay of the study area. The weighted (by TEUs) mean centers of the shipment origin points falling within each grid cell are used as representative aggregation points of shipment origins.

Error in calculating distances from weighted mean centers of 100-kilometer grid cells as origins is expected to be far less than from the weighted mean centers of states as origins. This does leave the problem of interpretability unaddressed, as a 100-kilometer grid cell unit has in itself no inherent meaning here. However, this problem could only be conquered through the use of administrative units, and due to the crucial nature of distance in this analysis, the reduction of error in measuring distance supersedes the importance of using inherently interpretable spatial units.

As with any form of aggregation of disaggregate point data to polygon spatial units, there are limitations which must be addressed. There are endless possibilities of different results that may be obtained by varying grid cell size, shape, and placement/alignment.

In geography, the concepts of modifiable areal unit problem and boundary effects speak to these issues, which are essential to spatial analysis. As Fotheringham, Brunsdon, and Charlton (2000) explain, it has been widely demonstrated that the aggregation of data into different systems of zones (or spatial units) can lead to different results in analysis, and therefore different conclusions; this is what is referred to as the zonal effect of the modifiable areal unit problem (MAUP) (Fotheringham, Brunsdon, and Charlton, 2000, p. 237; Openshaw, 1984). Related is the scale effect portion of the MAUP problem, in which it can be demonstrated that different results can be obtained by aggregation at different levels of spatial resolution (Fotheringham, Brunsdon, and Charlton, 2000, p. 237). This is addressed by using either disaggregate data, or the most spatially disaggregate level possible, and then visually demonstrating the effects of both scale and zoning elements of the MAUP problem.

In terms of scale, though the vector grid cell aggregation system is chosen as a compromise method (as previously described), it is certainly possible to demonstrate the dual differences of scale and zoning systems by comparing an analysis with aggregation of shipment origins to states and an analysis with shipment origins aggregated to the 100 km grid cells. However, it seems that a more useful comparison would be that of a sensitivity analysis of grid cell size using 50- and 100-km grid cells. These grid cell sizes were chosen to approximate a maximum one- and two-hour drive, respectively, from any point within the grid cell to any other point within the grid cell. This is touched upon in Sect. 9, as part of the larger research effort/extension.

Also, it should be noted as a future extension of the analysis, these results could also be compared to that of the disaggregate analysis; however, the sensitivity analysis noted previously may give an indication of the sensitivity of this relationship. As noted in Fotheringham, Brunsdon, and Charlton (2000), a study by Fotheringham and Wong (1991) found that some relationships may be more sensitive than others to data aggregation. Also, differently-shaped vector grid cells, as well as different alignments of the grid, could be investigated as future extensions of this analysis. However, these issues will be relegated to future investigation, as comparing the different types of aggregation schemes themselves will be the focus here.

Roy and Thill (p. 343) also suggest market segmentation, when it is unlikely or unknown whether the variance within the unit of aggregation is greater than the variance between units of aggregation. While the Reverse Spatial Interaction Regression Analysis in this dissertation is performed on the total O-D flows dataset, this analysis is repeated using flows disaggregated into six commodity categories (Appendix A), in accordance with this notion. Holguín-Veras and Patil (2007, p. 60) explain that within freight transport models, commodity type stands as a proxy for market segment.

The segmentation of the analysis by commodity category is graphically conceptualized in Fig. 11 below, in the form of a data cube. In envisioning this segmentation of the O-D flow cases, origins and destinations are represented on the x-and y-axes, whereas the z-axis represents the commodity category. The analysis will be performed both in terms of the O-D flow for each commodity category (Appendix A), and also in terms of the total O-D flow (or sum of commodity category O-D flows).

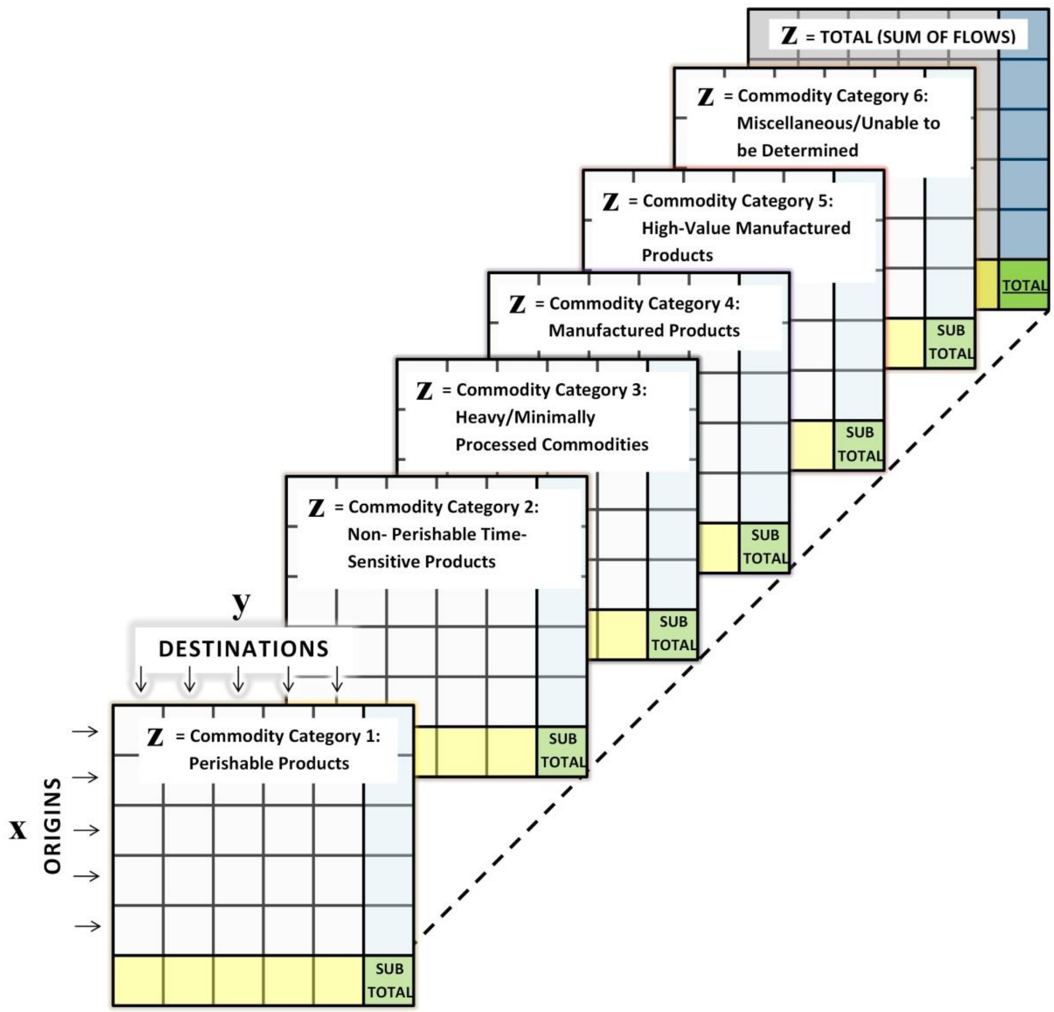


FIGURE 11: Data cube (Graphic by: K. Tiller)

5.5. Variables

The Reverse Spatial Interaction Regression Analysis follows from Plane's negative power function form of the equation for reverse calibration of the doubly constrained spatial interaction model (Eq. 9). The same equation and variable types are used for both the analysis of total flows and the analysis of flows disaggregated by commodity category. Table 4 lists the variables for the Reverse Spatial Interaction Regression Analysis and their equivalents in Plane's (1984) analysis.

The dependent variable is actual distance between each shipment origin point and each port of export. Use of actual distance as the dependent variable in the context of this dissertation - intended to derive expected distances conceptualized here as trade impedance - is supported by the use of distance as a proxy for transportation cost in previous studies (Bergstrand, 1985; Clark, Dollar, & Micco, 2004). More generally, it is supported by the endorsement of the appropriateness of distance as a measure of separation in spatial interaction models (Chisholm and O'Sullivan, 1973). The actual distance variable relies upon an absolute conceptualization of distance, whereas the output of the expected distance is conceptualized as a relative distance (see Sect. 4.2) but expressed in the units of absolute distance. (Here actual geodesic distance is calculated via the Calculate Geometry function within ArcMap 10.1 to determine the length of OD flow polylines, generated via the XY to Line function. Distance is measured in meters.)

The next term in the equation is the intercept for the regression, representing value of the dependent variable (distance) in terms of complete absence of independent variable (shipment flow volume).

Independent variables include the origin and destination dummy variable terms, as noted in Plane's analysis. These terms add k-1 origin dummy variable terms and ℓ-1 destination dummy variable terms, where k and ℓ are shipment origin aggregation points (weighted mean centers of 100 km grid cells) and ports of export, respectively. Each dummy variable in the estimation indicates whether there exists outflow from the origin indicated or inflow to the destination indicated.

Finally, the O-D flow volume variable provides the measure of shipment flow by volume (TEUs) between origin and destination for each case, within the context of total outflow volume from the origin (in TEUs) and total inflow volume to the destination (in TEUs). This variable is measured by the natural log of the product of the outflow total for shipment origin point k and the inflow total for port of export ℓ, divided by the total shipment flow from shipment origin point k to port of export ℓ. This variable demonstrates the interchangeability of associational and attributal properties (or in Plane's model, regions). As Plane states, "Any variable that affects the general attractiveness of a region should influence total inflow and outflow, while any variable measuring the specific attractiveness of a region relative to another region should influence an inferred distance" (p. 246).

In order to derive the expected distances using Plane's (1984) reversal of the doubly-constrained spatial interaction model (negative power function form) via a linear regression procedure, Eq. 9 in this document (Eq. 16, Plane, 1984, p. 247) is estimated through regression (used to estimate distance conditional upon the predictor values). Therefore, each (former) dummy variable representing participation of an origin or destination in an OD flow case (value=1), versus non-participation (value=0), instead becomes a scale-level variable.

To address parametric identification issues with the dataset as described above, particularly resulting from zero M_{ij} flows, the estimation is done with $M_{ij} = 0.01$ TEU instead. This approach allows us to account for cases where no flows are recorded in the dataset, which is meaningful information about the state of the freight interaction system in South America.

Additionally, the flow volume (M_{ij}) - including that for flows calculated with M_{ij} = 0.01 TEU (replacing zero-flow/M_{ij} = 0 interaction volumes) - is included in the analysis as the case weighting variable. Thus all results and residuals are calculated as part of a weighted least squares analysis. Residuals resulting from the estimation procedure are calculated as standardized weighted residuals, taking the product of each residual by the square root of the weighting variable (M_{ij}, including 0.01 replacement for zero-volume flows) and dividing by the standard error of the estimate. The standardized weighted residuals are then used to determine which flows have expected distance values significantly different from their actual distance values, for each dataset segmentation.

TABLE 4: Reverse spatial interaction regression analysis - variables

Variable Type	**Variable in Plane's 1984 Reverse Calibration Equation for Doubly-Constrained SI model**	**Operationalized Variable (for this Dissertation)**	**Concept Represented**	**Data Source**
Dependent variable: **O-D Actual Distance**	$\ln(\hat{d}_{ij})$ Natural log of distance between origin state k to destination state ℓ.	Distance (meters) between weighted mean centers of vector grid cells (shipment origin points) and South American ports of export (destinations).	Distance as a proxy for transportation cost (Bergstrand, 1985; Clark, Dollar, & Micco, 2004); Appropriateness of distance as a measure of separation in spatial interaction models (Chisholm & O'Sullivan, 1973).	Containerized waterborne U.S. imports originating in South America - October 2006 dataset (PIERS/UBM Global 2007). Basemap data source is Global Insight Plus (v. 2009.2) (Europa Technologies, 2009).
Independent variable: **Origin Dummy Variable**	$\sum a_k U_k$ Dummy variable indicating migration outflow from a particular state of origin.	Dummy variable indicating shipment outflow from a particular shipment origin point.	Correspond to balancing factors for doubly constrained spatial interaction model (Plane, 1984).	As indicated above; origins are weighted mean centers of 100-kilometer vector grid cells (weighted by shipment volume in TEUs of individual shipments originating within cell).

Table 4 (continued)

Variable Type	**Variable in Plane's 1984 Reverse Calibration Equation for Doubly-Constrained SI model**	**Operationalized Variable (for this Dissertation)**	**Concept Represented**	**Data Source**
Independent variable: **Destination Dummy Variable**	$\sum b_\ell V_\ell$ Dummy variable indicating migration inflow to a particular state of destination.	Dummy variable indicating shipment inflow to a particular port of export (destination point).	Correspond to balancing factors for doubly constrained spatial interaction model (Plane, 1984).	As indicated above; destinations are South American ports of export.

Table 4 (continued)

Variable Type	**Variable in Plane's 1984 Reverse Calibration Equation for Doubly-Constrained SI model**	**Operationalized Variable (for this Dissertation)**	**Concept Represented**	**Data Source**

Independent variable: **O-D Flow Volume Term**	c $\ln(O_iI_j/m_{ij})$
	
	Natural log of:
	• The product of:
	  ○ Outflow total for origin state k;
	  ○ Inflow total for destination state ℓ;
	• Divided by the total migration flow from origin state k to destination state ℓ.
		Natural log of:
		• The product of:
		  ○ Outflow total for shipment origin point k in TEUs;
		  ○ Inflow total for port of export ℓ in TEUs;
		• Divided by the total shipment flow (in TEUs) from shipment origin point k to port of export ℓ.
		
		(Both a total flow analysis and separate analyses using flows disaggregated by commodity category will be performed.)
			Interchangeability of associational and attributal properties of origins/ destinations (Plane, 1984).
			
			(Potential differences in shipment flow distribution based upon commodity type [Hummels, 2001; Roy & Thill, 2004; Inter-American Development Bank, 2010].)
				As indicated above.
				
				(Commodity categories are aggregations of Harmonized Tariff Codes from the United States International Trade Commission [2007-Revision 2].)

Regression Intercept	a_0 Intercept	Intercept	Regression Y-intercept (Plane, 1984).	As indicated above.

5.6. Dataset

For the total dataset comprised of 30,005 original shipments (78,296.51 TEUs) aggregated to 100-km origin grid cells, there are 424 shipment origin aggregation points on the South American landside and 34 South American ports of export, creating a matrix of 14,416 potential interactions. However, there are only 680 O-D combinations for which shipment flow volume is observed. The total dataset includes all commodity types. Figure 12 displays the volume of the total dataset by country of origin (based upon original shipment points). Tables and discussions from this point forward indicating a country reference the port-of-export country (the country in which the port of export is situated). Table 5 displays the port-of-export country share, and Table 6 displays the top-five port shares, for the total dataset and commodity segments.

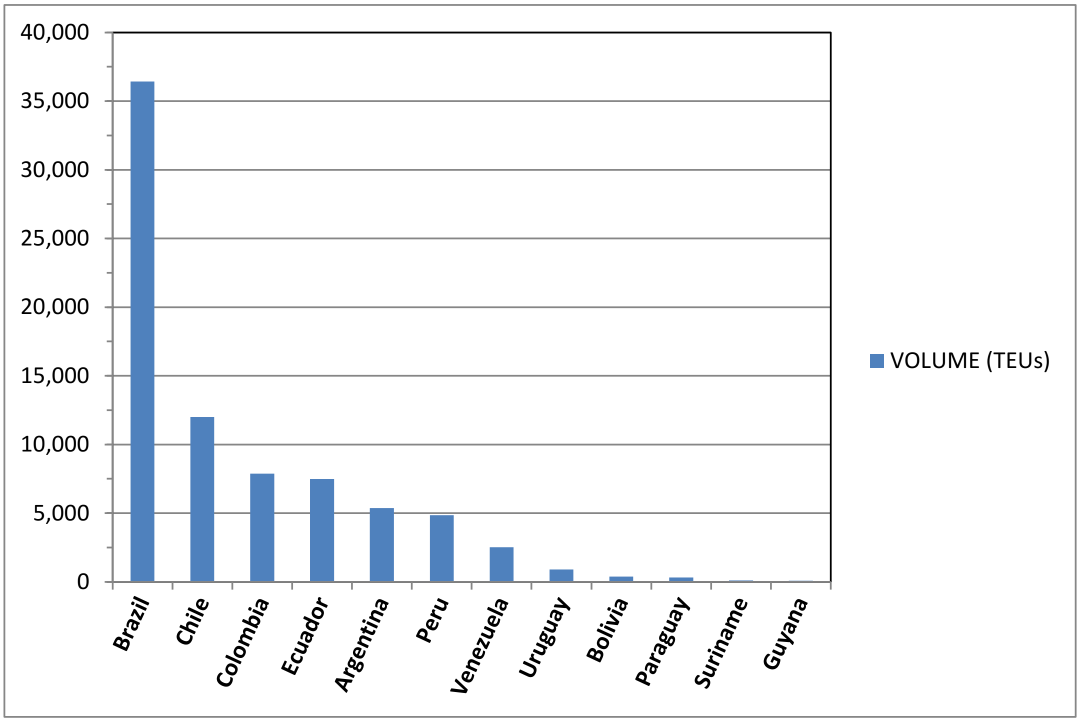


FIGURE 12: Total dataset – volume by country of shipment origin (based upon original shipment points)

TABLE 5: Largest port-of-export country shares of commodity category total volume

Largest Port-of-Export Country Shares of Commodity Category Total Volume													
Total Flow		**Commodity Category 1**		**Commodity Category 2**		**Commodity Category 3**		**Commodity Category 4**		**Commodity Category 5**		**Commodity Category 6**	
All Shipments		Perishable		Non-Perishable Time-Sensitive		Heavy Commodities/ Minimally Processed Natural Resources		Manufactured		High-Value Manufactured		Miscellaneous/ Unable to be Determined	
78,296.51 TEUs 680 flows		26,905.07 TEUs 391 flows		544.09 TEUs 45 flows		31,266.06 TEUs 346 flows		7,480.03 TEUs 196 flows		9,800.65 TEUs 272 flows		2,300.61 TEUs 190 flows	
Country	**%**	**Country**	**%**	**Country**	**%**	**Country**	**%**	**Country**	**%**	**Country**	**%**	**Country**	**%**
Brazil	46.7	Ecuador	25.1	Brazil	62.9	Brazil	54.6	Brazil	64.8	Brazil	71.4	Brazil	61.2
Chile	16.6	Brazil	21.9	Colombia	18.5	Chile	24.5	Colombia	12.4	Colombia	10.3	Argentina	10.0
Colombia	10.0	Chile	15.2	Argentina	11.8	Colombia	6.6	Chile	9.5	Peru	6.2	Colombia	9.0
Ecuador	9.6	Colombia	13.2	Uruguay	2.0	Argentina	4.8	Argentina	7.8	Venezuela	3.6	Chile	6.9
Argentina	6.5	Peru	12.2	Peru	1.8	Venezuela	4.7	Venezuela	2.8	Argentina	3.4	Peru	4.4
Peru	6.1	Argentina	8.8	Chile	1.3	Peru	2.0	Peru	1.8	Chile	3.3	Venezuela	4.3
Venezuela	3.2	Uruguay	1.9	Ecuador	1.3	Ecuador	1.6	Uruguay	0.5	Ecuador	1.2	Ecuador	3.3
Uruguay	1.2	Venezuela	1.5	Venezuela	0.4	Uruguay	0.9	Suriname	0.3	Uruguay	0.4	Uruguay	0.7
Suriname	0.1	Guyana	0.1	Guyana	0.0	Suriname	0.2	Ecuador	0.2	Guyana	0.1	Guyana	0.3
Guyana	0.1	Suriname	0.1	Suriname	0.0	Guyana	0.1	Guyana	0.0	Suriname	0.0	Suriname	0.0

TABLE 6: Top-five largest port shares of commodity category total volume

Top-Five Largest Port Shares of Commodity Category Total Volume													
Total Flow		**Commodity Category 1**		**Commodity Category 2**		**Commodity Category 3**		**Commodity Category 4**		**Commodity Category 5**		**Commodity Category 6**	
All Shipments		Perishable		Non-Perishable Time-Sensitive		Heavy Commodities/ Minimally Processed Nat. Resources		Manufactured		High-Value Manufactured		Miscellaneous/ Unable to be Determined	
78,296.51 TEUs 680 flows		26,905.07 TEUs 391 flows		544.09 TEUs 45 flows		31,266.06 TEUs 346 flows		7,480.03 TEUs 196 flows		9,800.65 TEUs 272 flows		2,300.61 TEUs 190 flows	
Port	**%**	**Port**	**%**	**Port**	**%**	**Port**	**%**	**Port**	**%**	**Port**	**%**	**Port**	**%**
Santos (Brazil)	15.5	Guayaquil-Puerto Bolivar (Ecuador)	24.9	Santos (Brazil)	33.2	San Vicente-Lirquen (Chile)	19.4	Santos (Brazil)	34.4	Santos (Brazil)	28.4	Santos (Brazil)	22.0
Guayaquil-Puerto Bolivar (Ecuador)	9.5	Callao (Peru)	10.1	Rio Grande (Brazil)	20.8	Santos (Brazil)	11.4	Salvador (Brazil)	9.5	Rio Grande (Brazil)	8.9	Buenos Aires (Argentina)	10.0
San Vicente-Lirquen (Chile)	8.5	Santos (Brazil)	9.3	Cartagena (Colombia)	17.8	Itajai (Brazil)	11.2	Cartagena (Colombia)	9.1	Itajai (Brazil)	8.1	Rio de Janeiro-Sepetiba Bay (Brazil)	9.8
Buenos Aires (Argentina)	6.3	Buenos Aires (Argentina)	8.5	Buenos Aires (Argentina)	11.8	Sao Francisco do Sul (Brazil)	8.8	Buenos Aires (Argentina)	7.8	Sao Francisco do Sul (Brazil)	6.9	Rio Grande (Brazil)	8.3
Itajai (Brazil)	5.9	San Antonio (Chile)	7.7	Sao Francisco do Sul (Brazil)	3.0	Vitoria (Brazil)	8.2	Rio de Janeiro-Sepetiba Bay (Brazil)	7.0	Cartagena (Colombia)	6.7	Itajai (Brazil)	6.3

RESULTS: EXISTENCE/LOCATION STEP

Part 1 of Research Questions #1 and #2 asks: How are the observed O-D flows by volume associated with actual distance between origins and destinations in the domestic/landside portion of the shipment path within the South American region - for the total dataset, and for each commodity category? This question is answered by the R^2 values obtained through regression estimation. For the total dataset (all shipments), as well as for all segments of the dataset (for each of the six commodity categories), the R^2 values range between 0.962 and 0.990. (These are presented in model summaries for each segment of the dataset in Appendix F.) Therefore, O-D flows by volume are highly associated with actual distance between origins and destinations in the domestic/landside portion of the shipment path within the South American region for all segments of the dataset, as would be expected.

The results for remaining research questions related to the reverse spatial interaction estimation – that is, Research Questions #1 (Part 2) and #2 (Part 2) – are answered in turn for each segment of the dataset, below. Reverse spatial interaction regression results (counts and volumes by segment and expected-to-actual distance ratio category) are presented in Table 7 for all data segments. Expected-to-actual distance (EtA) ratio ranges by commodity category are presented in Table 8.

It should be noted here that when "ratio" is used, this refers to EtA ratio (expected-to-actual distance ratio), unless otherwise noted. Where "higher-ratio flows" is used, it is inclusive of both the high-ratio (with EtA ratios over 1.10 but less than or equal to 2.00) and very high-ratio (those with EtA ratios over 2.00) EtA ratio flow subsets.

Proportionate flows are those with EtA ratios over 0.90 and less than or equal to 1.10.

Low-ratio flows are those with EtA ratios less than or equal to 0.90.

TABLE 7: Reverse spatial interaction estimation results - flows segmented by commodity type and EtA ratio category

DATA SEGMENT (AGGREGATED OBSERVED FLOWS)	FLOWS COUNT / VOLUME (TEUs)	PROPORTIONATE FLOWS			HIGH RATIO FLOWS COUNT / VOLUME (TEUs)	VERY HIGH RATIO FLOWS COUNT / VOLUME (TEUs)	HIGHER-RATIO FLOWS			LOW RATIO FLOWS		
		COUNT	VOLUME (TEUs)	PERCENT OF SEGMENT BY VOLUME			(HIGH + VERY HIGH) COUNT	VOLUME (TEUs)	PERCENT OF SEGMENT BY VOLUME	COUNT	VOLUME (TEUs)	PERCENT OF SEG-MENT BY VOLUME
All Shipment Flows	680 / 78,296.51	358	65,800.40	84.0 %	127 / 5,897.40	63 / 498.30	190	6,395.70	8.2 %	132	6,100.41	7.8 %
1: Perishable	391 / 26,905.07	235	23,946.53	89.0 %	71 / 1,666.56	26 / 59.90	97	1,726.46	6.4 %	59	1,232.08	4.6 %
2: Non-Perishable Time-Sensitive	45 / 544.09	28	534.39	98.2 %	10 / 6.73	2 / 0.16	12	6.89	1.3 %	5	2.81	0.5 %
3: Heavy/ Minimally Processed Nat. Resources	346 / 31,266.06	199	26,688.98	85.4 %	69 / 1,691.89	19 / 128.42	88	1,820.31	5.8 %	59	2,756.77	8.8 %
4: Manufactured	196 / 7,480.03	123	7,157.85	95.7%	37 / 175.30	11 / 9.44	48	184.74	2.5 %	25	137.44	1.8 %
5: High-Value Manufactured	272 / 9,800.65	139	7,825.45	79.8 %	67 / 980.85	16 / 118.17	83	1,099.02	11.2 %	50	876.18	8.9 %
6: Miscellaneous	190 / 2,300.61	101	1,891.57	82.2 %	44 / 185.93	15 / 17.04	59	202.97	8.8 %	30	206.07	9.0 %

TABLE 8: Expected-to-actual distance (EtA) ratio ranges by commodity category

Expected-to-Actual Distance (EtA) Ratio Ranges by Commodity Category													
Total Flow		Commodity Category 1		Commodity Category 2		Commodity Category 3		Commodity Category 4		Commodity Category 5		Commodity Category 6	
All Shipments		Perishable		Non-Perishable Time-Sensitive		Heavy Commodities/ Minimally Processed Nat. Resources		Manufactured		High-Value Manufactured		Miscellaneous/ Unable to be Determined	
Min:	0.00	Min:	0.00	Min:	0.74	Min:	0.07	Min:	0.20	Min:	0.08	Min:	0.23
Max.	19.13	Max.	50.71	Max.	13.69	Max.	18.42	Max.	14.06	Max.	44.68	Max.	12.06
Mean:	1.28	Mean:	1.35	Mean:	1.40	Mean:	1.22	Mean:	1.29	Mean:	1.30	Mean:	1.33
Median:	1.00	Median:	1.00	Median:	1.00	Median:	1.00	Median:	1.00	Median:	1.00	Median:	1.00

6.1. Results: Reverse Spatial Interaction Regression – Total Dataset (All Shipments)

Research Question #1, Part 2 asks: Are expected distance values disproportionately high as compared to actual physical distance in the domestic/landside portion of the shipment path within South America? This places the focus upon the high-ratio flows (flows with high expected distance as compared to actual distance), which are of the most interest for this study. High-ratio flows indicate where trade impedance is disproportionately high; as much of the literature related to trade in this region cites high transport costs/high trade barriers as major detriments to trade and the economy overall, it makes sense that identifying these flows is the primary concern. It is anticipated that expected distance values are not disproportionately high as compared to actual physical distance. (Stated differently, it is anticipated that expected distance values are proportionate to - or lesser than - actual physical distance.)

For the total flow dataset, the R^2 for the fully estimated model is 0.985. (Model summaries can be found in Appendix F.) The histogram for the total dataset (all shipments) (Fig. 13), displays the distribution (count) of flows by expected-to-actual (EtA) distance ratios. For 358 out of 680 aggregated observed flows by count, the null hypothesis must be accepted that expected distances are not disproportionately high, as compared to actual physical distance. These are the flows for which the expected-to-actual (EtA) distance ratio is between 0.90 and 1.10. Figure 14 demonstrates that these 358 proportionate flows comprise 52.6% of the dataset by count, but 84.0% of the dataset by volume (65,800.40 TEUs). Figure 15.a. presents a map displaying the total dataset of aggregated observed flows by volume (TEUs).

In comparison, 132 low-ratio flows (EtA distance ratios less than or equal to 0.90) comprise 19.4% of the dataset by count and 7.8% of the dataset by volume (6,100.41 TEUs). The 127 high-ratio flows comprise a comparable portion (7.5%) of the dataset by volume (5,897.40 TEUs). The portion of total dataset volume made up of very high-ratio flows is very low, at just 0.6% (498.30 TEUs), comprised of 63 flows (Figs. 13 and 14).

For all South American port-of-export countries (countries of the destination ports, not countries of shipment origin) in the dataset, the vast majority of their total dataset export flow volume is comprised of flows with proportionate EtA ratios (Fig. 14). All port-of-export countries in the dataset have percentages ranging from 75.0% to 99.1% of their export flow totals in the proportionate EtA ratio category (between 0.90 and 1.10). The port-of-export countries with the lowest percentages of flow totals in the proportionate ratio category are those of Brazil (75.0% of export flow volume) and Colombia (80.4% of export flow volume); those with the highest percentages of flow totals in the proportionate ratio category are Peru (99.0% of export flow volume) and Venezuela (99.1% of export flow volume).

In viewing the maps of the EtA ratios of the total dataset (Figs.15.a.-d.), it is visually apparent that the proportionate EtA ratio flows (Fig. 15.d.) generally appear to reflect the flow pattern of the overall dataset – with a few key exceptions. These exceptions include:

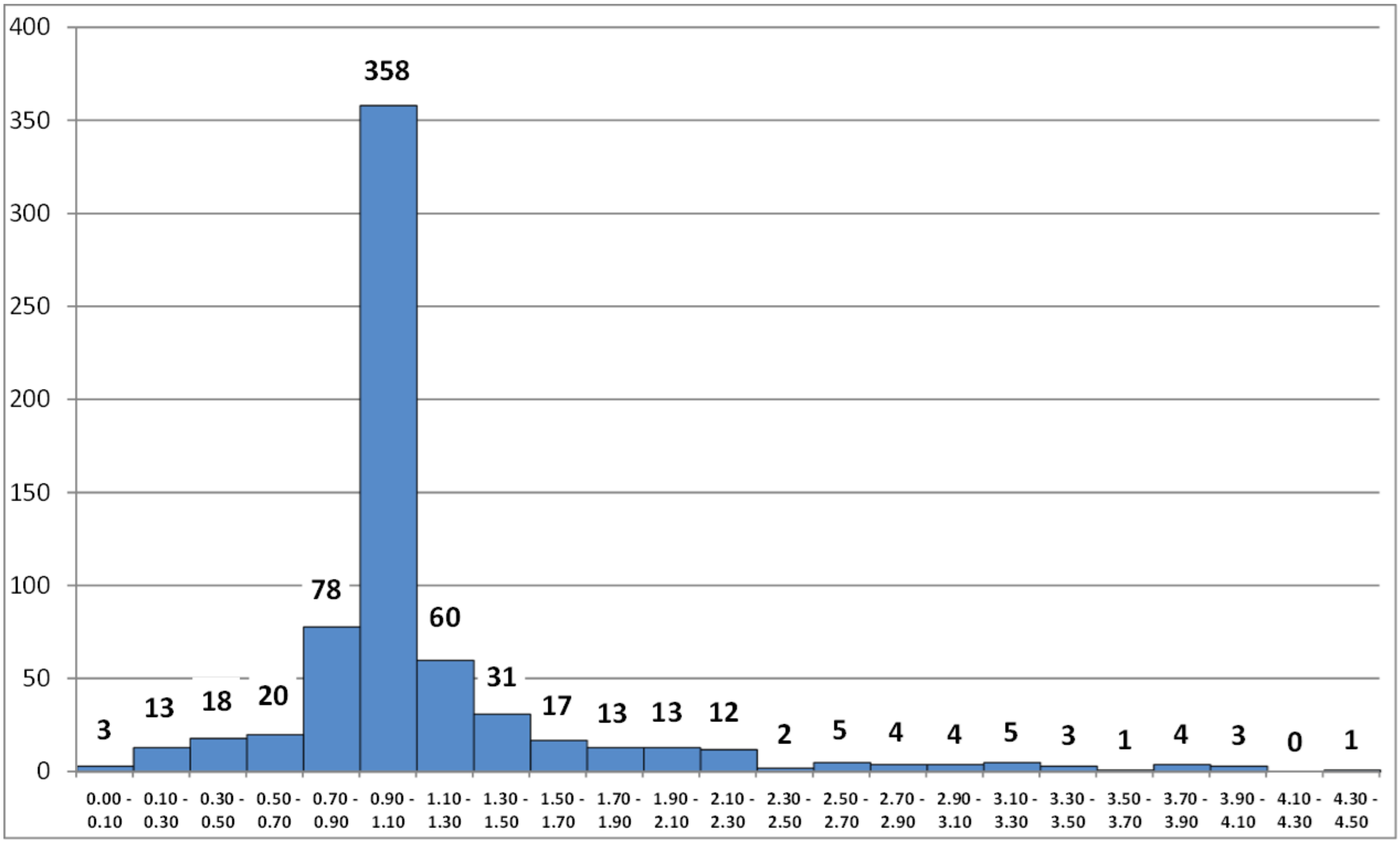


FIGURE 13: Histogram: total dataset (all shipments) – 680 aggregated observed flows by count *
*12 observations (aggregated observed flows) out of 680 have values ranging between 4.50 - 19.13; these observations are not displayed on the histogram.

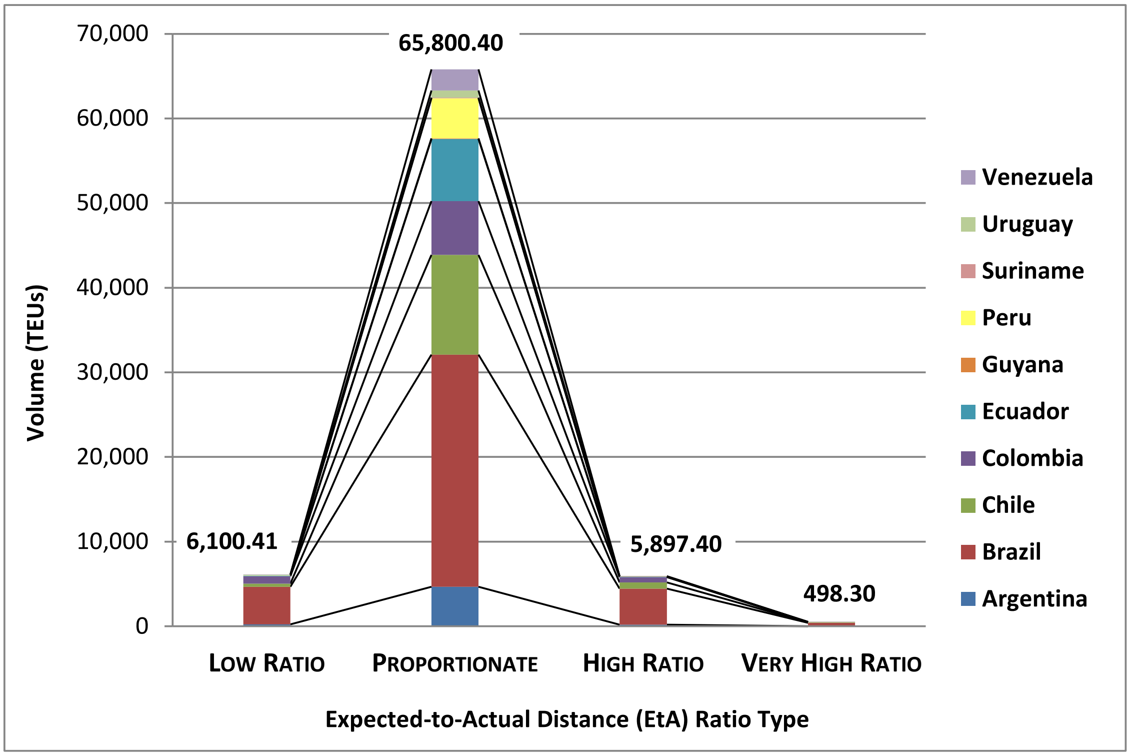


FIGURE 14: Total dataset (all shipments) – volume and port of export country by EtA ratio type

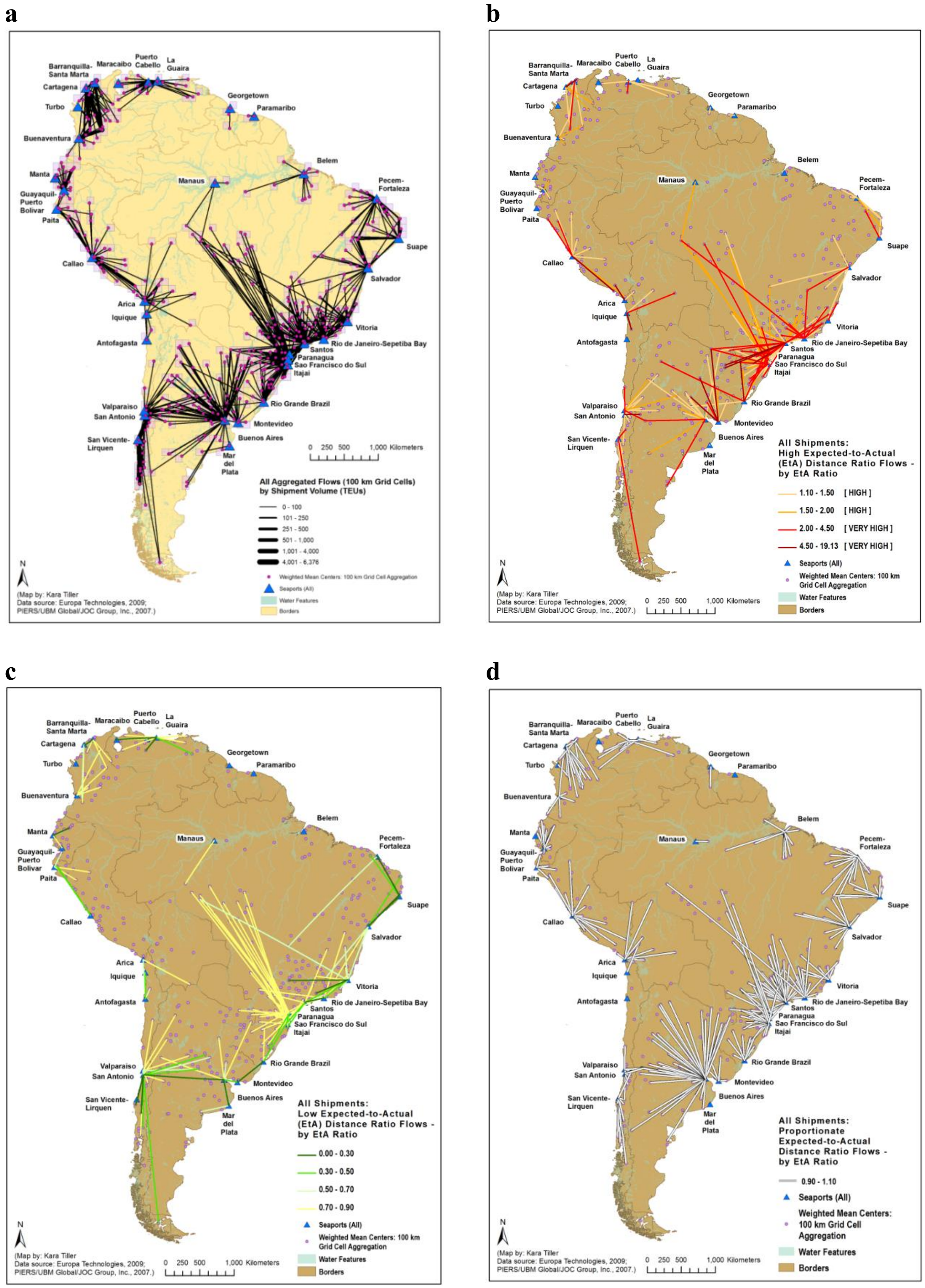


FIGURE 15: Reverse spatial interaction regression results – total dataset (all shipments): (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

a) The longer flows from the wood producing region of the Brazilian interior are not as well-represented in the proportionate category (or in the higher-ratio category, Fig. 15.b.). Although the higher-ratio category *appears* to include more of the flows hailing from this region, the proportionate flow total and the higher-ratio flow total from this region each sum to approximately 38.00 TEU. In comparison, in the low-ratio category, these regional flows are more numerous and voluminous - over 4.5 times that, at 171.95 TEUs.

b) The average distance for proportionate flows is the shortest among ratio categories, at 402.90 km, followed by the average distance for high-ratio flows (472.89 km); very high-ratio flows (586.16 km) - with the longest being that for low-ratio flows (671.39 km).

c) The average volume of proportionate ratio flows is far higher than those of other categories, at 183.80 TEUs. This is followed by the average volume of high-ratio flows (46.44 TEUs); low-ratio flows (46.22 TEUs); and very high-ratio flows (7.91 TEUs).

d) In the southeastern/south region of Brazil, spanning the area from the ports of Rio Grande to Vitoria (and including Itajai, Sao Francisco do Sul, Santos, and pseudo-port Rio de Janeiro-Sepetiba Bay), the general angle of proportionate flows (Fig. 15.d.) is one of movement from the interior toward the coast (southeast-directed). However, this is different from the orientation of higher-ratio flows (Fig. 15.b.) in that region - which generally are at southwest/northeast oriented angles. Some of the flows display coastwise movement, but more generally, the pattern appears to represent flows which bypass or divert away from nearest ports. The low-ratio flows in this area have an even more pronounced coastwise orientation, especially for the lower-ratio flows among these. (However, the low-ratio flows with EtA ratios ranging from 0.70 to 0.90 are more similar to interior-to-coast type flows, and include the majority of export volume associated with the longer flows from Brazil's wood-producing region.)

The ports in this region of Brazil are also the largest players in terms of the volume of higher-ratio flows for the total dataset: Vitoria; the Rio de Janeiro-Sepetiba Bay pseudo-port; Santos; Paranagua; Sao Francisco do Sul; and Itajai. It appears that many of these flows are bypassing nearer ports in order to reach Santos (Fig. 15.b.) While Santos itself accounts for less of the higher-ratio flow volume (3.9% or 250.66 TEUs) than do most of the other ports above, the combination of higher volumes of higher-ratio flows to the other southeastern ports - combined with the numerous visually apparent (though lower total volume) higher-ratio flows bypassing them to reach Santos – seems worthy of further investigation. When considering the *count* of aggregated higher-ratio flows to each port of export, Santos is the port of export destination for a much higher *number* of aggregated higher-ratio flows (26 flows) than any other port. The second-highest number of higher-ratio flows is destined for pseudo-port Rio de Janeiro-Sepetiba Bay (16 flows).

One would assume initially that this port-bypass situation could simply be explained by the fact that Santos is a larger port with a larger hinterland. However, the larger extent of the Santos hinterland (without regard to the longer flows from Brazil's wood-producing region) is easily viewed in the map of proportionate ratio flows. Here, the hinterland is delineated as larger than that of other ports - *but without the same degree of near-port bypass* (Figs. 16.a. and b.). These figures make it apparent that the proportionate flows versus higher-ratio flows exporting from Santos have differing origins, and thus the spatial distribution of flows in each category differs in regard to this port.

However, higher-ratio flows exporting from Santos are more lateral, hailing from the west/southwest as far as Asuncion, Paraguay and western Misiones Province, Argentina - and as far northeast as the state of Espirito Santo, Brazil. The difference in the spatial distribution of proportionate-ratio flows (Fig. 16.a.) versus higher-ratio flows (Fig. 16.b.) is clearly illustrated. It is anticipated that the results of the proposed research extension (Step 4/Association Step) may provide insight into this result.

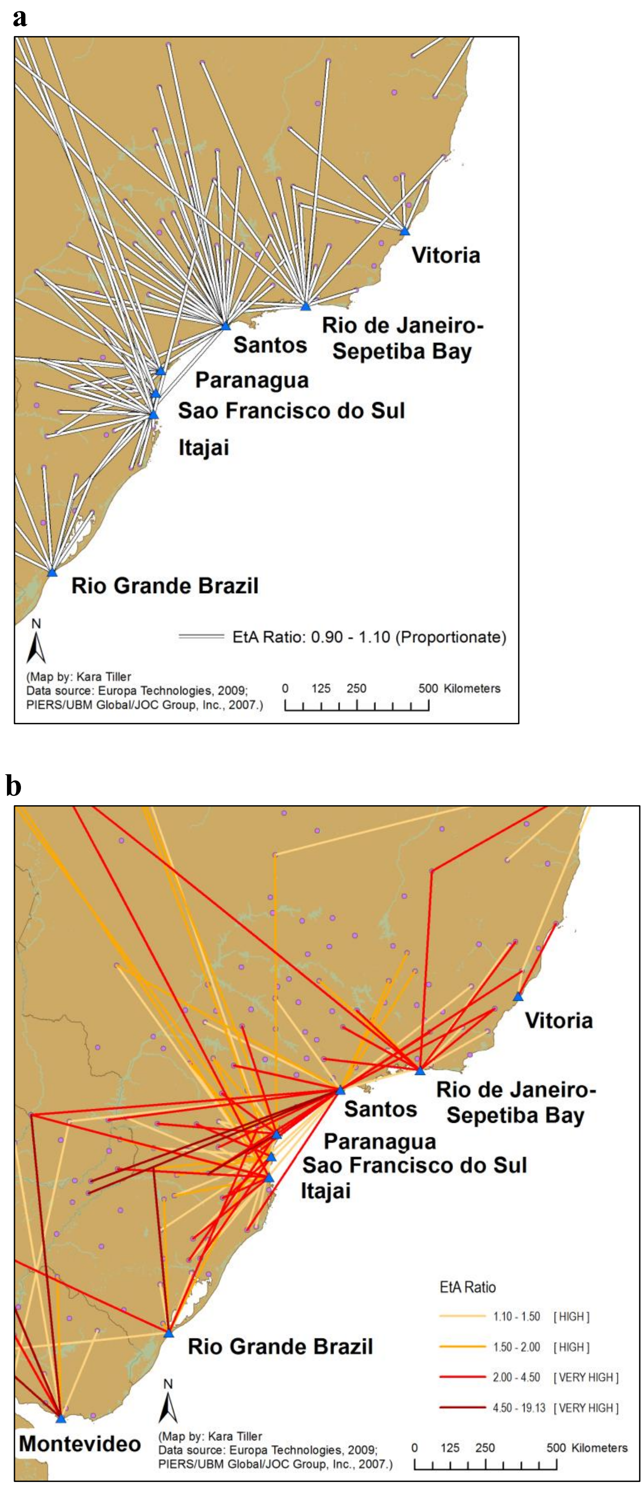


FIGURE 16: Ports of southeastern Brazil - reverse spatial interaction regression results - total dataset (all shipments): (a.) proportionate flows (b.) higher-ratio flows

e) In general, it appears at first sight that the hinterlands for the proportionate ratio flows are more singular (associated with one port of export) than hinterlands for higher-ratio flows. In other words, without assuming causation or direction of causality for such, the high EtA ratios and hinterland overlap of flows appear to occur together, whereas hinterland overlap/near-port bypass is far less evident in the proportionate-ratio flow category for the total flow dataset.

f) Additionally, it appears that for the total flow dataset (all shipments), the lowest values among the low-ratio flows could potentially be associated with coastwise moves (an exception to this being in terms of Chilean flows). This suggests that, as mode and route cannot be verified from the PIERS database/manifest data values, it is possible that these may be feedered waterborne flows – not representative of landside travel. If this is the case, it may indicate that a substantially lower portion of landside travel is actually low-ratio (as the lowest of the current low-ratio flows may involve a non-landside mode of transport). Creative means of determining mode from the data could be investigated in the future, as the pre-export mode(s) of travel are not provided in the data.

As it is theorized that expected distances are not disproportionately high as compared to actual physical distance, the null hypothesis must also be accepted for the 132 low-ratio flows (Fig. 13) comprising 7.8% of the dataset by volume (6,100.41 TEUs) (Fig. 14). It is visually apparent in the map of low-ratio flows (Fig. 15.c.) that of the flows with the lowest ratios (expected distance 70% or less of the actual distance), generally these flows fall into three typologies. Either they are: (1) movements along the coastline (or suggest coastwise movements); (2) they are Chilean flows; or, (3) they are flows originating from Argentina but exporting from the Western coast of South America via Chilean ports. Low-ratio flows with EtA ratios ranging from 0.00 to 0.70 have a mean flow distance of 606.0 km and a mean flow volume of 29.14 TEU - whereas low-ratio flows with EtA ratios ranging from 0.70 to 0.90 (closer to the proportionate range) have higher mean flow distance (716.7 km) and higher mean flow volume (58.04 TEUs).

The set of 190 higher-ratio flows (8.2% of the dataset by volume at 6,395.70 TEUs) is of the most interest for this analysis (Fig. 14). Although in Fig. 14, high and very high ratios are segmented separately, the map in Fig. 15.b. presents higher-flows ratios more generally, with the flows in orange (EtA ratios ranging from 1.10 to 2.00) labeled as high, whereas the flows in red (EtA ratios over 2.00) are labeled as very high. (As a group they are discussed as higher-ratio flows unless differentiated/compared.) For these 190 flows, the null hypothesis must be rejected and the alternative hypothesis must be accepted that expected distances are disproportionately high as compared to actual physical distance.

It is interesting to view the results of high-ratio flows in terms of ports. When high- and very high-ratio flows are taken together under the general banner of higher-ratio flows, seven ports emerge as those exporting sizable portions of the higher-ratio flow volume. These seven ports alone export nearly 80% of the higher-ratio flow volume. They range from exporting 28.0% to 5.4% of higher-ratio flow volume, in this order: Vitoria (Brazil); Sao Francisco do Sul (Brazil); Itajai (Brazil); Valparaiso (Chile);

the Rio de Janeiro-Sepetiba Bay pseudo-port (Brazil); Paranagua (Brazil); and Buenaventura (Colombia). These differ, for the most part, from the top-five port shares for the overall dataset (all flows/all ratios) listed in Table 6; only Itajai is a top port for all flows and higher-ratio flows alone.

The port of export with the largest proportion of flow volume among higher ratio flows is Vitoria, Brazil, which exported 28.0% of higher-ratio flow volume; however, this is not apparent on the map. This is due to the fact that one of the three aggregated flows comprising this figure is a short (11.7 km), high-volume flow (1,687.10 TEUs). It is further notable that this flow has an EtA ratio of 1.11, near the bottom of the higher-flow range.

When very high ratio flows are explored separately, 79.1% of the very high-ratio flow volume is exported from five ports: the Rio de Janeiro-Sepetiba Bay pseudo-port (Brazil); Itajai (Brazil); Santos (Brazil); Paranagua (Brazil); and the San Vicente-Lirquen pseudo-port (Chile). It is important to consider, however, that the total export flow volume for the very high-ratio portion of the total dataset amounts to just 498.30 TEUs.

As noted in the total flow dataset histogram (Fig. 13), 12 aggregated flows out of 680 have extremely high EtA ratio values, ranging between 4.50 and 19.13, which are not displayed on the histogram. These are all very low-volume flows; none of these flows is over 2.55 TEUs. The most extreme flows (EtA ratios > 10.00), are flows exporting most commonly from Montevideo (Uruguay), with one flow exporting from Arica (Chile). Other ports with flows featuring in the extreme high-ratio group include (with ratios from 8.25 to 5.01, respectively): Iquique (Chile); Callao (Peru); Cartagena (Colombia); Paranagua and Santos (Brazil); San Antonio (Chile); and Rio Grande (Brazil).

6.2. Results: Reverse Spatial Interaction Regression – Commodity Category 1: Perishable Products

For Commodity Category 1 (Perishable Products), the R^2 for the fully estimated model is 0.984. For perishable products, it is anticipated that expected distance values are not disproportionately high as compared to actual physical distance. The histogram for the Commodity Category 1/Perishable Products dataset (Fig. 17), displays the distribution of flows by EtA distance ratio. For 294 out of 391 aggregated observed flows by count, the null hypothesis must be accepted that for this data segment, expected distances are not disproportionately high as compared to actual physical distance. Flows with proportionate EtA ratios comprise 89.0% of the dataset by volume (23,946.53 TEUs), and low-ratio flows comprise 4.6% of the dataset by volume (1,232.08 TEUs) (Fig. 18); thus 93.6% of the dataset by volume is demonstrated not to have disproportionately high expected distances.

All port-of-export countries in the data segment have percentages ranging from 70.8% to 100.0% of their export flow in the proportionate ratio category. The lowest percentages of port-of-export country flow totals in the proportionate ratio category are those of Colombia (70.8% of export flow volume) and Chile (81.1% of export flow volume). The highest percentages in the proportionate ratio category are claimed by Guyana and Suriname, both of which have 100.0% of export flow volume in the proportionate ratio category (though it should be noted that these are very low total export volumes – 27.00 and 19.58 TEUs, respectively).

In viewing the maps of the flow volume (Fig. 19.a.) and EtA ratios (Figs.19.b.-d.) for perishable products, just as in the total dataset analysis, it is apparent that the proportionate EtA ratio flows (Fig. 19.d.) generally represent the flow pattern of the overall dataset, with exceptions. This includes the very different spatial arrangement of flows by ratio level for the southeastern ports of Brazil. As noted in the total dataset analysis, the angle of proportionate flows (Fig. 19.d.) is generally one of movement from the interior toward the coast (directed ether south or southeast). Again, this is different from the orientation of higher-ratio flows (Fig. 19.b.) in that region, which generally are at southwest/northeast oriented angles. Though this is demonstrated in the maps, it is essential to reference this to volume, noting that the higher-ratio flows exporting from Santos (for example) for this data segment total just 58.20 TEUs, whereas proportionate ratio flows exporting from Santos for this data segment total 2,426.99 TEUs. Still, the difference in pattern is unmistakable.

Additionally, one difference notable from the proportionate flow pattern for this data segment is that longer flows associated with particular ports are low-ratio flows. These include Chilean ports of Valparaiso and San Antonio, as well as longer flows exporting from Buenaventura (Colombia), and southeastern Brazilian ports of Paranagua and Itajai, as well as northeastern Brazilian ports Pecem-Fortaleza (pseudo-port), Suape, Salvador, and Vitoria.

In general, the flow length in this data segment does not vary as much as flow length in some other data segments; still, differences between ratio categories exist in terms of their mean length and volume of flows. As in the total dataset, for perishable products, proportionate flows have a far higher mean flow volume (101.90 TEUs) than do other flow categories – next is high-ratio flows at 23.47 TEUs, followed by low-ratio flows at 20.88 TEUs. As is the usual result in this analysis, the very high-ratio flow category has a very low mean flow volume (2.30 TEUs). For the perishable products, it is high-ratio flows which have the shortest mean flow distance among categories (362.62 km), followed by proportionate flows (377.28 km), very high-ratio flows (397.75 km); and low-ratio flows (518.04 km).

Only 97 flows are higher-ratio, totaling 6.4% of data segment volume (1,726.46 TEUs). For these flows, the null hypothesis must be rejected, and the alternative hypothesis must be accepted that expected distances are disproportionately high as compared to actual physical distance.

The high ratio subset (EtA ratios between 1.10 and 2.00) has the following ports as the top exporters of high-ratio perishable products volume: Valparaiso (Chile) (24.0%); Cartagena (17.6%), Buenaventura (11.5%), and pseudo-port Barranquilla-Santa Marta (6.4%) (Colombia); San Antonio (Chile) (5.9%); and Mar del Plata (Argentina) (5.4%). Again these differ from the top-five port shares for the overall segment, as listed in Table 6. The very high ratio subset is very low-volume (59.90 TEUs), its largest export ports include Suape (Brazil); Callao (Peru); and pseudo port Rio de Janeiro-Sepetiba Bay (Brazil).

The most extreme values among higher ratio flows (those with an EtA ratio over 4.50) range from 5.65 to 50.71 for the perishable products segment, and involve ports of export Montevideo (Uruguay); Buenos Aires (Argentina); San Antonio (Chile); and Santos (Brazil).

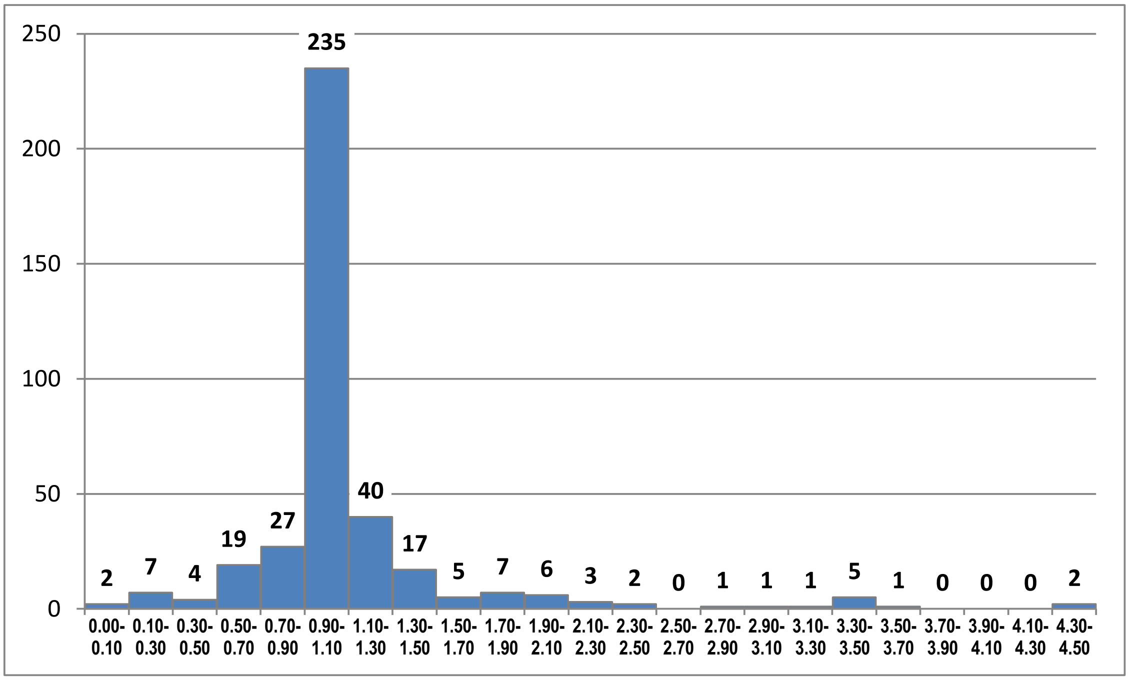


FIGURE 17: Histogram: Commodity Category 1 - Perishable Products - 391 aggregated observed flows by count*

* 6 observations (aggregated observed flows) out of 391 have values ranging between 4.50 – 50.72; these observations are not displayed on the histogram.

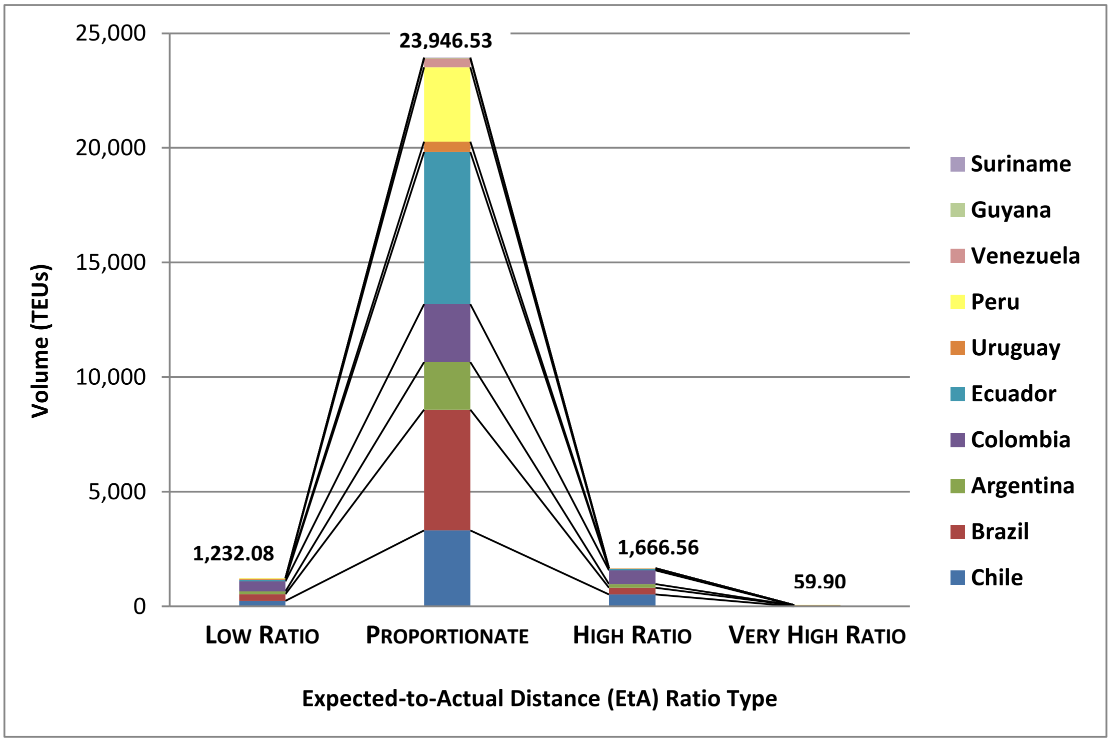


FIGURE 18: Commodity Category 1 - Perishable Products: volume and port of export country by EtA ratio type

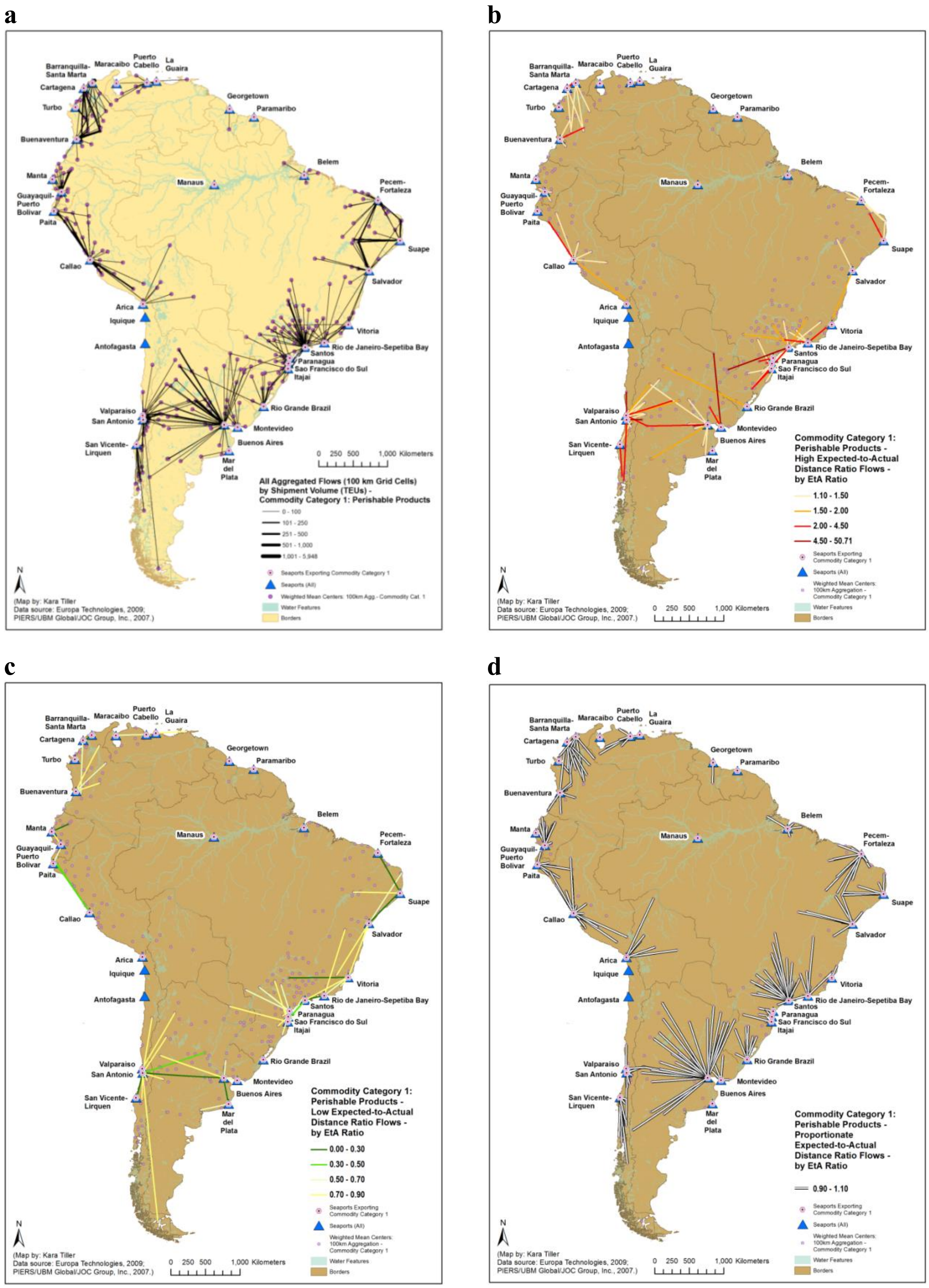


FIGURE 19: Reverse spatial interaction regression results - Commodity Category 1: Perishable Products: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.3. Results: Reverse Spatial Interaction Regression – Commodity Category 2: Non-Perishable Time-Sensitive Products

For the Commodity Category 2 (Non-Perishable Time-Sensitive Products), the R^2 for the fully estimated model is 0.990. For non-perishable time-sensitive products, it is anticipated that expected distance values are not disproportionately high as compared to actual physical distance. The histogram for the data segment (Fig. 20) displays the distribution of flows by EtA distance ratio. For 33 out of 45 aggregated observed flows by count, the null hypothesis must be accepted that for non-perishable time-sensitive products, expected distances are not disproportionately high as compared to actual physical distance.

Flows with proportionate EtA ratios comprise 98.2% of the dataset by volume (534.39 TEUs), and low-ratio flows comprise 0.5% of the dataset by volume (2.81 TEUs) (Fig. 20); thus 98.7% of the non-perishable time sensitive products data segment by volume is demonstrated not to have disproportionately high expected distances. Again, just as in the analyses of previous segments, the largest ratio category for these flows is the proportionate ratio category (Fig. 21; Fig. 22.d.). All port-of-export countries in the data segment have percentages ranging from 73.9% (Uruguay) to 100.0% (Colombia, Ecuador, and Venezuela) of their export flow in the proportionate ratio category.

In terms of the five low-ratio flows (Fig. 22.c.) for Non-Perishable Time-Sensitive Products, all are associated with the ports of Buenos Aires (Argentina) and Montevideo (Uruguay).

The smallest ratio category for Non-Perishable Time-Sensitive products is that of the 12 higher-ratio flows (totaling 6.89 TEUs) (Fig. 22.b.). For these 12 flows, the null hypothesis must be rejected, and the alternative hypothesis accepted that expected distances are disproportionately high as compared to actual physical distance. The flows with the highest EtA ratios export from Santos (Brazil) (13.69) and Valparaiso (Chile) (2.56). All higher-ratio flows are very low-volume.

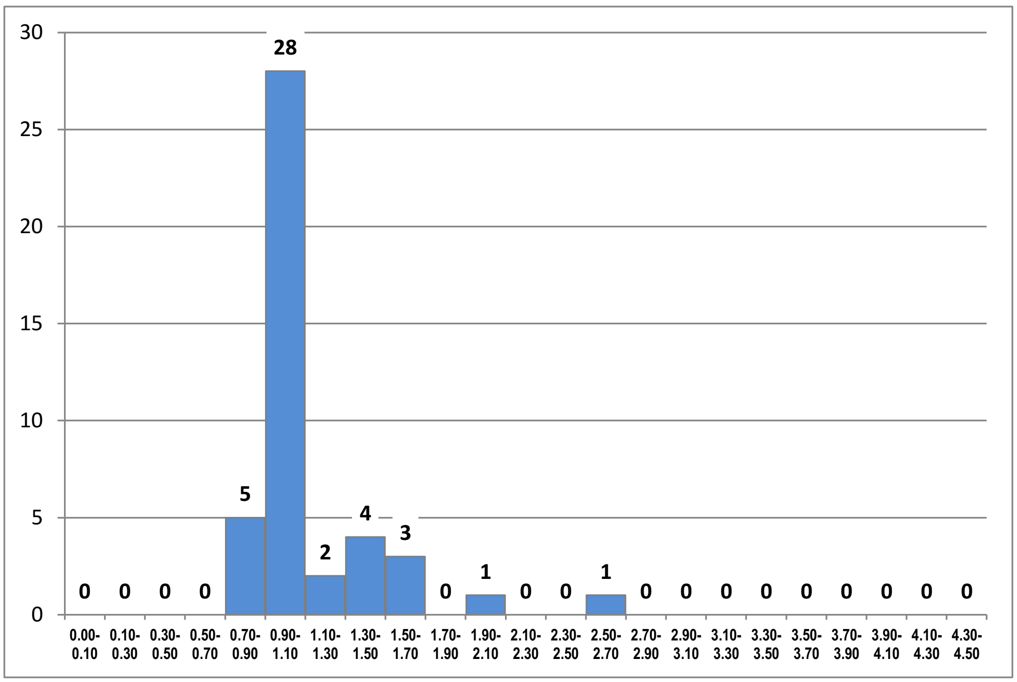


FIGURE 20: Histogram: Commodity Category 2 – Non-Perishable Time-Sensitive Products – 45 aggregated observed flows by count *
*1 observation (aggregated observed flow) out of 45 has a value of 13.69; this observation is not displayed on the histogram.

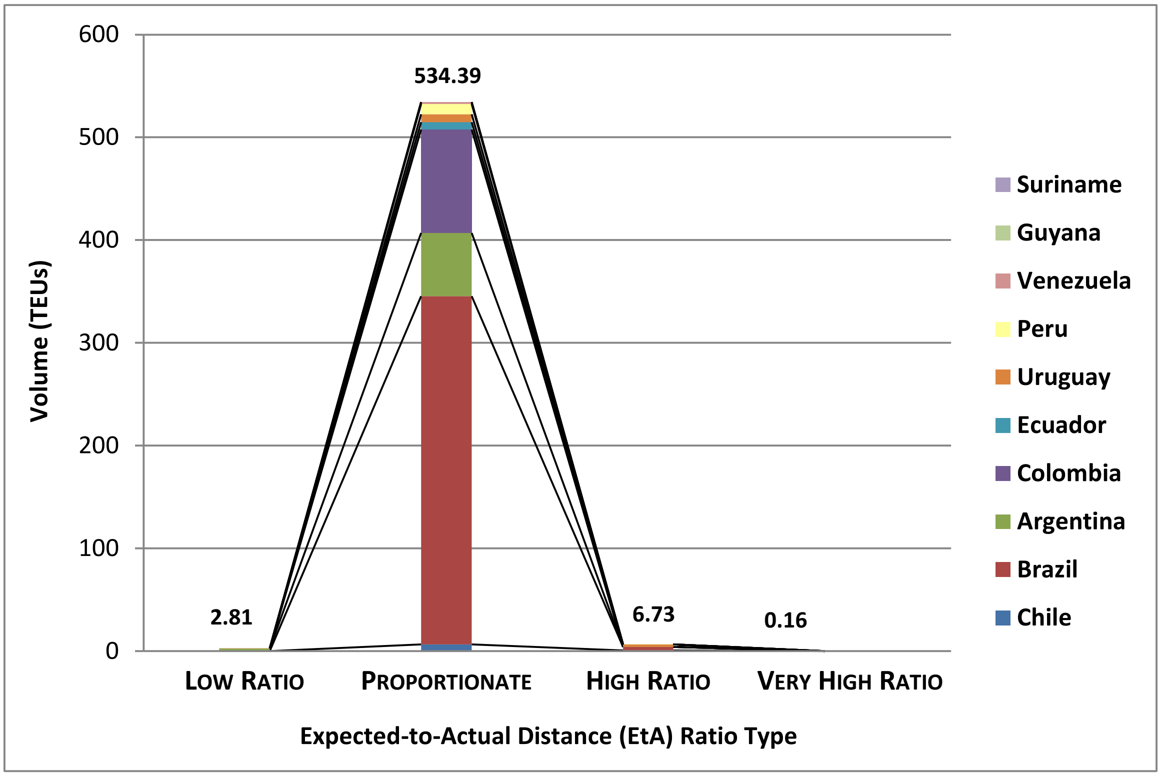


FIGURE 21: Commodity Category 2 - Non-Perishable Time-Sensitive Products: volume and port of export country by EtA ratio type

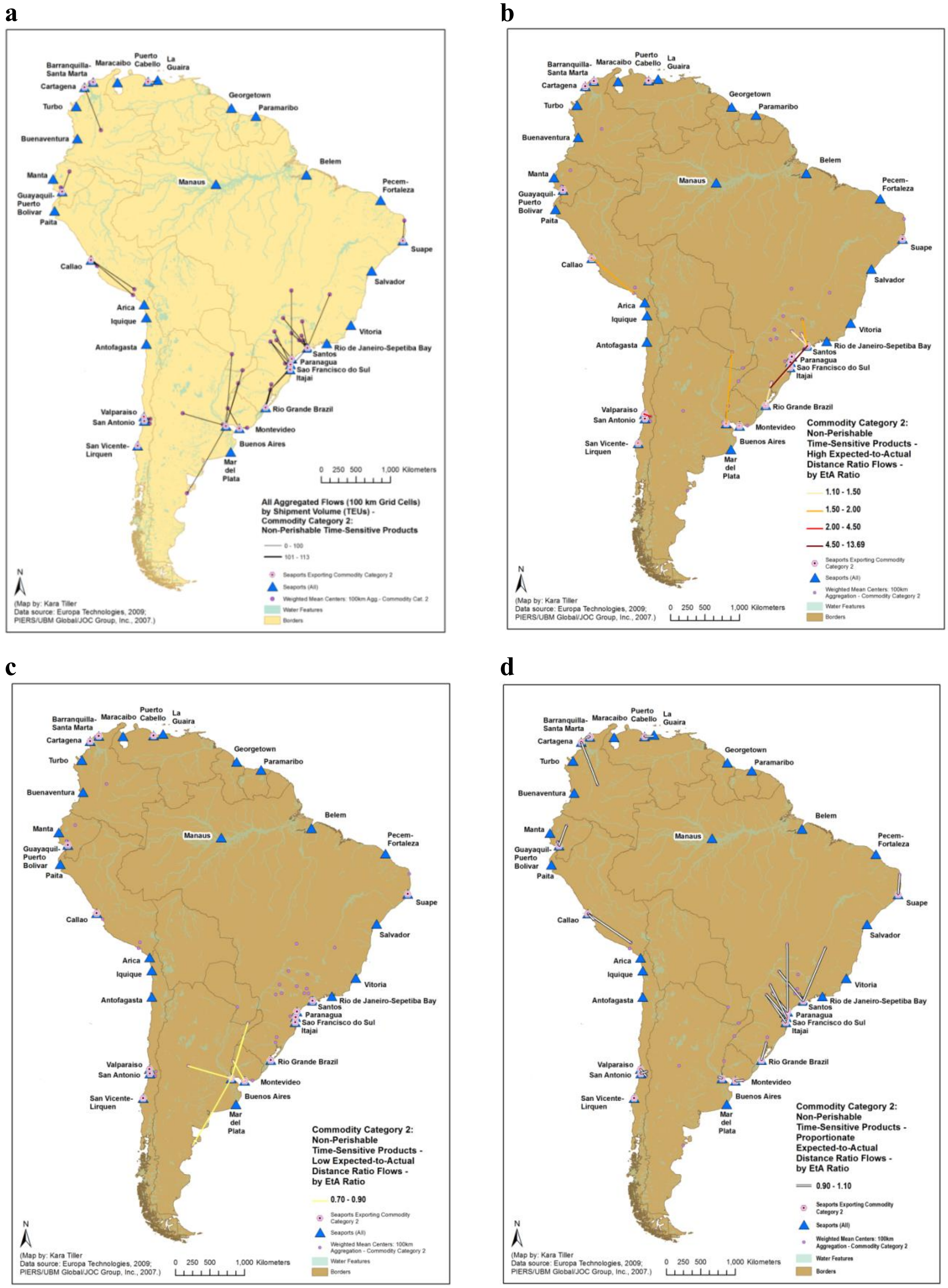


FIGURE 22: Reverse spatial interaction regression results - Commodity Category 2: Non-Perishable Time-Sensitive Products: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.4. Results: Reverse Spatial Interaction Regression – Commodity Category 3: Heavy Commodities/Minimally Processed Natural Resources

For Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources), the R^2 for the fully estimated model is 0.987. Unlike the data segments analyzed up to this point, for heavy/minimally processed commodities, it is anticipated that expected distance values *are* disproportionately high as compared to actual physical distance. However, for the majority of the flows (258 proportionate- and low-ratio flows), the null hypothesis must be accepted that expected distance is not disproportionately high, as compared to actual physical distance (Fig. 23).

The proportionate ratio category (199 flows comprising 85.4% of the data segment volume at 26,688.98 TEUs [Fig. 24]) and the low-ratio category (59 flows comprising 8.8% of the data segment volume at 2,756.77 TEUs) represent the majority of flow count and volume for the data segment. The higher-ratio flows category has a greater flow count (88) than the low-ratio flows – however, the volume total for higher-ratio flows is the least among categories (1,820.31 TEUs) (5.8% of the data segment volume). Thus only for these 88 higher-ratio flows, the null hypothesis must be rejected, and the alternative hypothesis accepted that expected distances are disproportionately high as compared to actual physical distance.

All port-of-export countries in the data segment have percentages ranging from 74.2% (for Brazil) to 100.0% (for Peru and Ecuador) of country flow volume total in the proportionate ratio category. Among categories, the proportionate category has the lowest mean flow distance (412.6 km) and the highest mean flow volume (134.12 TEUs) (Fig. 25.d.); this mean flow volume is far higher than the mean volume for other ratio categories. (The low-ratio category has the highest mean flow distance at 612.3 km and a mean flow volume of 46.72 TEUs, and the higher-ratio category has a mean flow distance of 481.2 km and the lowest mean flow volume among ratio categories, at 20.69 TEUs.)

Regarding the higher-ratio flows, 95.5% of heavy commodity/minimally processed natural resource higher-ratio flows are exported through Brazilian ports. (Yet Brazil's overall portion of the heavy commodity/minimally processed natural resource data segment total volume is only 54.6%.) This is apparent in comparing the map of flow volume (Fig. 25.a.) to the map of higher-ratio flows (Fig. 25.b.).

Itajai exported 34.8% of higher-ratio flow volume for the data segment, followed by Sao Francisco do Sul (22.0%), Paranagua (17.1%), pseudo-port Rio de Janeiro-Sepetiba Bay (11.5%), Santos (6.2%), and Salvador (2.9%). For each of these ports, their shares of the higher-ratio flow volume is much higher than their share of the total flow volume – with the exception of Santos.

The extreme high-ratio flows, ranging from 18.42 to 4.55, exported (respectively) from Cartagena, Itajai, pseudo port San Vicente-Lirquen (the highest-volume exporter for the data segment), Paranagua, and a second flow exporting from Itajai.

Overall, 96.3% of heavy commodity/minimally processed natural resource lower-ratio flows are exported through Brazilian ports. It is notable that Sao Francisco do Sul exported an even higher proportion of the low-ratio flows (23.9%) than it exported for the higher-ratio flows, and that several of the ports noted just above also exported sizable portions of low-ratio flow volume. Vitoria exported the largest share of low-ratio flow volume for the heavy commodities/minimally processed natural resources data segment, at 30.9% of the category volume. The spatial distribution of low-ratio flows (Fig. 25.c.) does not appear to differ greatly from that of the higher-ratio flows, with the exception of flow length differences, and the flows to Vitoria being somewhat more apparent in the low-ratio flows map.

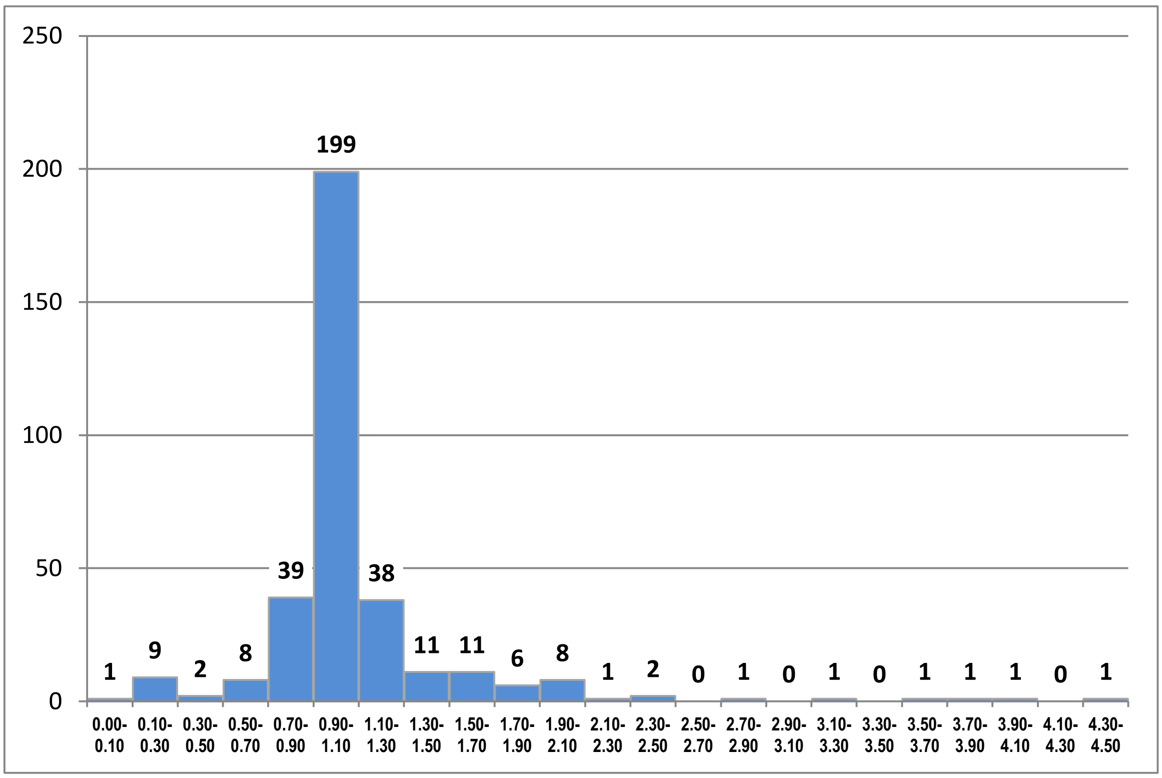


FIGURE 23: Histogram: Commodity Category 3: Heavy Commodities/ Minimally Processed Natural Resources – 346 aggregated observed flows by count *
*5 observations (aggregated observed flows) out of 346 have values ranging between 4.55 - 18.42; these observations are not displayed on the histogram.

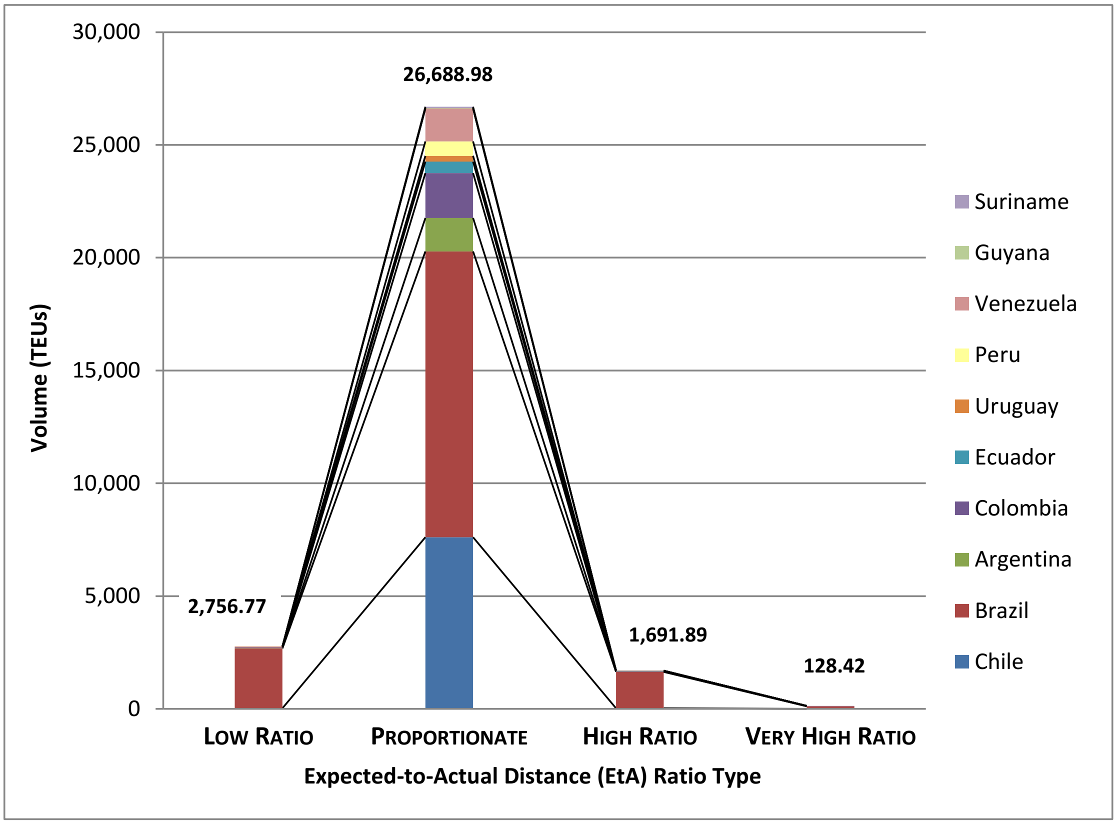


FIGURE 24: Commodity Category 3 - Heavy Commodities/Minimally Processed Natural Resources: volume and port of export country by EtA ratio type

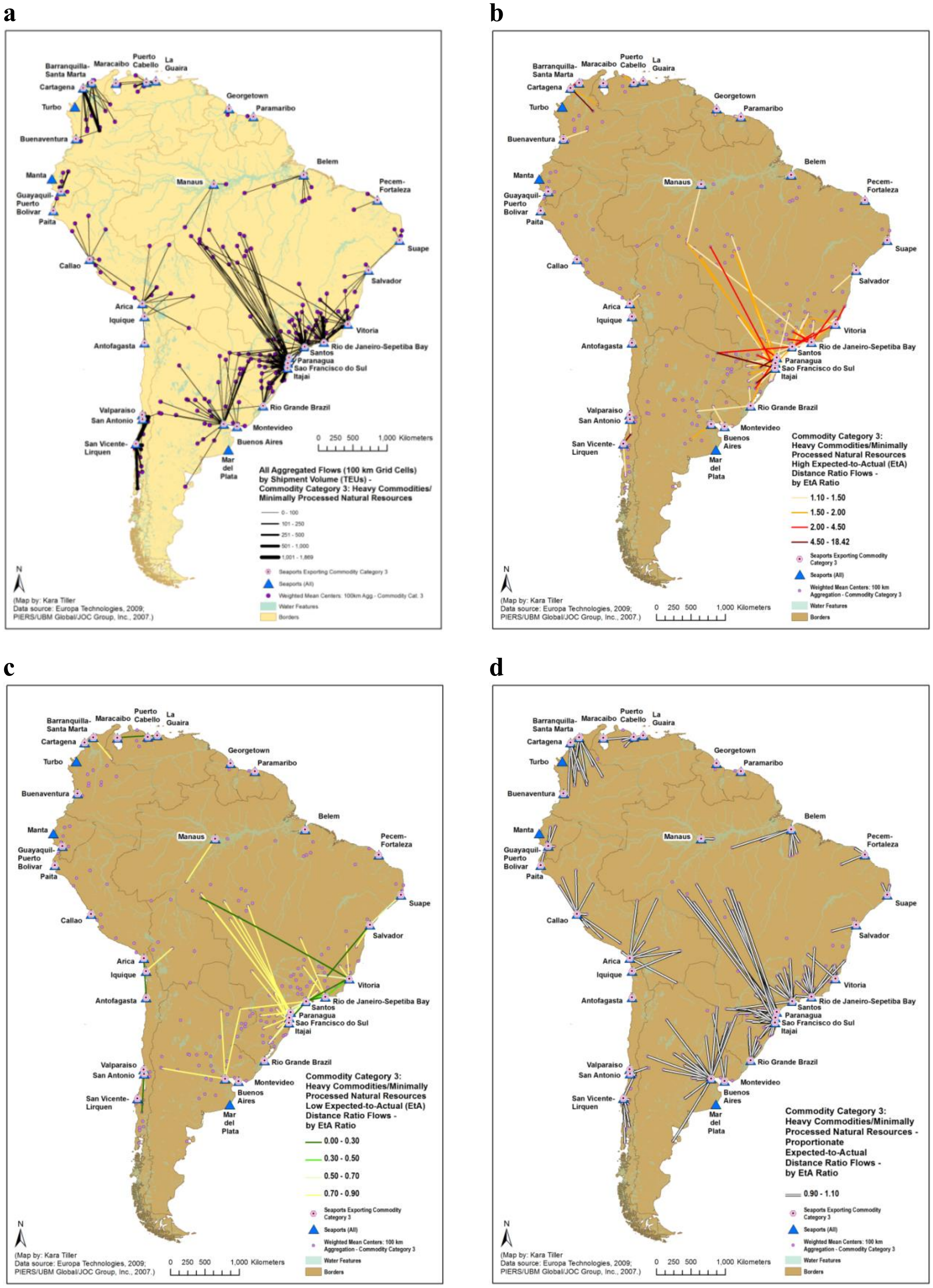


FIGURE 25: Reverse spatial interaction regression results - Commodity Category 3: Heavy Commodities/ Minimally Processed Natural Resources: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.5. Results: Reverse Spatial Interaction Regression – Commodity Category 4: Manufactured Products

For Commodity Category 4 (Manufactured Products), the R^2 for the fully estimated model is 0.988. For manufactured products, it is anticipated that expected distance values are not disproportionately high as compared to actual physical distance. The histogram for the Commodity Category 4/Manufactured Products data segment (Fig. 26), displays the distribution of flows by EtA distance ratio. For 148 out of 196 flows by count, the null hypothesis must be accepted that for the Commodity Category 4/Manufactured Products data segment, expected distances are not disproportionately high as compared to actual physical distance.

The proportionate ratio category comprises 95.7% of the segment by volume (123 flows totaling 7,157.85 TEUs) (Fig. 27). All port-of-export countries in the data segment have percentages ranging from 90.1% (for Ecuador) to 100.0% (for Suriname) of country flow volume total in the proportionate ratio category.

In comparing the map of all flows in the dataset by volume (Fig. 28.a.) to the maps of flows by ratio category (Figs. 28.b.–d.), it is immediately apparent that the set of flows in the proportionate ratio category (Fig. 28.d.) comprises the overwhelming majority of flow volume for the manufactured products data segment, and thus appears quite similar to the spatial distribution of the set of all flows in the dataset.

However, the higher-ratio flows (Fig. 28.b.) are clearly shorter flows, with the majority (64.2% of category volume) exporting from Brazilian ports. Yet the highest of these ratios, ranging from 14.06 to 4.96, respectively, export from Montevideo (Uruguay); Iquique (Chile); Cartagena (Colombia); Buenos Aires (Argentina); and

Buenaventura (Colombia). Still, as in all other data segments, the very high-ratio flows have a very low total volume of 9.44 TEUs. The 48 higher-ratio flows (including high- and very high-ratio flows) are those for which the null hypothesis must be rejected, and the alternative hypothesis accepted that expected distances are disproportionately high as compared to actual physical distance.

In comparison, the low-ratio flows (25 flows comprising 1.8% of the data segment volume, at 137.44 TEUs) are generally longer flows; these flows have a mean flow distance over double that of the high-ratio flows. Brazil exports 60.9% of the volume of this category, followed by Chile which exports 23.3% of the low-ratio category export volume for manufactured products.

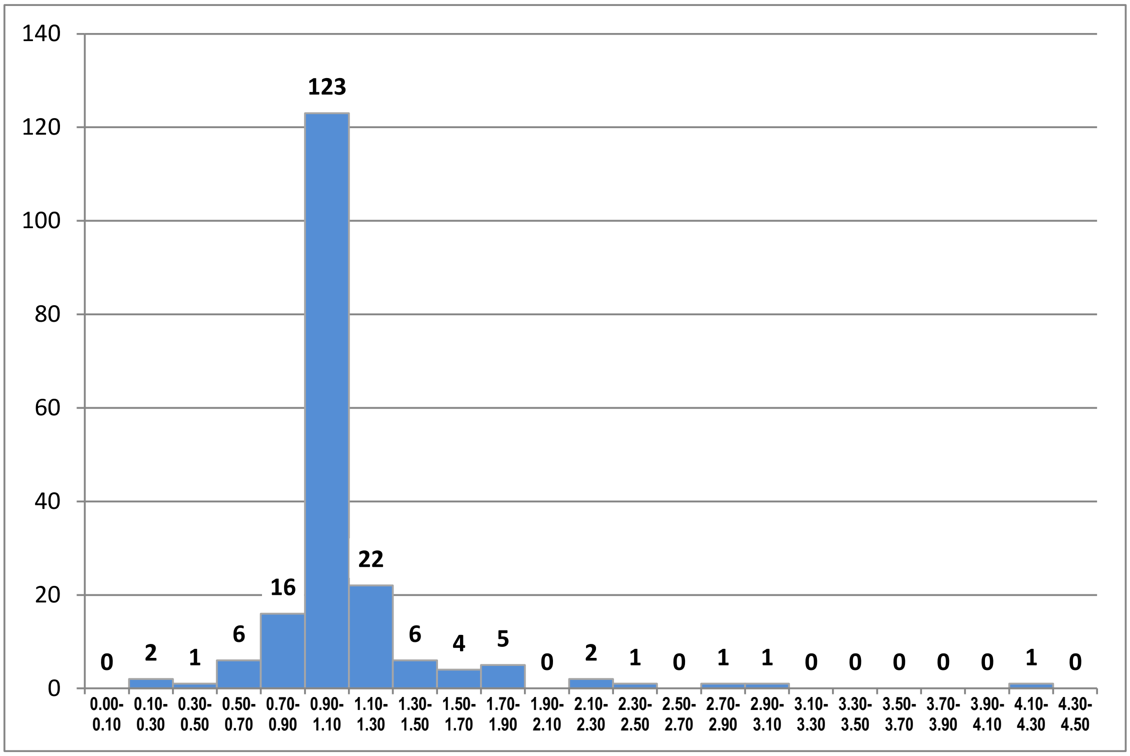


FIGURE 26: Histogram: Commodity Category 4: Manufactured Products – 196 aggregated observed flows by count *
*5 observations (aggregated observed flow) out of 196 have values ranging between 4.96 - 14.06; these observations are not displayed on the histogram.

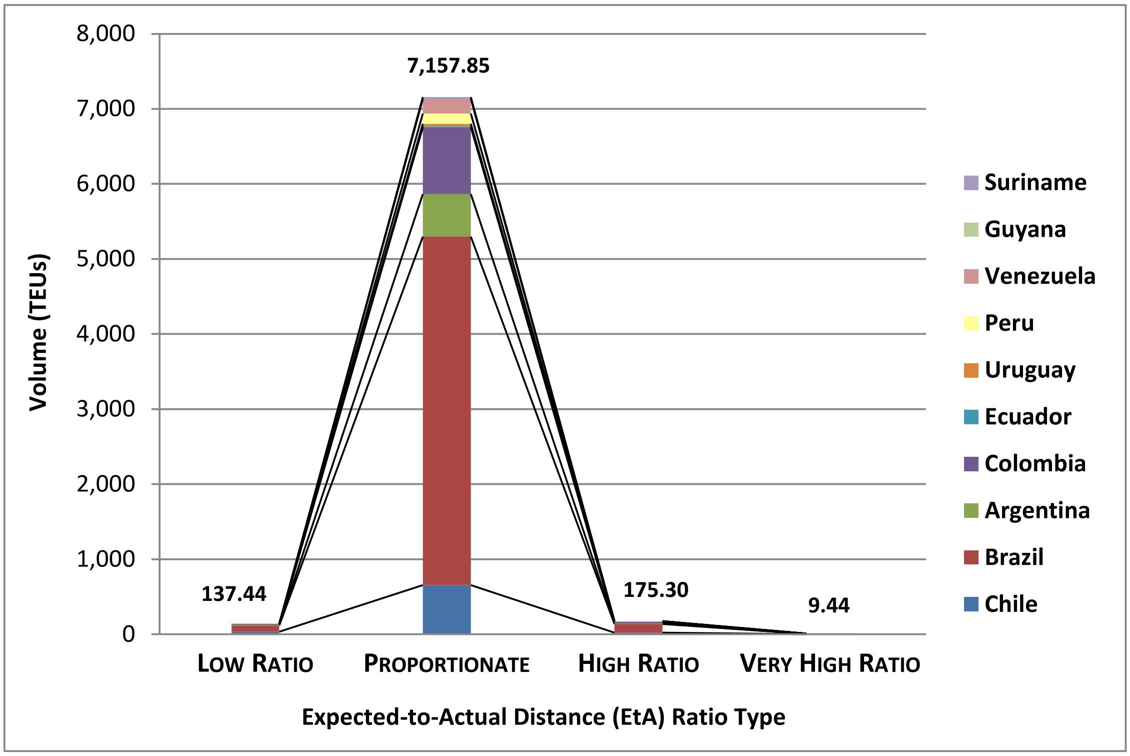


FIGURE 27: Commodity Category 4: Manufactured Products - volume and port of export country by EtA ratio type

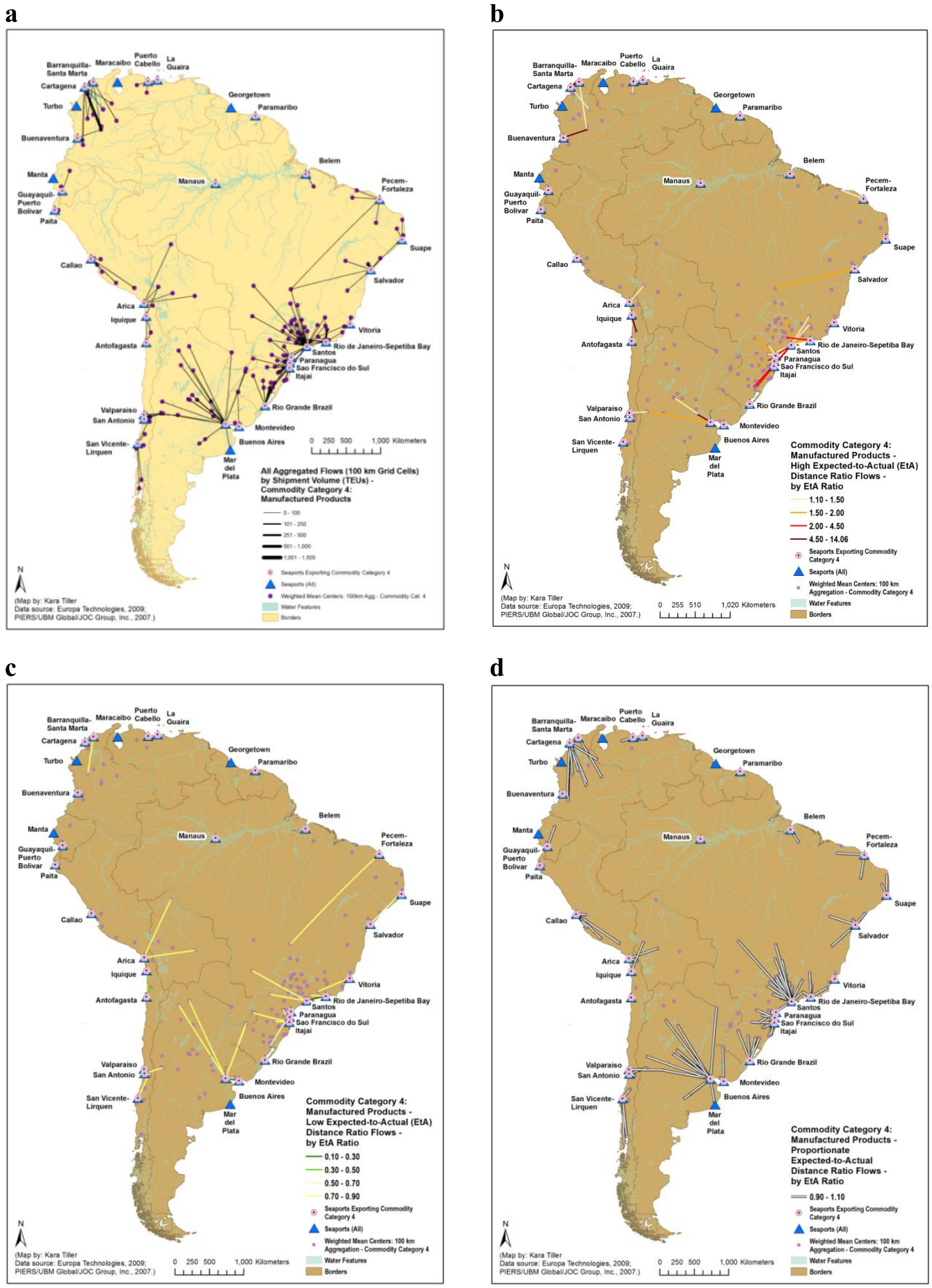


FIGURE 28: Reverse spatial interaction regression results - Commodity Category 4: Manufactured Products: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.6 Results: Reverse Spatial Interaction Regression – Commodity Category 5: High-Value Manufactured Products

For Commodity Category 5 (High-Value Manufactured Products), The R^2 for the fully estimated model is 0.966. High-Value Manufactured Products is one of two commodity categories for which it is anticipated that expected distance values are disproportionately high as compared to actual distance. However, for the majority of the high-value manufactured products segment (139 proportionate-ratio and 50 low-ratio flows, totaling 7,825.45 TEUs and 876.18 TEUs, respectively), the null hypothesis must be accepted that expected distances are not disproportionately high as compared to actual physical distance. For the 83 higher-ratio flows, totaling 1,099.02 TEUs, the null hypothesis must be rejected, and the alternative hypothesis accepted that for high-value manufactured products, expected distances are disproportionately high as compared to actual physical distance (Fig. 29).

Among commodity categories, high-value manufactured products evidence the lowest portion of total flow volume (79.8%) in the proportionate ratio category (Fig. 30). In another diversion from other segments, for port-of-export countries, the range shifts to include both 0.0% (for Guyana and Suriname, or Uruguay at 0.3%) to 98.7% (for Peru) of country flow volume in the proportionate category. Even for countries with a majority of proportionate flow volume, there are some (relatively) lower percentages in the proportionate category – for example Brazil (75.5%) and Chile (69.3%). Although these are still overwhelming proportionate majorities, the percentages are lower as compared to the same countries' proportionate shares of flow volume in other data segments.

Additionally, proportions of higher-ratio flows as percentages of country export flow volume are quite different than for other data segments. For high-value manufactured products, Guyana and Suriname have 100.0% of their export flow volume identified as high-ratio; Uruguay has 82.4% of its export flow volume identified as high-ratio, and Chile has 22.0% identified as high-ratio.

Brazil's share of non-proportionate export flows in the segment are nearly split between low-ratio flows (11.4% of Brazil's data segment flow volume) and higher-ratio flows (11.5% of Brazil's data segment flow volume). This is interesting to view at the port level. With some key exceptions, the maps for the Commodity Category 5/High-Value Manufactured Products (Figs. 31.a.-d.), visually appear to be quite similar to their counterparts for Commodity Category 4: Manufactured Products (Figs. 31.a.-d.) - the majority of flows is proportionate, and high-ratio flows are shorter and appear to be concentrated in Brazil's southeastern port region. For high-value manufactured products, like perishable products, higher-ratio flows have shortest mean flow distance (unlike other data segments for which proportionate-ratio flows have the shortest mean flow distance).

Overlapping hinterland/near-port bypass/diversion among both high- and low-ratio flows is apparent in Brazil's southeastern port region. As viewed to some degree in other segments, it appears that these flows demonstrate an avoidance of Rio Grande, Brazil, which is the near-port which several flows divert from to reach farther ports to the north. Additionally, ports which have both high- and low-ratio flows vary in shares by category.

Also visually apparent as a difference between Commodity Category 4/ Manufactured Products and Commodity Category 5/High-Value Manufactured Products in the maps are several longer, low-ratio flows to the ports of Buenos Aires and Montevideo (Fig. 31.c.). However, the share of total low-ratio flow volume exporting from Buenos Aires is low – just 24.96 TEUs (2.8%). (Likewise, though they are visually apparent, the low-ratio flows exporting from Montevideo have a low total volume of just 6.96 TEUs.)

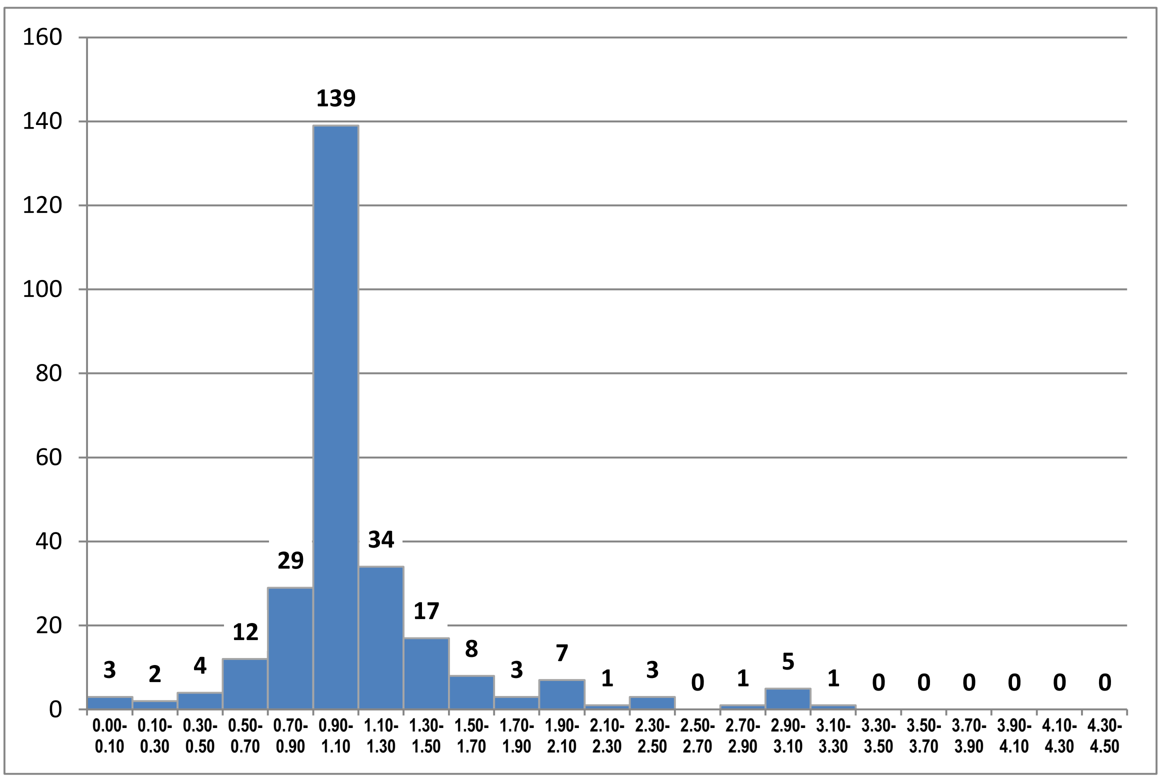


FIGURE 29: Histogram: Commodity Category 5 - High-Value Manufactured Products – 272 aggregated observed flows by count *
*3 observations (aggregated observed flow) out of 272 have values ranging between 5.07 - 44.68; these observations are not displayed on the histogram.

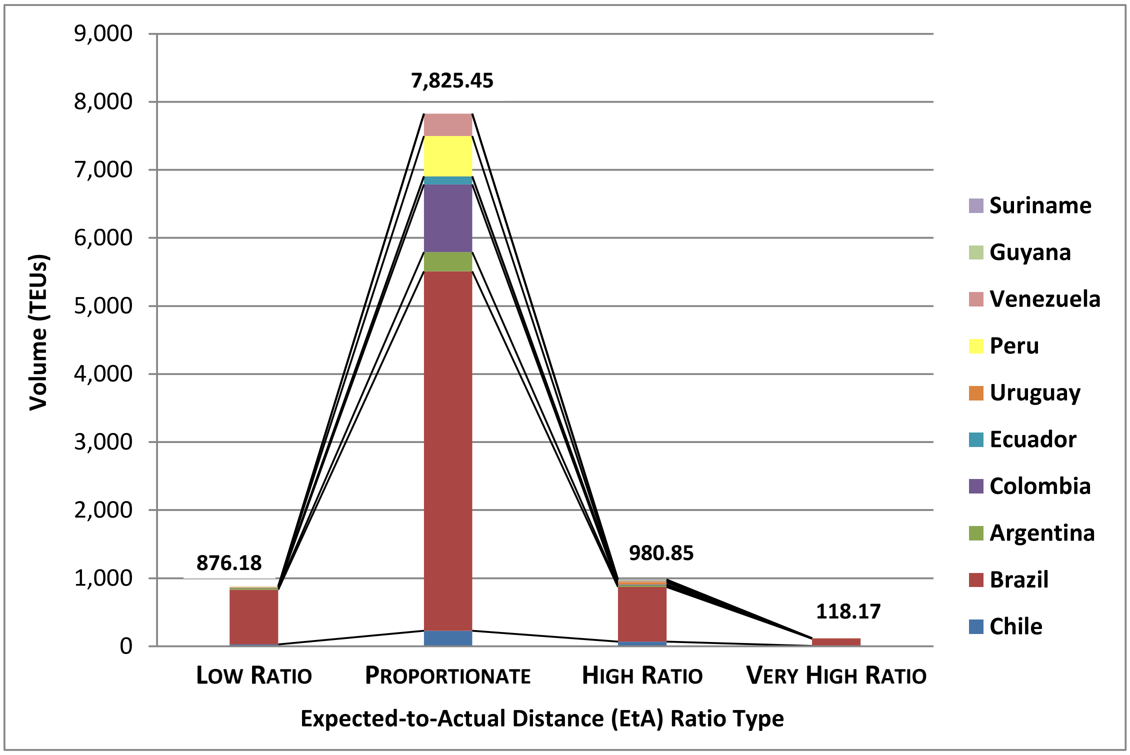


FIGURE 30: Commodity Category 5 - High-Value Manufactured Products - volume and port of export country by EtA ratio type

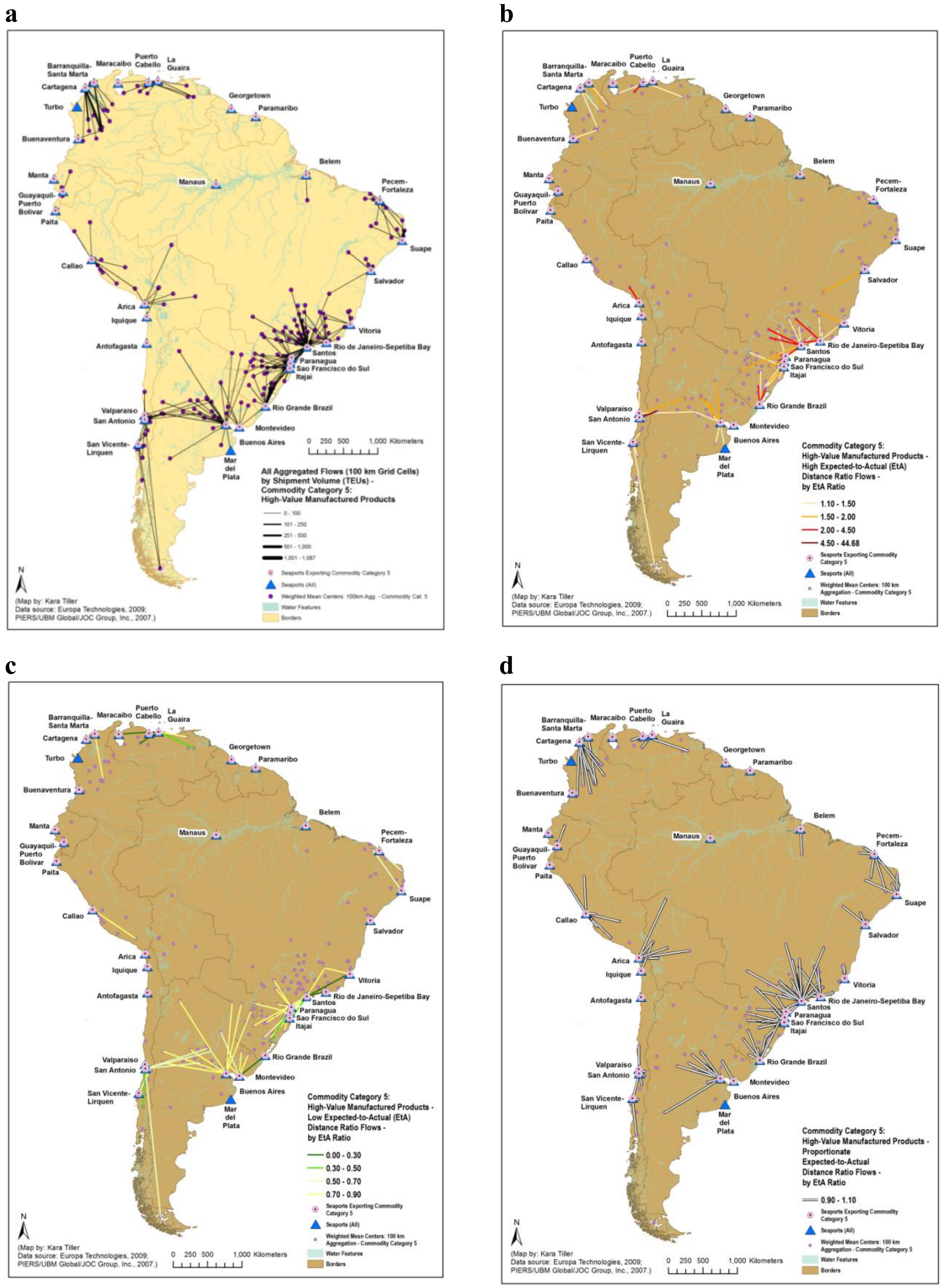


FIGURE 31: Reverse spatial interaction regression results – Commodity Category 5: High-Value Manufactured Products: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.7. Results: Reverse Spatial Interaction Regression – Commodity Category 6: Miscellaneous/Unable to be Determined Commodities

For Commodity Category 6 (Miscellaneous/Unable to be Determined Commodities), The R^2 for the fully estimated model is 0.962. It is anticipated that expected distance values are not disproportionately high as compared to actual physical distance for miscellaneous/unable to be determined commodities. The histogram for the Commodity Category 6/Miscellaneous/Unable to be Determined Commodities data segment (Fig. 32), displays the distribution of flows by EtA distance ratio. For 131 out of 190 aggregated observed flows by count, the null hypothesis must be accepted that for the miscellaneous/unable to be determined commodities data segment, expected distances are not disproportionately high, as compared to actual physical distance.

The proportionate ratio category is comprised of 101 flows comprising 82.2% of the data segment volume, at 1,891.57 TEUs (Fig. 33). Port-of-export countries in the data segment have percentages ranging from 0.0% (for Guyana and Suriname) to 96.9% (for Peru) of country flow volume total in the proportionate ratio category. The low-ratio category is comprised of 30 flows comprising 9.0% of the data segment volume, at 206.07 TEUs (Fig. 33). In the flow volume and ratio category maps (Figs. 34.a.-d.), it is apparent that although the Miscellaneous/Unable to be Determined Commodities data segment is a “catch-all” category, the differences in spatial distribution and arrangement of flows by ratio category are readily apparent. As expected, proportionate ratio flows have the highest mean flow volume for the data segment at 18.73 TEUs. The higher-ratio category has the shortest mean flow distance (276.3 km) and lowest mean flow volume, while the low-ratio category has the longest mean flow distance at 637.3 km.

For the 59 higher-ratio flows (comprising 8.8% of the data segment at 202.97 TEUs) (Figs. 34.a.-d.), the null hypothesis must be rejected that expected distances are not disproportionately high as compared to actual physical distance; the alternative hypothesis must be accepted that for the Commodity Category 6/Miscellaneous/Unable to be Determined Commodities data segment, expected distances are disproportionately high as compared to actual physical distance. The highest of these ratios, ranging from 12.06 to 4.55, respectively, are associated with flows exporting from Manaus (Brazil); Mar del Plata (Argentina); Cartagena (Colombia); Buenos Aires (Argentina); Santos and Itajai (Brazil).

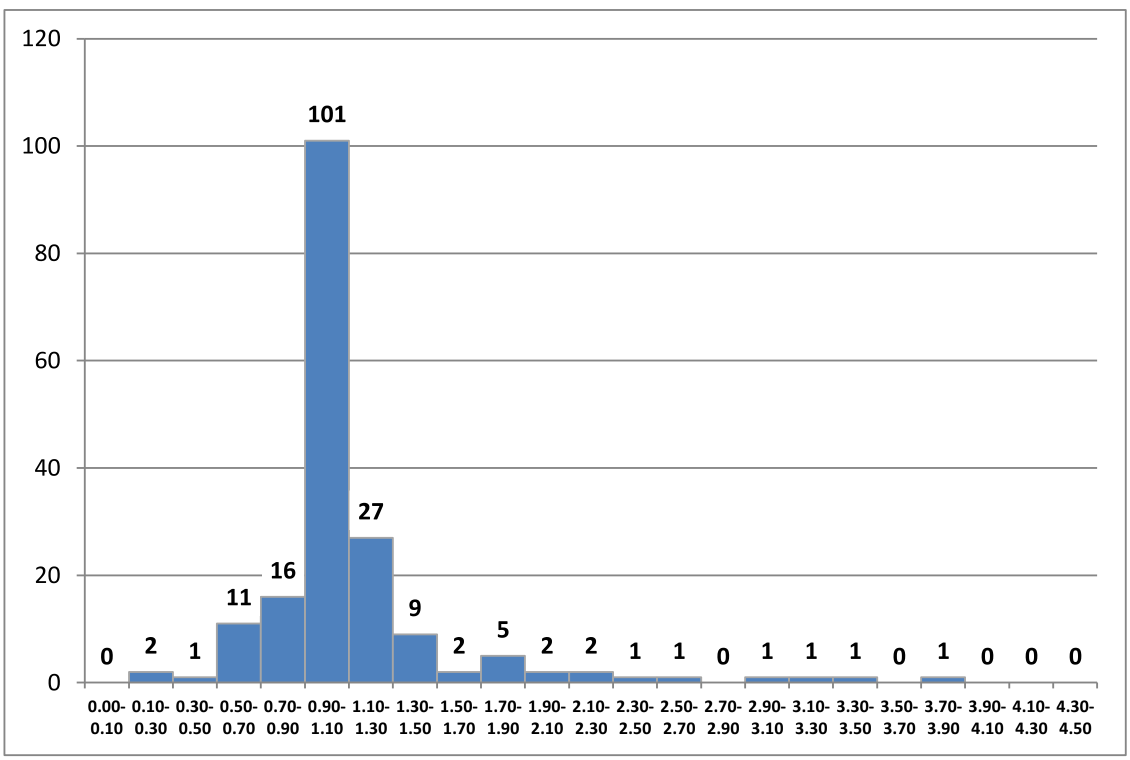


FIGURE 32: Histogram: Commodity Category 6: Miscellaneous/Unable to be Determined Commodities – 190 aggregated observed flows by count *
*6 observations (aggregated observed flow) out of 190 have values ranging between 4.55 – 12.06; these observations are not displayed on the histogram.

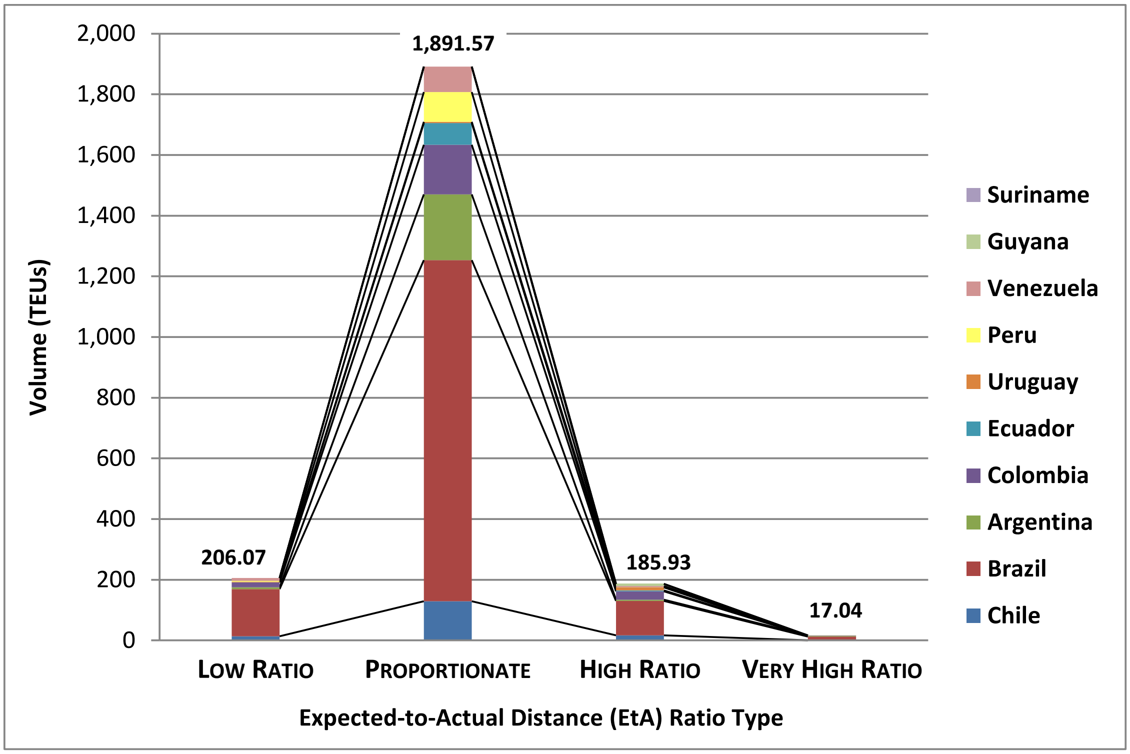


FIGURE 33: Commodity Category 6: Miscellaneous/Unable to be Determined Commodities – volume and port of export country by EtA Ratio Type

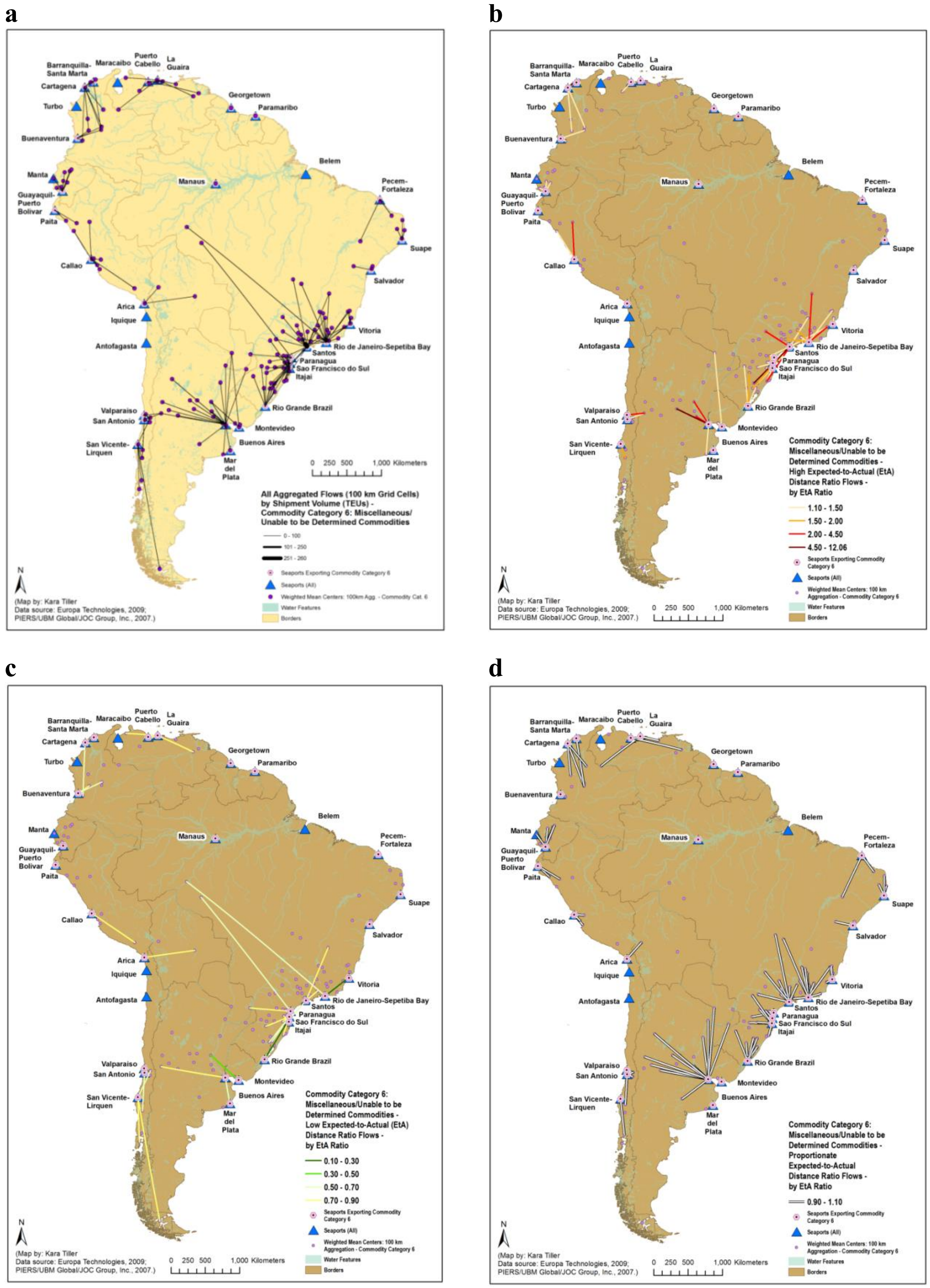


FIGURE 34: Reverse spatial interaction regression results – Commodity Category 6: Miscellaneous/Unable to be Determined Commodities: (a.) by volume (b.) higher-ratio flows (c.) low-ratio flows (d.) proportionate flows

6.8. Conclusions: Reverse Spatial Interaction Regression

For all data segments, the vast majority of flow volume does not demonstrate disproportionately high trade impedance. The higher-ratio category (inclusive of high- and very high-ratio flows) is shown to be the second-highest volume ratio category for most data segments, however it is the lowest-volume ratio category for Commodity Categories 3 (heavy commodities/minimally processed natural resources) and 6 (miscellaneous/unable to be determined). For all data segments, however, the total volume of the high-ratio category is far lower than the total volume of the proportionate-ratio category. For non-perishable time-sensitive goods, only 1.3% of the shipment volume is higher-ratio, indicating disproportionate trade impedance.

Proportionate-ratio flows have the highest mean volume among ratio categories and include the vast majority of the total O-D flow volume for every data segment. Additionally, proportionate-ratio flows more often have the shortest mean flow distance (with the exception of the cases of perishable products, high-value manufactured goods, and miscellaneous/unable to be determined commodities, for which higher-ratio flows have shortest mean distance).

Although higher-ratio flows are not the majority by flow count or volume, the total volume for the higher-ratio category is substantial for some segments, and usually compares in magnitude to the total volume of low-ratio flows. The main indicator that there may be systematic differences underlying low-, proportionate-, and higher-ratio category flows is the clear difference which is observed, for some data segments, between the spatial distribution and arrangement of the flows by ratio category.

These patterns provide general ideas about regions of origin, ports, flow directions/areas, and interactions with commodity type, which may have a relationship with disproportionate trade impedance. For example, greater volumes of higher-ratio flow occur with certain port of export/commodity type combinations. These observations may be only "flags" for other factors – urban congestion, particular problematic transportation corridors, or other factors – that underlie the observable patterns.

This initial exploration can only begin to identify these, and is a way of generating more specific questions in moving from the continental scale, to a finer-scale resolution. These observed differences in flow patterns, made apparent through analysis and visualization, can point to (or narrow down) problem areas for trade impedance as well as areas of trade facilitation.

Some particular points which emerged in this step include the following. For the ports of southeastern Brazil, in terms of the total dataset, perishable products, and manufactured products subset, high-ratio flows demonstrate a pattern that differs from that of proportionate and low-ratio flows in the region. It appears that many of these flows are bypassing nearer ports in order to reach Santos. For the proportionate flows exporting from Santos, interior-to-port (south-/southeast-directed) flows are most common. Higher-ratio flows exporting from Santos are more lateral, hailing from the west/southwest as far as Asuncion, Paraguay and western Misiones Province, Argentina – and as far east as the state of Espirito Santo, Brazil. High EtA ratios and hinterland overlap of flows appear to occur together; hinterland overlap/near-port bypass is far less evident in the proportionate-ratio flows. It is also seen demonstrated to some degree in high-value manufacturing flows, an avoidance of Rio Grande, Brazil, which is the near-port.

Additionally, the lowest values among the low-ratio flows could potentially be associated with coastwise moves – if so, this may indicate that a substantially lower portion of landside travel is actually low-ratio.

RESULTS: SPATIAL AUTOCORRELATION ANALYSIS

7.1. Introduction

In this step, the spatial pattern that emerges from the distribution of expected-to-actual distance (EtA) ratios for O-D flows is examined first using the Global Moran's I statistic (Cliff & Ord, 1972, 1981; Anselin, 1988; Getis, 2010) as an indicator of the existence and degree of global spatial autocorrelation in each data segment (for the total dataset and each of six commodity category segments). The analysis assigns the *mean* and *median* of EtA ratios for flows to their associated origin points, and also to their associated destination ports.

Therefore, in the Spatial Pattern step of this analysis, both the mean and median of EtA ratios for flows are investigated at origins and at destinations with which the flows are associated. Two notions are important to maintain throughout the final assessment of this analysis.

First, the mean and median are both used in order to present the complete picture of the central tendency of EtA ratios at each origin or destination point. The mean may be skewed by extreme EtA ratios – particularly as the range for high EtA ratios is wider than the range for low ratios, and as the extreme values of high EtA ratios are very high in terms of absolute value, for some cases. The mean may better indicate where these extreme values occur, yet the median may be regarded as the better indicator of the typical value in such cases. Thus it is reasonable to perform analysis on the basis of both measures. It is also reasonable to focus particularly upon results which are similar for both measures. Differences in the results between measures can be specifically interpreted in terms of the measure represented (in other words, a result for the mean analysis which indicates high-value clusters/outliers where the median does not, could be skewed by [an] extreme EtA ratio value[s]). Though both measures are reviewed in results, for the local Moran's I (cluster/outlier) analysis, final conclusions present the median EtA ratio results only, in order to provide the most conservative presentation of results.

A more ideal method of exploring the spatial pattern of O-D flows is paired-point cluster correspondence (as in Lu and Thill, 2003). (Conceptually, this multiple-point set/multi-location event methodology also accords well with an analysis of spatial autocorrelation for known origin and destination points, for which the actual route between points is unknown – and for which the assumption of such would be further abstraction.) However, this is a significantly resource- and skill-intensive research endeavor of its own merit, which stands as a potential future extension of this analysis. Baseline results obtained as described above could later be compared with those garnered from a more intensive focus on vector autocorrelation analysis.

Thus for the present spatial pattern analysis, global spatial autocorrelation will be investigated separately for both the mean and median of the expected-to-actual (EtA) distance ratios for flows associated with each of the origins (weighted mean centers of 100-km grid cells) and destinations (ports of export). This will be accomplished through a tool available in ArcGIS 10.2, the Spatial Autocorrelation/Global Moran's I tool.

Following this, local spatial autocorrelation analysis (Ord & Getis, 1995; Anselin, 1995; Sokal et al., 1995; Getis, 2010) is used to indicate significant clusters of low and high values and outliers, again using mean and median of the EtA ratios for flows at each associated origin and destination. The local analysis is performed using the Cluster and Outlier Analysis/Anselin Local Moran's I tool in ArcGIS 10.2.

Both global and local analyses are performed using an inverse distance conceptualization of spatial relationships, Euclidean distance, and a threshold distance of zero, with row standardization. The spatial weight matrix W is thus based upon the distance decline function, and therefore the weight entered into each cell representing the O-D interaction (I,j) is the inverse of distance between the origin and destination indicated (Getis, 2010).

The distribution of the mean and median EtA ratios for flows among their associated shipment origin points and ports of export is based upon the assumption of randomization. Therefore, this portion of the analysis addresses the following research question: For the set of observed O-D flows, what is the spatial pattern in the distribution of expected-to-actual (EtA) distance ratios?

With that said, this analysis assigns measures of central tendency of a trade impedance (expected-to-actual distance) ratio for flows, *to their endpoints*. Thus, results should not necessarily be interpreted as relating to occurrences at, or characteristics of, the endpoints themselves (origin points or ports of export). Rather, these results delineate, to some degree, a smaller group of features with which such results may be associated. For example, using this methodology, an area of high-high clusters identified at origin and/or destination points indicates a group of features which can be examined further, as when either of these endpoints is indicated, the associated flows are of course indicated as well. The bounds of this group of features may delineate a region which emerges as a target for further examination in terms of trade impedance.

As in many analyses involving interaction data, it is difficult to view a large number of flows or interactions – here, at the continental scale – and to determine which features, results, and/or characteristics should command attention. Among all origins, destinations, and flows and their attributes (such as orientation/direction or relationship with other features, including proximity or bypass), what are the essential aspects in terms of these results? How can results be coalesced into meaningful information about the research questions?

Though flows were aggregated, segmented by commodity category and parsed by EtA ratio category for visualization in the Step 1/Existence Step (Reverse Spatial Interaction Regression) – nevertheless there remains in each map a large number of flows, their associated features, and the gamut of results, the extent of which is difficult to process meaningfully. This dataset is not within the realm of Big Data (though it perhaps proposes a scalable methodology), however, with interactions which have multi-scalar interpretations, and which must be understood within multiple levels of trade and transportation context (including sectoral and political/economic factors, and directionality [Fig. 5]) – interpretation is best focused a richer selection of results from which meaning may be gleaned. Relatedly, it is a challenge and an ultimate goal of this analysis to determine how to drill down from a continental extent of flows to the sub-continental scale.

While it is true that local cluster and outlier analysis may be an imperfect means of investigating local spatial autocorrelation of flows (more ideally investigated through vector autocorrelation), nevertheless, it appears that this particular application of local spatial autocorrelation methods (assigning mean and median EtA ratios to flow endpoints) may provide a means of identifying a richer selection of results. This simple approach provides cluster and outlier results which also function to derive relevant endpoints and associated flows – indicating a potential region of interest – in what will be termed here the *cluster/outlier-context feature selection*.

This method provides assistance in drilling down to a more meaningful subset of results and spatial extent, without losing a view to the actual flows with which the endpoint-assigned values (EtA ratios) actually represent:

(1) The essential features – origins and destinations – with results of interest (flow values' measures of central tendency at endpoints) are pre-selected by their identification as high-high clusters or high-value outliers in the analysis;

(2) The flow context is maintained (via selection of flows associated with endpoints identified as clusters or outliers above); and

(3) Simultaneously, the region of focus is narrowed.

As compared to viewing high numbers of flow stratified by value ranges and/or segmentation by a few easily conceptualized categories (such as aggregated commodity categories), this provides a means of identifying relevant results, features, and a spatial sub-region, simultaneously – providing context based upon flow values.

The cluster/outlier-context feature selection can assist with exploring the nature of high-value clusters. In this context, the following results are obtained, and conclusions drawn, from the Spatial Pattern/spatial autocorrelation step of this analysis.

7.2. Results for the Total Dataset (All Shipments)

7.2.1. Global Spatial Autocorrelation Analysis

The global analysis of both mean and median EtA ratios at origins indicate clustering, with less than a 1.0% likelihood that the clustered pattern could be the result of chance. These statistics are as reported in Table 9. For mean EtA ratios at origins, Moran's I = 0.023227; z-score = 4.350828; and p-value = 0.000014. For median EtA ratios at origins, Moran's I = 0.022136; z-score = 4.214211; and p-value = 0.000025. Therefore in terms of the total dataset, for both the mean and median EtA ratio analyses at origins, the null hypothesis must be rejected that the spatial pattern of the distribution of EtA ratios is random; the alternative hypothesis must be accepted that the pattern is not random. The pattern is such that overall similar mean and median values are found in close proximity. Positive autocorrelation exists in terms of flow origins.

For destinations, results differ to some degree for mean and median EtA ratios. Results for the global analysis of mean EtA ratios at destinations indicate a random pattern (Moran's I = -0.021648, z-score = 0.180594, and p-value = 0.856687); thus for mean EtA ratios at destinations, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

Results for the analysis of median EtA ratios at destinations indicate a dispersed pattern, however this is only significant at the 10.0% level (Moran's I = -0.130810, z-score = -1.721205, and p-value = 0.085214), meaning that there is less than a 10.0% likelihood that the dispersed pattern could be the result of random chance. As more stringent (5.0%) statistical significance levels are needed here, this result – while informative as to the tendency of the distribution toward dispersion – cannot be taken to indicate a significant difference from randomness at the 5.0% level. Therefore, at destinations, for median EtA ratios, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

7.2.2. Local Spatial Autocorrelation Analysis

For the spatial Cluster and Outlier/Anselin Local Moran's I analysis, a result of "not significant" (Figs. 35.a. and b.) indicates that the mean or median EtA ratio at a certain origin or destination point is not more pronounced than that expected for a random distribution. Significant results indicate similarity (spatial clustering of high or low values) or dissimilarity (spatial outlier status: low values surrounded by high values, or vice versa) that is more pronounced than expected for a random distribution.

In terms of the local cluster and outlier analysis results for the total dataset, in the analysis of mean EtA ratios at origins (Fig. 35.a.), 411 out of 424 origins are found not to be significant. For the 13 origins whose mean EtA ratios are found to be significant, 11 of the origins are found to be significant in terms of similarity, as high-high clusters. The two regions delineated by these clusters are similar to those delineated by the analysis of median EtA ratios at origins, in which 8 high-high clusters are evidenced (Fig. 35.b.).

One of these regions delineated by high-high clusters spans from 150 km north of Antofagasta in northern Chile into southern regions of Peru, a span within which lie the ports of Iquique and Arica, Chile. The results evidenced in this region are the same for both the mean and median EtA ratio analyses at origins – an area of 4 high-high clusters punctuated by the same 2 low-high outliers. The second region which is similarly delineated by both mean and median analyses at origins extends from Asuncion, Paraguay to parts of western Misiones province, Argentina, to an area near Goya, Argentina on the Rio Parana. In the mean EtA ratio analysis at origins, however, the delineation of this region extends to the state of Santa Catarina, Brazil, and contains 3 additional high-high clusters (for a total of 7 high-high clusters in this region).

Additionally, in the analysis of median EtA ratios at origins, a high-low outlier is identified in Brazil, approximately 530 km northwest of Vitoria. Therefore – counting as significant results this outlier, the 8 high-high clusters, and 2 low-high outliers already detailed - for the analysis of median EtA ratios at origins (Fig. 35.b.), 413 out of 424 origins are found not to be significant.

The majority of destinations (ports of export) are found not to be significant in the similarity or dissimilarity of their mean and median EtA ratios, as well; 31 and 32 ports are found not to be significant for the mean and median EtA ratio analyses (Figs. 35.a.-b.), respectively. In the analysis of median EtA ratios at destinations, Antofagasta (Chile) is identified as a low-high outlier – a port with a low EtA ratio, surrounded by ports with high EtA ratios. Both the analyses of mean and the median EtA ratios at destinations identify the port of Mar del Plata (Argentina) as being a low-high outlier as well. In the mean analysis of EtA ratios at destinations, another outlier is identified – the high-low outlier of Montevideo, Uruguay (a port with a high EtA ratio, surrounded by ports with low EtA ratios). Additionally, the mean analysis of EtA ratios at destinations identifies Arica, Chile as a high-high cluster.

In the Reverse Spatial Interaction Regression results, lateral flows to the port of Santos were noted as part of the characteristic high-ratio spatial pattern. However, in viewing results of the analysis of median EtA ratios at origins, specifically the Asuncion/Misiones/Rio Parana sub-continental region of high-high clusters and their associated features (cluster/outlier-context feature selection, Fig. 36.a.), the occurrence of disproportionate trade impedance is somewhat more illuminated.

Specifically, it becomes apparent that: (1) it is not direction/orientation of flows per se that appears to potentially be associated with disproportionately high trade impedance, but rather an interaction of particular origins and flows with particular ports. This is clear in the comparison of high-high cluster origin flows to Buenos Aires, which are low or proportionate, versus flows to Montevideo (and other ports) for which EtA ratios of flows in this feature selection are high or very high); and yet (2) the high EtA ratios are not clustered at these ports themselves. This appears to indicate that the flows which are destined for them (which the EtA ratios represent) should be the features of interest in terms of these results.

With a fairly proximate (at the sub-continental scale) collection of high-high clusters, it appears that cluster/outlier-context feature selection assists in delineating a region of focus at a scale at which meaning can emerge. As discussed below, Fig. 36.b. provides another example.

Clarity is lent by the cluster/outlier-context feature selection for the second high-high origin cluster discussed in the results, in northern Chile/southern Peru (Fig. 36.b.). Again, median EtA ratios at ports themselves are not clustered; rather, an interaction of particular origins and flows with particular ports appears to be associated with high trade impedance.

It appears as well that rather than one region, there are actually two sub-regions identified. The area north of Arica and inland (potentially extending to, though not including, the port of Callao) is delineated as a potential area of disproportionate trade impedance/high-high EtA ratio clustering. The area immediately south of Iquique appears as a potential area of disproportionate trade impedance/high-high EtA ratio clustering as well; to the south of this high-high origin cluster, a proportionate flow and a low-value outlier origin and destination (at the port of Antofagasta) delineate where this potential area of trade impedance is bounded.

Though it is clear to see in Fig. 15.b. (Reverse Spatial Interaction Regression results) where many of these high-ratio flows occur, it is difficult to process and make meaning of this number of flows at the continental scale, even when stratified by EtA ratio level. Investigating via cluster/outlier-context feature selection appears to assist in pointing out regions of focus.

This regional view of meaningful results would be interesting to compare with the results of a future extension proposed here as Step 4 (Association Step) for the larger research endeavor.

TABLE 9: Spatial autocorrelation analysis – Global Moran's I analysis results for mean and median EtA ratios at origins and destinations

DATA SEGMENT **(AGGREGATED OBSERVED FLOWS)**	**FLOWS** **COUNT / VOLUME (TEUs)**	**NUMBER OF ORIGINS**	**RESULTS: ORIGINS – MEAN OF EtA RATIOs** **(FOR FLOWS DEPARTING THIS ORIGIN)**	**RESULTS: ORIGINS – MEDIAN OF EtA RATIOs** **(FOR FLOWS DEPARTING THIS ORIGIN)**	**NUMBER OF DESTINATIONS**	**RESULTS: DESTINATIONS – MEAN OF EtA RATIOs (FOR FLOWS ARRIVING AT THIS DESTINATION)**	**RESULTS: DESTINATIONS – MEDIAN OF EtA RATIOs** **(FOR FLOWS ARRIVING AT THIS DESTINATION)**
All Shipment Flows (Total Dataset)	680 / 78,296.51	424	Clustered: z = 4.350828 p = 0.000014 I = 0.023227 E(I) = -0.002364 Var = 0.000035	Clustered: z = 4.214211 p = 0.000025 I = 0.022136 E(I) = -0.002364 Var = 0.000034	34	Random: z = 0.180594 p = 0.856687 I = -0.021648 E(I) = -0.030303 Var = 0.002297	Random (5% level): z = -1.721205 p = 0.085214 I = -0.130810 E(I) = -0.030303 Var = 0.003410
1: **Perishable Products**	391 / 26,905.07	283	Random: z = 0.250189 p = 0.802441 I = -0.001846 E(I) = -0.003546 Var = 0.000046	Random: z = 0.244456 p = 0.806877 I = -0.002462 E(I) = -0.003546 Var = 0.000020	32	Random: z = 1.453749 p = 0.146016 I = 0.004689 E(I) = -0.032258 Var = 0.000646	Random: z = 0.204060 p = 0.838306 I = -0.020581 E(I) = -0.032258 Var = 0.003274
2: **Non-Perishable Time-Sensitive Products**	45 / 544.09	41	Random: z = 0.294814 p = 0.768136 I = -0.016370 E(I) = -0.025000 Var = 0.000857	Random: z = -0.483173 p = 0.628973 I = -0.050694 E(I) = -0.025000 Var = 0.002828	16	Random: z = -0.400610 p = 0.688707 I = -0.108505 E(I) = -0.066667 Var = 0.010907	Random: z = 0.058618 p = 0.953256 I = -0.059812 E(I) = -0.066667 Var = 0.013674
3: **Heavy Commodities/ Minimally Processed Natural Resources**	346 / 31,266.06	223	Random: z = -0.340376 p = 0.733573 I = -0.007694 E(I) = -0.004505 Var = 0.000088	Random: z = 0.181841 p = 0.855708 I = -0.003080 E(I) = -0.004505 Var = 0.000061	31	Random: z = -0.892913 p = 0.371904 I = -0.078558 E(I) = -0.033333 Var = 0.002565	Random: z = 0.404749 p = 0.685662 I = -0.006490 E(I) = -0.033333 Var = 0.004398

Table 9 (continued)

DATA SEGMENT **(AGGREGATED OBSERVED FLOWS)**	**FLOWS** **COUNT / VOLUME (TEUs)**	**NUMBER OF ORIGINS**	**RESULTS: ORIGINS – MEAN OF EtA RATIOs** **(FOR FLOWS DEPARTING THIS ORIGIN)**	**RESULTS: ORIGINS – MEDIAN OF EtA RATIOs** **(FOR FLOWS DEPARTING THIS ORIGIN)**	**NUMBER OF DESTINATIONS**	**RESULTS: DESTINATIONS – MEAN OF EtA RATIOs** **(FOR FLOWS ARRIVING AT THIS DESTINATION)**	**RESULTS: DESTINATIONS – MEDIAN OF EtA RATIOs** **(FOR FLOWS ARRIVING AT THIS DESTINATION)**
4: **Manufactured Products**	196 / 7,480.03	152	Random: z = -0.231084 p = 0.817250 I = -0.009578 E(I) = -0.006623 Var = 0.000164	Random: z = -0.168287 p = 0.866358 I = -0.008757 E(I) = -0.006623 Var = 0.000161	29	Random: z = -0.853352 p = 0.393464 I = -0.092075 E(I) = -0.035714 Var = 0.004362	Random: z = 0.761744 p = 0.446213 I = -0.000863 E(I) = -0.035714 Var = 0.002093
5: **High-Value** **Manufactured Products**	272 / 9,800.65	196	Random: z = 0.882208 p = 0.377664 I = -0.003125 E(I) = -0.005128 Var = 0.000005	Random: z = 1.157640 p = 0.247011 I = -0.002526 E(I) = -0.005128 Var = 0.000005	32	Random: z = -0.623583 p = 0.532902 I = -0.063781 E(I) = -0.032258 Var = 0.002555	Random: z = 0.164683 p = 0.869194 I = -0.028293 E(I) = -0.032258 Var = 0.000580
6: **Miscellaneous/** **Unable to be Determined** **Commodities**	190 / 2,300.61	145	Random: z = 0.183667 p = 0.854274 I = -0.004268 E(I) = -0.006944 Var = 0.000212	Random: z = 0.236359 p = 0.813154 I = -0.003506 E(I) = -0.006944 Var = 0.000212	28	Random: z = -0.017720 p = 0.985862 I = -0.038122 E(I) = -0.037037 Var = 0.003746	Random: z = 0.092374 p = 0.926401 I = -0.031404 E(I) = -0.037037 Var = 0.003719

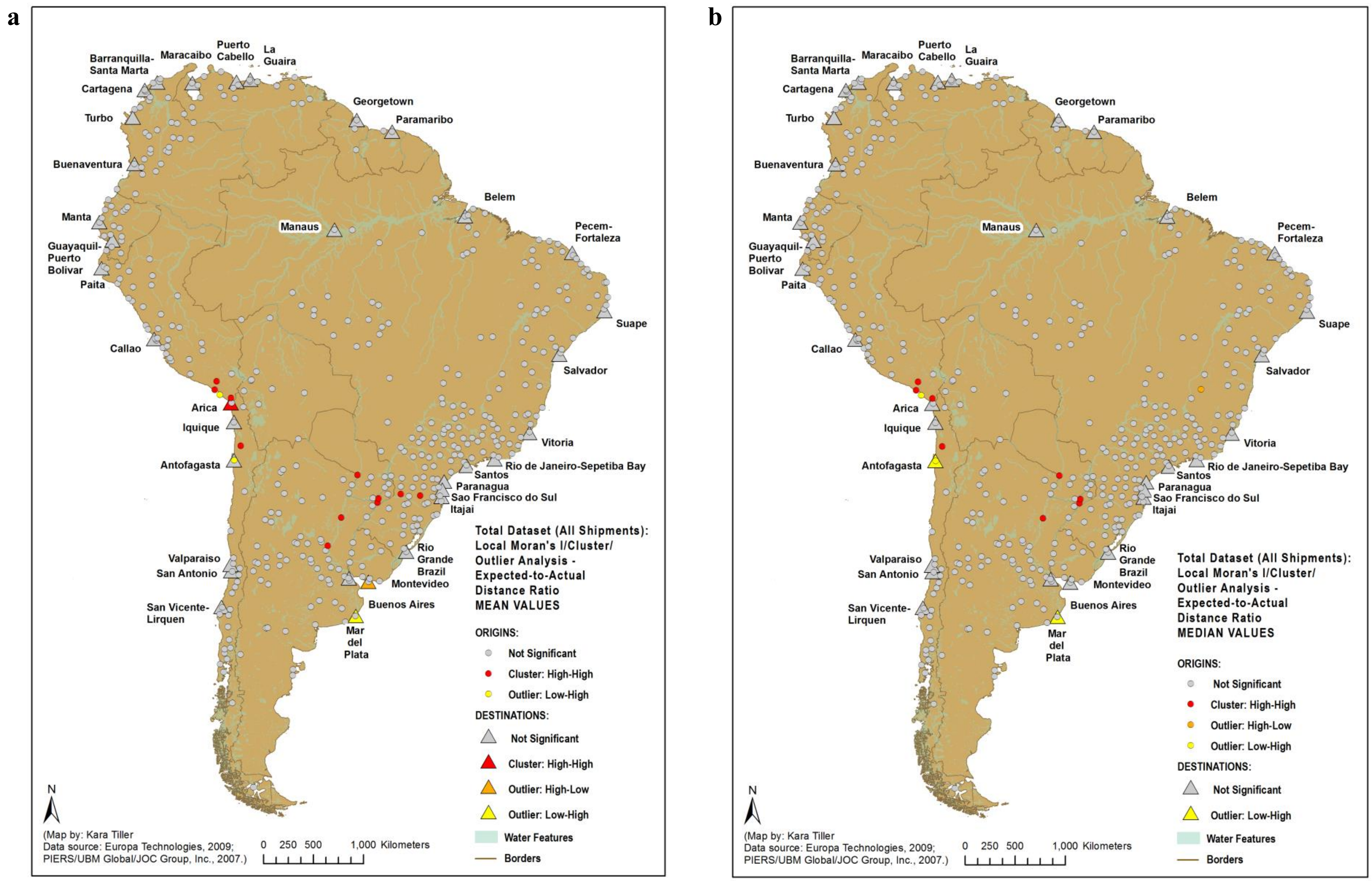
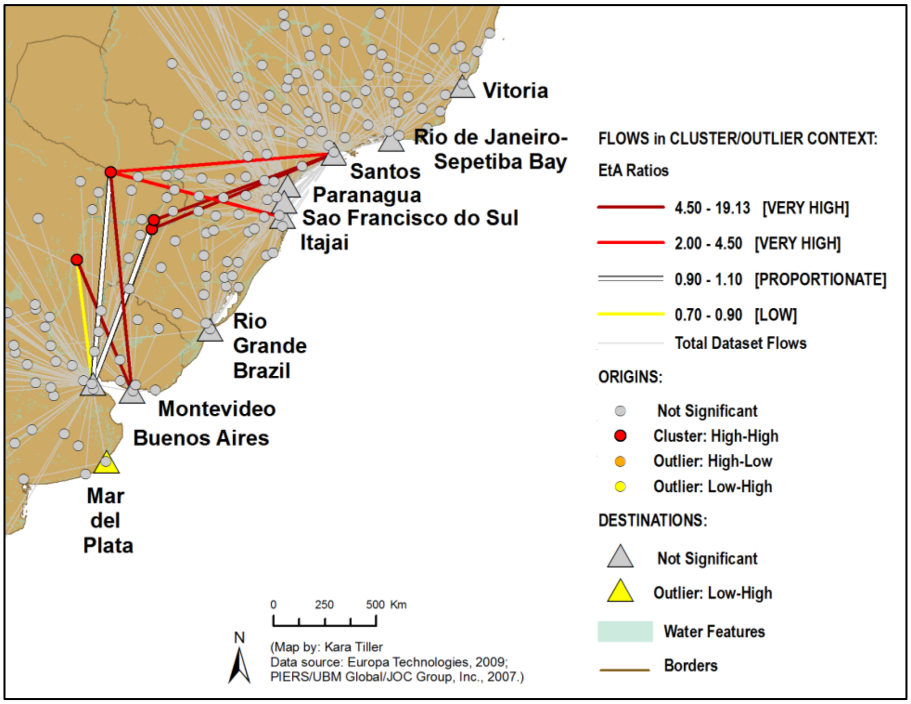


FIGURE 35: Total dataset (all shipments): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

a



b

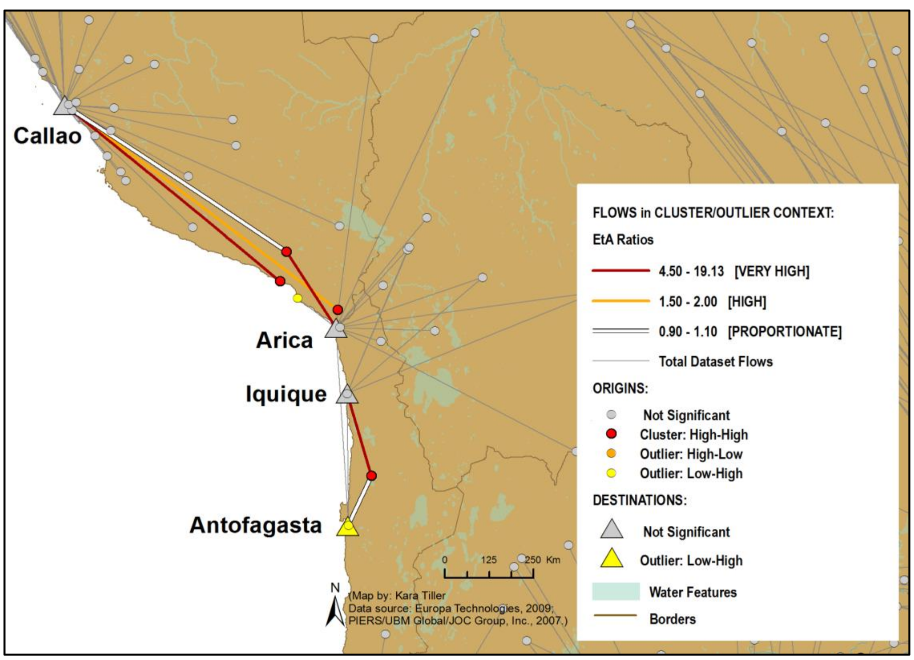


FIGURE 36: Total dataset (all shipments): Local Moran's I/cluster/outlier analysis – median values and cluster/outlier-context feature selection with EtA ratios: (a) high-high origin cluster 1; (b) high-high origin cluster 2.

7.3. Results for Commodity Category 1 (Perishable Products)

7.3.1. Global Spatial Autocorrelation Analysis

For the global spatial autocorrelation (Moran's I) analysis of the Commodity Category 1 (Perishable Products) segment, both mean and median EtA ratios at origins have results which indicate a random pattern (Table 9). Thus for both mean and median EtA ratios at origins, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

The global analysis of mean and median EtA ratios at destinations also have results which indicate a random pattern (Table 9). Therefore it is also true that for both mean and median EtA ratios at destinations, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

7.3.2. Local Spatial Autocorrelation Analysis

In terms of local spatial autocorrelation analysis for the Commodity Category 1 (Perishable Products) segment, for mean EtA ratios at origins (Fig. 37.a.), 280 origins out of 283 are found not to be significant in their similarity or dissimilarity. There are three statistically significant results, and all are found in Argentina. One high-high origin cluster is located approximately 205 km west of Buenos Aires. Two outliers are found in this region, as well – one low-high outlier (low value surrounded by high values) is in very close proximity to the port of Buenos Aires, and one high-low outlier (high value surrounded by low values), is found approximately 24 km to the south.

For the analysis of median EtA ratios at origins, 279 origins out of 283 are found not to be significant in their similarity or dissimilarity. Just as in the mean analysis, all four of the significant results are in the general vicinity of the Rio de la Plata area. For the origins, the results are the same for the median analysis as for the mean analysis, both for outliers and clusters, but with one addition. An additional high-high cluster is found approximately 335 km northwest of Montevideo on the Rio Uruguay (near Paysandu, Uruguay).

In the analysis of the mean EtA ratios at destinations, a high-high cluster is identified at the port of Buenos Aires; additionally, a low-high outlier is evidenced at the port of Mar del Plata, Argentina. The remaining 30 ports in the mean analysis are found to be not significant. However, in the analysis of the median EtA ratios at destinations, all 32 ports exporting Commodity Category 1 (Perishable Products) are found not to be significant in their similarity or dissimilarity.

In viewing the cluster/outlier-context feature selection for the local analysis of median EtA ratios (not pictured here, for brevity), it is demonstrated that one extremely high-value flow is associated with each of the two origin high-high clusters (205 km west of Buenos Aires and 335 km northwest of Montevideo on the Rio Uruguay) and one high-low outlier (24 km south of Buenos Aires). The remainder of flows associated with these clusters are proportionate. Two of these high-value flows are exported via Buenos Aires, while another is exported through Montevideo – yet neither of these ports was identified as significant in terms of similarity or dissimilarity. The origins and flows which are high-ratio within the cluster/outlier context appear to delineate a potential area of focus for trade impedance within the Rio de la Plata region, in terms of perishable products.

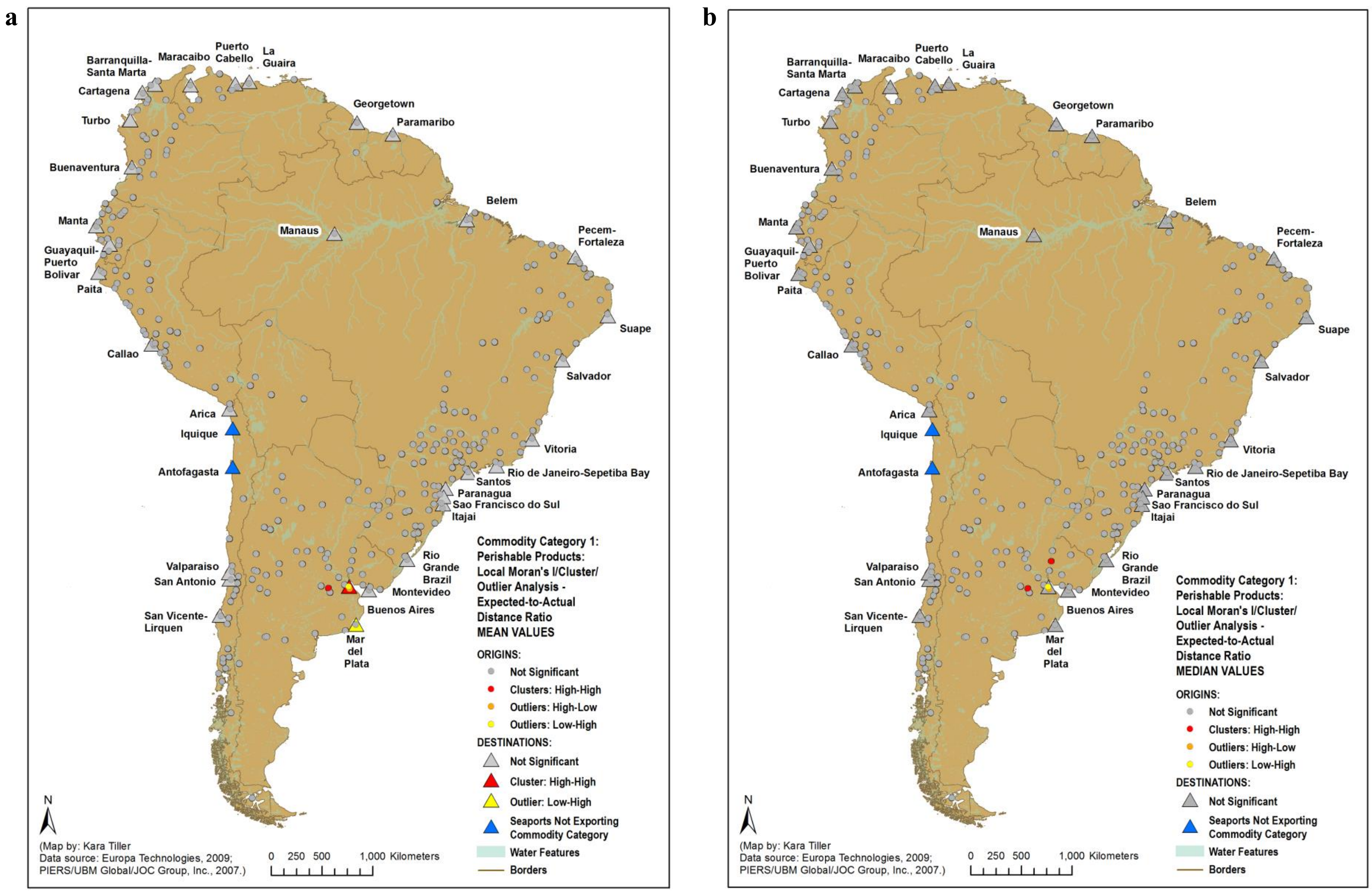


FIGURE 37: Commodity Category 1 (Perishable Products): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

7.4. Results for Commodity Category 2 (Non-Perishable Time-Sensitive Products)

7.4.1. Global Spatial Autocorrelation Analysis

For the global Moran's I analysis of the Commodity Category 2 (Non-Perishable Time-Sensitive Products) data segment - a total of 45 flows - both mean and median of EtA ratios at origins and destinations have results which indicate a random pattern (with values as presented in Table 9). Thus for both mean and median EtA ratios at both origins and destinations, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

However, this data segment involves a low number of destinations (16 ports), and this number of destination input features is low enough for the global Moran's I analysis of destinations to produce untrustworthy results. This is the case for 3 out of 7 data segments (Commodity Categories 2, 4, and 6) for the analysis of destinations/ports only; each of these segments has less than 30 destinations as input features. (This does not apply to the analysis of origins for any data segment, as the number of origin input features is greater than 30 for all segments.)

7.4.2. Local Spatial Autocorrelation Analysis

For the cluster and outlier/Anselin local Moran's I analysis of Non-Perishable Time-Sensitive Products, in the analysis of mean EtA ratios at origins (Fig. 38.a.), 40 origins out of 41 are found not to be significant in their similarity or dissimilarity. One high-low outlier (high value surrounded by low) is found 265 km north of the port of Rio Grande, Brazil.

A similar result is found for the analysis of mean EtA ratios at destinations (Fig. 38.a.), as 15 out of 16 ports are not found to be significant. However, one high-low outlier is identified at the port of Santos.

As in the global analysis, for the cluster and outlier/Anselin local Moran's I analysis, the low number of destinations for this segment is not ideal for the analysis and may produce untrustworthy results; therefore the results for destination clusters and outliers should be taken in that context.

In the analysis of median EtA ratios at origins (Fig. 38.b.), just as in the analysis of mean ratios at origins, 40 out of 41origins are found not to be significant in their similarity or dissimilarity. The one significant origin is a high-low outlier located in the area of Asuncion, Paraguay.

The results of the analysis of median EtA ratios at destinations feature one high-high cluster at the pseudo-port of San Vicente-Lirquen, Chile. Also, one significant low-high outlier is found at the port of San Antonio, Chile. The remaining 14 ports are not found to have significant similarities or differences.

The local Moran's I analysis of median EtA ratios at origins indicates one significant high-low outlier near Asuncion, Paraguay. When associated features are selected, this origin is shown to be associated with one high-ratio flow which exports from Buenos Aires – again, potentially indicating the Rio de la Plata region as a of focus in terms of trade impedance for non-perishable time-sensitive products.

Though the caution noted above must be taken in terms of results for destinations, results of note for the analysis of median EtA ratios at destinations include the following. The high-high cluster at the pseudo-port of San Vicente-Lirquen, Chile, is associated with a high-ratio flow that originates approximately 11 km from the port. The low-high outlier at the port of San Antonio, Chile, is associated with one proportionate flow originating approximately 90 km from the port.

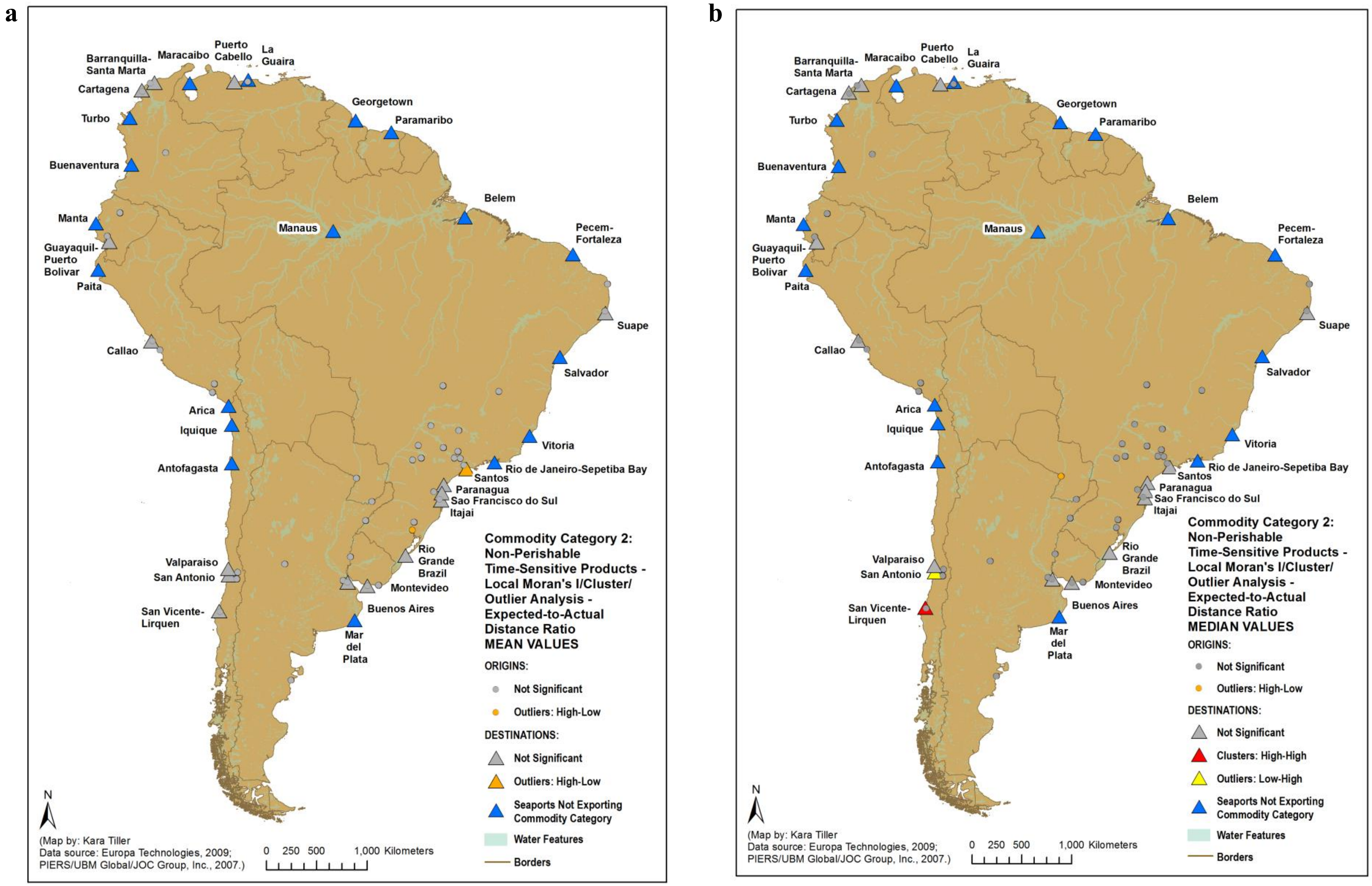


FIGURE 38: Commodity Category 2 (Non-Perishable Time-Sensitive Products): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

7.5. Results for Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources)

7.5.1. Global Spatial Autocorrelation Analysis

For the global Moran's I analysis of the Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources) Segment – a total of 346 flows – both mean and median EtA ratios at origins have results which indicate a random pattern (with Moran's I = -0.007694, z-score = -0.340376, and p-value = 0.733573 for mean EtA ratios, and Moran's I = -0.003080, z-score = 0.181841, and p-value = 0.855708 for median EtA ratios) (Table 9). Also, both mean and median EtA ratios at destinations have results which indicate a random pattern (with Moran's I = -0.078558, z-score = -0.892913, and p-value = 0.371904 for mean EtA ratios, and Moran's I = -0.006490, z-score = 0.404749, and p-value = 0.685662 for median EtA ratios). Thus for both mean and median EtA ratios at origins and destinations, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

7.5.2. Local Spatial Autocorrelation Analysis

For the cluster and outlier/Anselin local Moran's I analysis of Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources), in the analysis of mean EtA ratios at origins (Fig. 39.a.), 219 origins out of 223 are found not to be significant in their similarity or dissimilarity. One high-high cluster is found in the area of Cucuta, Colombia, proximate to the border with Venezuela. Additionally, 3 high-low outliers are evidenced: one close to this high-high cluster, but more interior to

Colombia; another at Asuncion, Paraguay; and another apparent on the map between the Chilean ports of San Vicente-Lirquen (pseudo-port) and San Antonio.

In the analysis of mean EtA ratios at destinations (Fig. 39.a.), 28 out of 31 ports are not found to be significant. Two high-low outliers are identified – at ports of Cartagena and San Vicente-Lirquen (pseudo-port) – and one low-high outlier is identified at the pseudo-port of Barranquilla-Santa Marta (Colombia).

In the analysis of median EtA ratios at origins (Fig. 39.b.), 218 origins out of 223 are found not to be significant in their similarity or dissimilarity. Two high-high clusters are evidenced – one near Cucuta, Colombia (as evidenced in the mean analysis), and one in close proximity to pseudo-port Barranquilla-Santa Marta (Colombia). Similar to the mean analysis, 2 high-low outliers are evidenced - one close to the high-high cluster at Cucuta, but more interior to Colombia, and another at Asuncion, Paraguay. Additionally, one low-high outlier is found in close proximity to the port of Maracaibo, Venezuela.

In the analysis of median EtA ratios at destinations (Fig. 39.b.), 28 out of 31 ports are found not to be significant in their similarity or dissimilarity. Two high-high clusters are evidenced – at the ports of Georgetown, Guyana, and Paramaribo, Suriname. One high-low outlier is evidenced at the port of Antofagasta, Chile.

The analysis for this data segment begins to clearly demonstrate how results may differ for different commodity categories. The cluster/outlier-context feature selection enabled by the results from the local Moran's I analysis of median EtA ratios guides a selection of flows lending greater clarity to the spatial pattern of the data segment.

Therefore it is possible to view, as in Fig. 40.a. that though some origins are identified as high-high clusters or high-value outliers, this may not necessarily implicate the origin itself in terms of trade impedance. When associated flows are selected and their EtA ratios are displayed, it is clear that the flows from these origins are high-ratio *only for flows which are destined for Cartagena* – at least for the heavy commodities/minimally processed natural resources data segment. Cartagena itself is not identified as a significant cluster or outlier in the median analysis results (viewed here) – but was identified by the mean analysis as a high-low outlier.

Overall, this analysis seems to suggest that it is the interactions between these flows and the port that should garner attention, and the area delineated by Cartagena and the high-high clusters may be identified as a region of focus for heavy commodities/minimally processed natural resources trade impedance.

The usefulness of this type of analysis to elucidate the nature of the spatial pattern can be viewed also in Fig. 40.b., as here Asuncion is identified as a high-low outlier. It cannot (necessarily) be assumed, however, that this implicates the Rio de la Plata region, in terms of trade impedance for this commodity category. When cluster/outlier-context feature selection is used to select and display the associated flows, it is clear that from this origin, there is a proportionate flow to Buenos Aires - but the high EtA-ratio flows are exporting from Santos and Itajai. This may suggest that the region of focus for trade impedance is the Paraguayan portion of the landside journey, as well as states of Parana, Santa Catarina, and Sao Paulo in Brazil (but not the Rio de la Plata region) for heavy commodities/minimally processed natural resources trade impedance.

The cluster/outlier-context feature selection for the high-high clusters and the high-low outlier evidenced in the analysis of median EtA ratios for destinations, is not shown here. This is because for the flows to these ports, origins are in fairly close proximity (some within kilometers of the ports, none more than 100 km away). Though these results are visually less interesting, they stand out due to the identification of the ports of Georgetown and Paramaribo as high-high clusters – which points to this region as a potential area of focus for trade impedance in terms of heavy commodities/minimally processed natural resources, as well.

a b

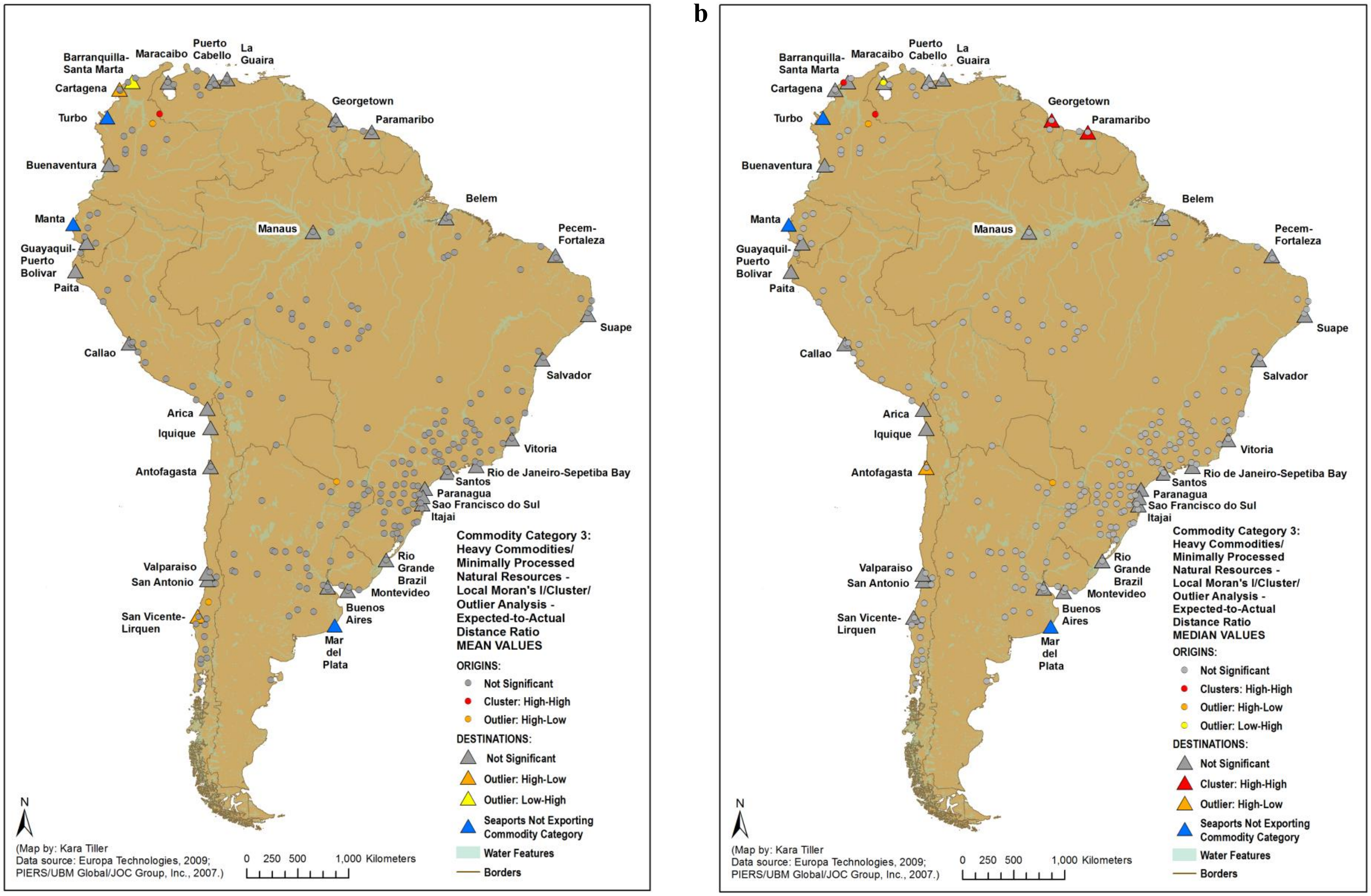


FIGURE 39: Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources): Local Moran's I/ cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

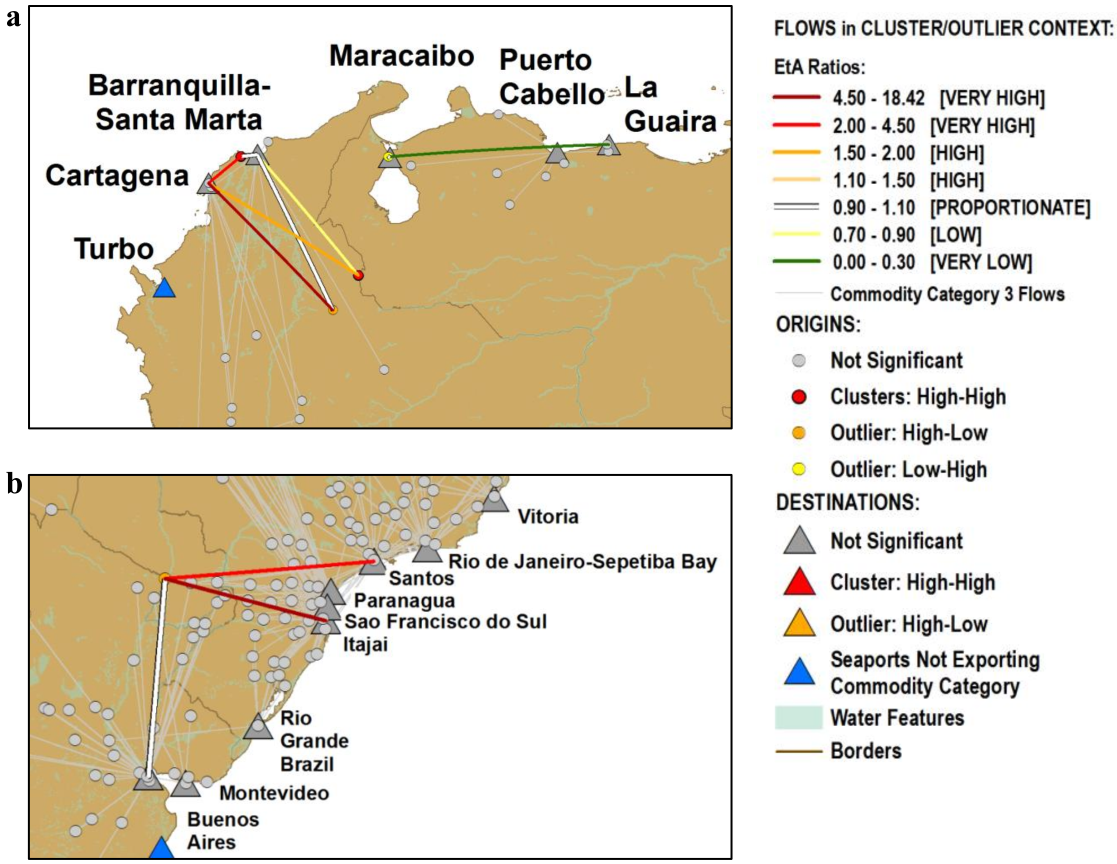


FIGURE 40: Commodity Category 3 (Heavy Commodities/Minimally Processed Natural Resources): Local Moran's I/cluster/outlier analysis – median values and cluster/outlier-context feature selection with EtA Ratios: (a) origin clusters/outliers: Colombia/Venezuela; (b) origin outlier: Asuncion, Paraguay

7.6. Results for Commodity Category 4 (Manufactured Products)

7.6.1. Global Spatial Autocorrelation Analysis

For the global Moran's I analysis of the Commodity Category 4 (Manufactured Products) data segment - a total of 196 flows - both mean and median EtA ratios at origins have results which indicate a random pattern (with Moran's I = -0.009578, z-score = -0.231084, and p-value = 0.817250 for mean EtA ratios, and Moran's I = -

0.008757, z-score = -0.168287, and p-value = 0.866358 for median EtA ratios) (Table 9). Also, both mean and median EtA ratios at destinations have results which indicate a random pattern (with Moran's I = -0.092075, z-score = -0.853352, and p-value = 0.393464 for mean EtA ratios, and Moran's I = -0.000863, z-score = 0.761744, and p-value = 0.446213 for median EtA ratios). However, it should be noted that the number of destinations (29 ports) is low enough in this data segment for this analysis to produce untrustworthy results.

7.6.2. Local Spatial Autocorrelation Analysis

In the cluster and outlier/Anselin local Moran's I analysis results for Manufactured Products, the analysis of both mean and median EtA ratios at origins (Figs. 41.a.-b.) have the same result: two high-low outliers, one located between the ports of Iquique and Antofagasta, Chile, and one located in close proximity to the port of Montevideo, Uruguay.

As in the global analysis, the low number of destinations for this segment (29 ports) is not ideal for the analysis, and may produce untrustworthy results; therefore the results for destination clusters and outliers should be taken in that context.

In the analysis of mean EtA ratios at destinations, two high-low outliers are found, whereas the remaining 27 ports are found not to be significant. These outliers are found at the ports of Iquique and Montevideo.

In the analysis of median EtA ratios at destinations, only one high-low cluster is identified, for the port of Buenaventura, Colombia; the remaining 28 ports are found not to be significant.

It is demonstrated that for the origin high-low outliers, one very high-ratio flow is identified between the origin proximate to Montevideo and the port itself. As all other flows to this port are proportionate, it appears that this sole high-ratio flow is the basis for the result that Montevideo stands as a high-low outlier for Manufactured Products.

For the origin high-low outlier located between the ports of Iquique and Antofagasta, one extreme high-ratio flow is destined for Iquique, with a low-ratio flow destined for Antofagasta. The analysis of mean EtA ratios at destinations had identified the port of Iquique as a high-low outlier, and though other flows were found to be proportionate, this was not its only high-ratio flow. This area near and including Iquique may potentially be a focus for trade impedance in terms of manufactured products.

For the port of Buenaventura, Colombia, identified as a destination high-low cluster, one extreme high ratio flow from the interior of Colombia (over 360 km away) appears to be responsible for its designation as such; another flow, originating more proximate to the port, has a proportionate EtA ratio.

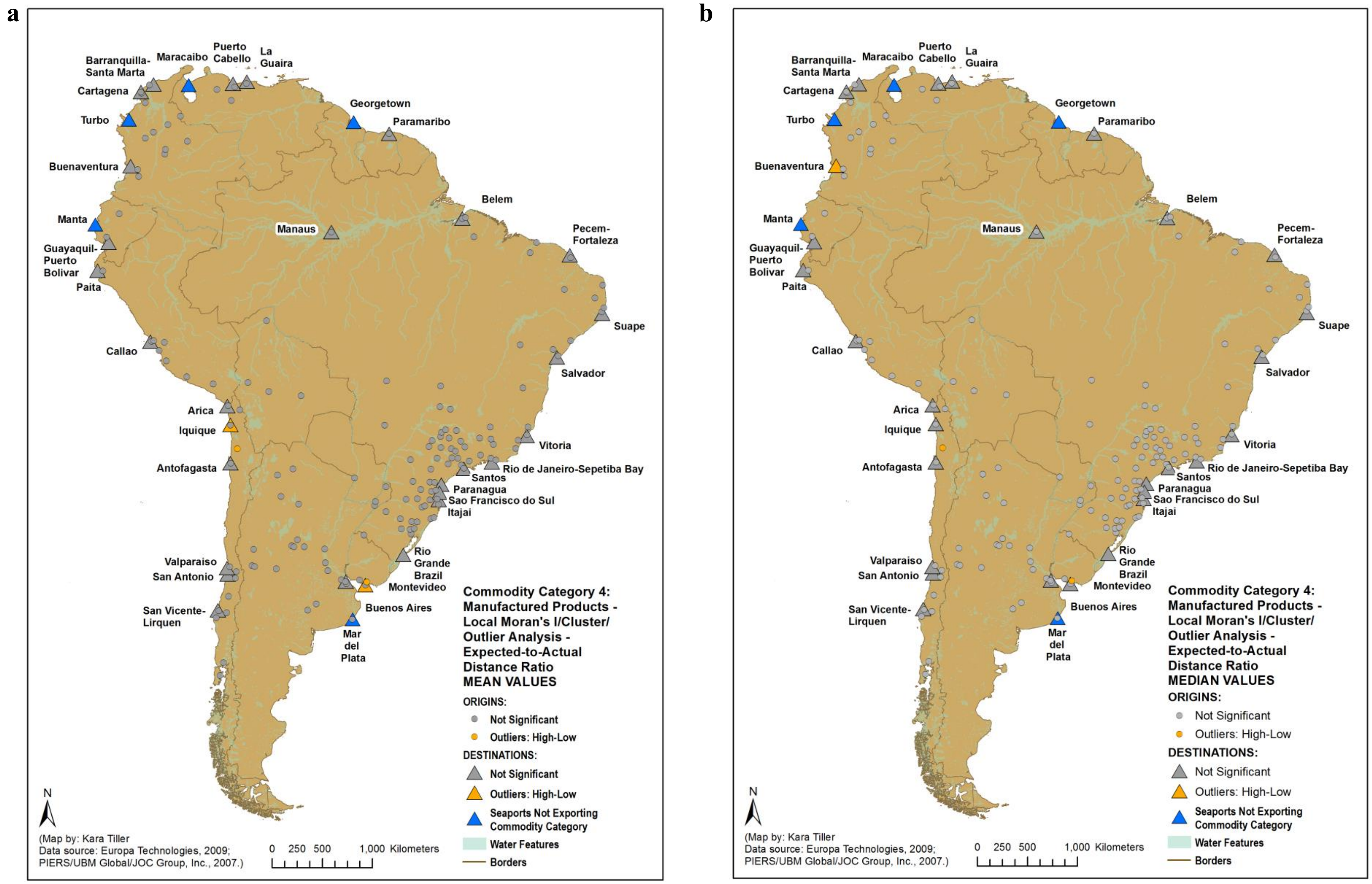


FIGURE 41: Commodity Category 4 (Manufactured Products): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

7.7. Results for Commodity Category 5 (High-Value Manufactured Products)

7.7.1. Global Spatial Autocorrelation Analysis

For the global Moran's I analysis of the Commodity Category 5 (High-Value Manufactured Products) data segment - a total of 272 flows - both mean and median EtA ratios at origins have results which indicate a random pattern (Table 9). Also, both mean and median EtA ratios at destinations have results which indicate a random pattern (Table 9). Thus for both mean and median EtA ratios at origins and destinations, the null hypothesis must be accepted that the spatial pattern of the distribution of EtA ratios is random.

7.7.2. Local Spatial Autocorrelation Analysis

For the cluster and outlier/local Moran's I analysis for High-Value Manufactured Products (Figs. 42.a.-b.), both mean and median EtA ratios at origins have the same result: one high-low outlier, located in close proximity to the port of Valparaiso, Chile. The remaining 195 origins, for both analyses, are found not to be significant in their similarity or dissimilarity.

In the analysis of mean EtA ratios at destinations, two high-low outliers are found, whereas the remaining 30 ports are found not to be significant. These outliers are found at the ports of Valparaiso, Chile and Maracaibo, Venezuela.

The analysis of median EtA ratios at destinations, likewise, also features the high-low outlier of the port of Maracaibo; the remaining 31 ports are found not to be significant.

In viewing the cluster/outlier-context feature selection for the local Moran's I analysis of median EtA ratios, the one high-low origin outlier approximately 30 km from the port of Valparaiso is associated with one extreme high-ratio flow (EtA ratio = 44.68). This makes apparent why the port of Valparaiso is identified as a high-low outlier in the mean analysis, as it is impacted by this extreme value.

The port of Maracaibo, identified as a high-low outlier, is associated with one very high EtA ratio flow, approximately 4 km from the port. This is the only inflow to this port in this commodity category.

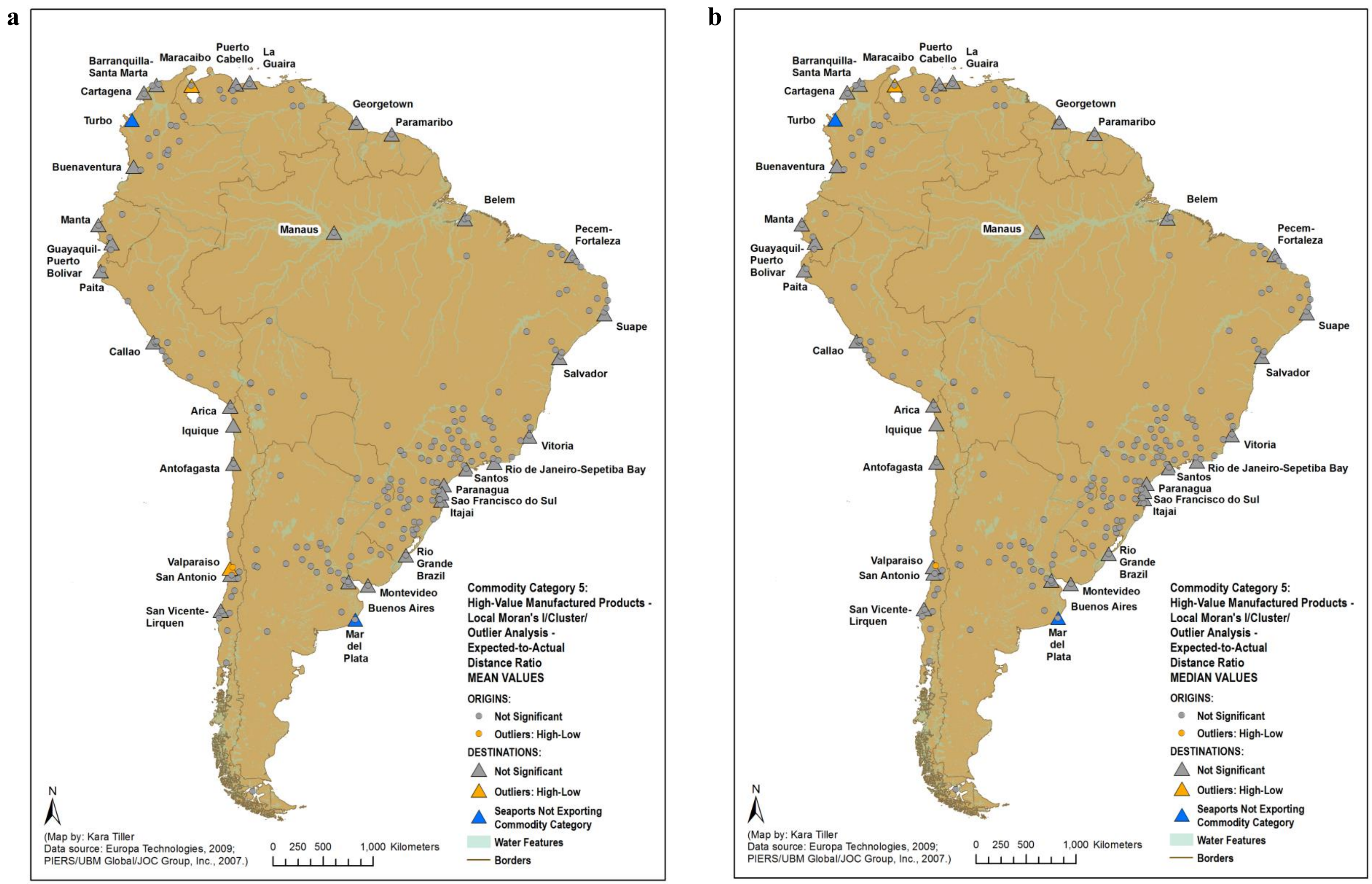


FIGURE 42: Commodity Category 5 (High-Value Manufactured Products): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

7.8. Results for Commodity Category 6 (Miscellaneous/Unable to be Determined Commodities)

7.8.1. Global Spatial Autocorrelation Analysis

For the global Moran's I analysis of the Commodity Category 6 (Miscellaneous/Unable to be Determined Commodities) data segment - a total of 190 flows - both mean and median EtA ratios at origins have results which indicate a random pattern (Table 9). Also, both mean and median EtA ratios at destinations have results which indicate a random pattern (Table 9). However, it should be noted that the number of destinations (28 ports) is low enough in this data segment for this analysis to produce untrustworthy results for destinations.

7.8.2. Local Spatial Autocorrelation Analysis

For the cluster and outlier/local Moran's I analysis for Miscellaneous/Unable to be Determined Commodities (Figs. 43.a.-b.), both mean and median EtA ratios at origins have the same result: one high-low outlier, located in close proximity to the port of Manaus, Brazil. The remaining 144 origins are found not to be significant for both analyses.

As in the global analysis, the low number of destinations for this segment is not ideal for the analysis and may produce untrustworthy results; therefore the results for destination clusters and outliers should be taken in that context.

In the analysis of both mean and median EtA ratios at destinations, two high-low outliers are found, whereas the remaining 26 ports are found not to be significant. These outliers are found at the ports of Manaus, Brazil and Mar del Plata, Argentina.

In viewing the cluster/outlier-context feature selection for the local Moran's I analysis of median EtA ratios, the one high-low origin outlier in close proximity to the port of Manaus, Brazil, is associated with a very high-ratio flow exporting from this port; this is the only outflow for this origin, and the only inflow for this port, in this data segment. Thus it is apparent how the port of Manaus is identified as a high-low outlier.

The port of Mar del Plata, identified as a high-low outlier, is associated with one very high EtA ratio flow, which originates 2.4 km from the port. The same origin has low-ratio outflow to another port, and thus it appears that potentially, the port of Mar del Plata and its immediate vicinity may be an area of focus for trade impedance in terms of miscellaneous/unable to be determined commodities.

a

b

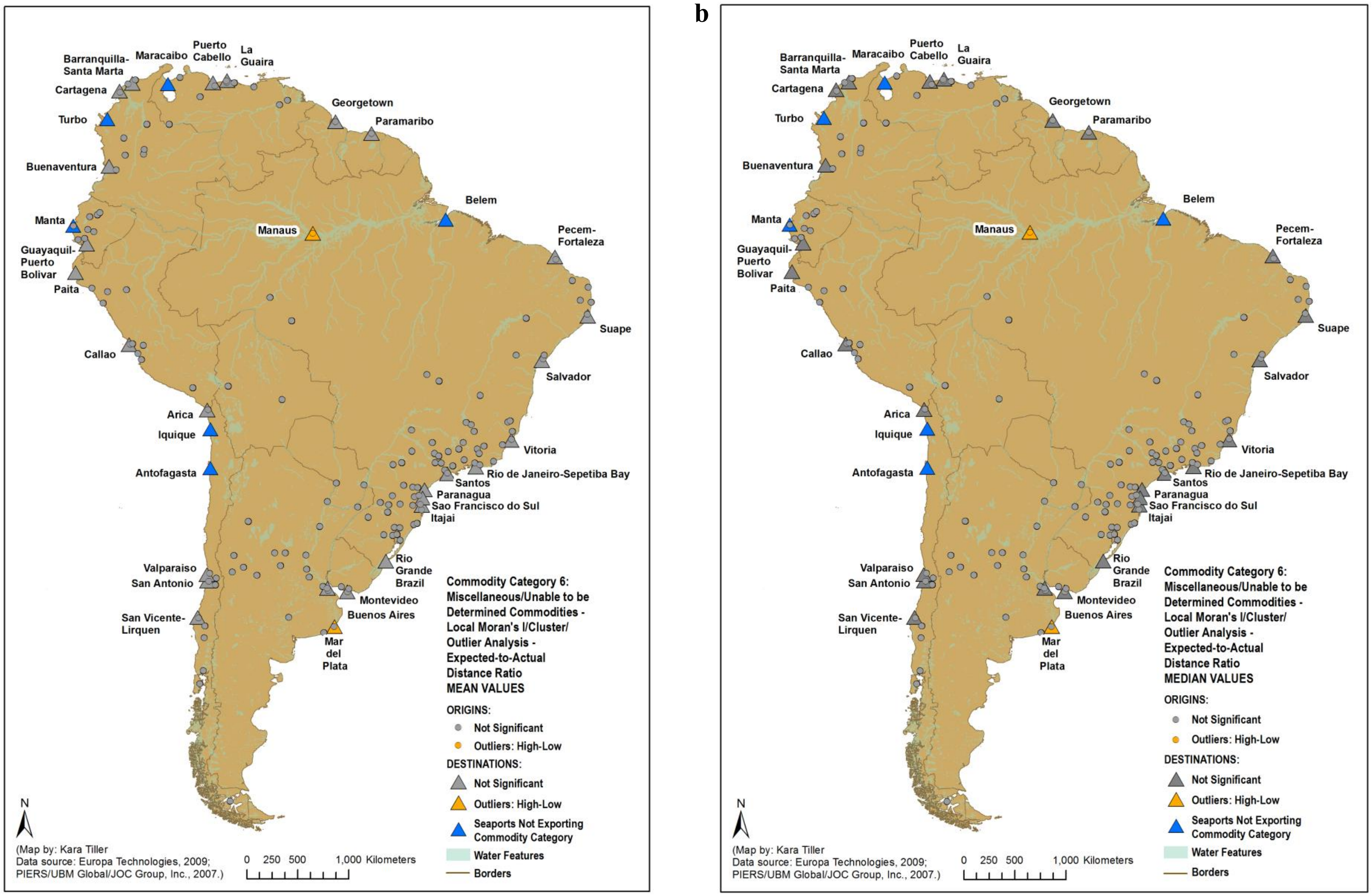


FIGURE 43: Commodity Category 6 (Miscellaneous/Unable to be Determined Commodities): Local Moran's I/cluster/outlier analysis – EtA ratios for flows at origins and destinations: (a) mean values; (b) median values

7.9. Conclusions: Spatial Autocorrelation Analysis

Global spatial autocorrelation analysis of both mean and median EtA ratios at origins indicate clustering, with less than a 1.0% likelihood that the clustered pattern could be the result of chance. The pattern is such that overall similar mean and median values are found in close proximity; positive autocorrelation exists in terms of flow origins for the total dataset. This is evident as well in the total dataset local spatial autocorrelation results for origins; two regions of high-high clusters are apparent – an Asuncion (Paraguay)/western Misiones Province (Argentina)/Rio Parana region and southern Peru/northern Chile region. Through cluster/outlier-context feature selection visualization, each of these is demonstrated to be further divisible into sub-regions. (In addition, a high-ratio outlier is found in Minas Gerais, Brazil.) Global spatial autocorrelation results for both mean and median EtA ratios at destinations for the total dataset indicate a random pattern significant at the 5.0% level, though the median results indicate a tendency toward dispersion.

To provide the most conservative measure with brevity, median results only are presented for commodity-based segment results. Global spatial autocorrelation results for all commodity-based segments show the spatial pattern of both mean and median EtA ratios at origins is random. Likewise, the spatial pattern of both mean and median EtA ratios at destinations is random. For 3 out of 7 segments (non-perishable time-sensitive, manufactured, and miscellaneous products), the number of destinations (16, 29, and 28) is low enough to produce untrustworthy results.

Local spatial autocorrelation results for the commodity-based segments are presented for high-ratio clusters and outliers only. This is because the extent of potential high trade impedance areas of focus – often delineated/punctuated by the occurrence of low-value outliers – has already been presented in the results (and is further summarized in final conclusions, Sect. 9). There are no low-ratio clusters evidenced in the results – only low-ratio outliers. For the commodity-based segments, local spatial autocorrelation results at origins/destinations are as follows.

For perishable products, there are two high-high origin clusters in the Rio de la Plata region, and mixed outlier results in close proximity to Buenos Aires. For non-perishable time-sensitive products, Asuncion (Paraguay) is a high-ratio origin outlier. For heavy/minimally processed commodities, there are five high-ratio clusters/outliers: one origin at Asuncion; two origins in Colombia's northeastern Andes region; one origin at pseudo-port Barranquilla-Santa Marta (Colombia); two destination clusters, at ports of Georgetown and Paramaribo; and one destination outlier at Antofagasta (Chile). For manufactured products, there are two high-ratio origin outliers, one located south of the port of Iquique, Chile, and one proximate to the port of Montevideo, Uruguay. For high-value manufactured products, there is one high-value origin outlier in close proximity to the port of Valparaiso (Chile), and one high-value outlier of the port of Maracaibo (Venezuela). For miscellaneous/unable to be determined commodities, there is one high-low outlier, located in close proximity to the port of Manaus, Brazil.

FINAL CONCLUSIONS

There are four points upon which this research will conclude; each touches upon a unique aspect of this analysis. These include a brief summary of the research and synthesis of the conclusions, and final conclusions regarding the contribution of this research in terms of the study area (particularly regarding the foundational conceptualization of relative trade space). In the next section (Sect. 9), future extensions to the research are detailed, with a focus upon future of the dataset.

8.1. Summary of Research and Conclusions

As the initial launch of a larger research endeavor, this study investigates claims in the regional economic integration literature that Latin American domestic transportation costs stand as a significant barrier to trade and integration. The prescription usually forwarded as a solution is transportation infrastructure improvement, however, such calls are made without benefit of the multi-scalar trade and transportation analyses necessary to investigate if, where, and/or how infrastructure improvement could reduce transport costs. In order to clarify and specify overarching claims of infrastructure improvement needs across modes and throughout the region, evidence of the existence and level of disproportionate trade impedance must be demonstrated to narrow the location of its occurrence, and to profile its spatial pattern.

This research completes these crucial first steps between the identification of high transport costs and the forwarding of the most efficient and effective solutions to address them. This study stands as a precursor to proposed future extensions which will move the investigation toward a finer spatial resolution (sub-continental regions) while simultaneously investigating factors associated with trade impedance throughout South America.

A unique conceptual approach by Plane (1984) is employed to this end. A reconfiguration and reverse calibration of the doubly constrained spatial interaction model as a regression model is used to obtain the result of interest - the expected value of the dependent variable (distance) for each case (O-D flow). In essence, this model outputs *distance as experienced* by trade flows, in a relative conceptualization of space - termed by Plane (1984) variously as generalized transaction cost, functional separation, or predicted, inferred, or expected distance; which Notteboom (2001) identified as revealed accessibility, specifically in terms of landside freight movement; and which Thill (2011) evolved more conceptually as nongeographic or relative distance, with much broader implications and application. This experienced separation may certainly be comprised of impedances other than transportation cost - political-economic, sectoral, and transportation factors have all been suggested in the literature as Latin American trade impediments. Therefore for this application, it is dubbed *trade impedance* (Kockelman & Ruiz-Juri, 2003; Navajas et al., 2010).

This study is built upon the essential foundation of a dataset locationally verified as to actual production origin point locations for containerized waterborne shipments. This dataset was developed through an innovative verification process addressing legacy data constraints in the field.

For this first foray into the use of this unique methodology and dataset, the sub-region of South America is analyzed, as its trade impedance factors are demonstrated in the data to be quite different than those of Central America, Caribbean island nations, and

Mexico. Therefore, this research investigates the existence, level, location, and spatial distribution of disproportionate trade impedance within the landside portion of South American export flows - between shipment origin point and port of export - for one month of U.S.-bound waterborne containerized shipments (October 2006). The relative "trade space" for flows is delineated, and points of particular concern are made apparent.

Upon this foundation, research questions were addressed via three steps. In Steps 1 (Existence) and 2 (Location), the reverse spatial interaction regression is used to obtain expected distance values, and the ratio of this value to the respective actual distance for each flow is visualized/mapped as a measure of the level of trade impedance (EtA, or expected-to-actual distance ratio).

The initial research question related to this step asks how O-D flows by volume are associated with actual distance; this question is answered by the R^2 values for the estimation regression analysis (the reverse doubly-constrained spatial interaction model formulated as a regression model). Results for the R^2 values for the total dataset and all commodity-based segments range between 0.962 and 0.990. This result is as expected.

The second research question involved in this step asks whether expected distance values are disproportionately high as compared to actual physical distance. The vast majority of U.S.-bound export trade flows by volume are found to have trade impedance proportionate to distance (ranging between 79.8% to 98.2% of segment flow volume), therefore the majority of flows are in accordance with the null hypothesis that expected distances are not disproportionately high.

For the total dataset, 8.2% of flows by volume have expected distance (trade impedance) which is disproportionately high. The share evidenced as disproportionately high ranges from 1.3% to 11.2% of total volume for commodity-based segments of the dataset. In general, this is comparable to, though somewhat higher than, the share of volume comprised of low trade impedance flows, with two exceptions. For both the heavy commodities/minimally processed natural resources segment and the miscellaneous/unable to be determined commodity segment, the high trade impedance flows comprise the lowest share of segment total volume (5.8% and 8.8%, respectively); for these segments, low trade impedance flows comprise higher shares of volume (8.8% and 9.0%, respectively).

Additionally, it was hypothesized that for both the heavy commodities/minimally processed natural resources segment and the high-value manufactured products segment, expected distance values would be disproportionately high as compared to actual physical distance. This was not evidenced by the results for the majority of flow volume. However, the high-value manufactured products segment does have the largest share of disproportionately high trade impedance volume among segments, at 11.2% of its segment total volume. All other segments (as well as the total dataset) were hypothesized to have proportionate expected distance – and, for the majority of flow volume in the total dataset and all segments, this is the case.

In terms of mapping and visualizing trade impedance in the location step, generally, proportionate flows evidence a much higher mean volume. Proportionate flows usually also claim the shortest mean flow distance - except in the case of perishable products, high-value manufactured products, and miscellaneous/unable to be determined commodities, for which high-ratio flows have the shortest mean flow distance.

For the total dataset, proportionate flows evidence less port hinterland overlap than do high trade impedance flows. For the total flows dataset and perishable products, a notable result is that the proportionate flows - versus higher-ratio flows - exporting from Santos have differing origins; many higher-ratio flows hail from the west/southwest as far as Asuncion, Paraguay and western Misiones Province, Argentina and as far northeast as the state of Espirito Santo, Brazil, often bypassing other nearer ports to reach Santos. High EtA ratios and hinterland overlap of flows appear to occur together, whereas hinterland overlap/near-port bypass is far less evident for proportionate-ratio flows. Also apparent is that for low-ratio flows in the total dataset, the lowest-ratio flows (except in terms of Chilean flows) appear to suggest coastwise movement, which may suggest that the mode is not landside/surface transport for these flows, and therefore that low-ratio flows may be a smaller portion of the dataset by volume. A creative means of determining mode from the data (if possible) could assist with further determination.

Additionally, all port-of-export countries in the manufactured data segment have percentages ranging from 90.1% (for Ecuador) to 100.0% (for Suriname) of country flow volume total in the proportionate ratio category, which is a higher (and more narrow) range than for other commodity categories. In the high-value manufactured segment, higher-ratio flows are shorter and concentrated in Brazil's southeastern port region; overlapping hinterlands/near-port bypass among both high- and low-ratio flows is apparent here, and flows appear to demonstrate an avoidance of the port of Rio Grande, Brazil.

In Step 3 (Spatial Pattern), mean and median trade impedance (EtA) ratios for flows were assigned to their associated endpoints (to both origins/points of production, then destinations/ports of export). Global spatial autocorrelation analysis was performed to determine the overall spatial distribution of trade impedance ratio means/medians at origins and destinations, and local spatial autocorrelation analysis was performed to identify clusters and outliers.

For the total dataset, global spatial autocorrelation analysis of both mean and median EtA ratios at origins indicate clustering, with less than a 1.0% likelihood that the clustered pattern could be the result of chance. The pattern is such that overall similar mean and median values are found in close proximity; positive autocorrelation exists in terms of flow origins. This is evident as well in the total dataset local spatial autocorrelation results for origins; two regions of high-high clusters are apparent – an Asuncion (Paraguay)/western Misiones Province (Argentina)/Rio Parana region and southern Peru/northern Chile region. Through cluster/outlier-context feature selection visualization, each of these is demonstrated to be further divisible into sub-regions. Global spatial autocorrelation results for both mean and median EtA ratios at destinations for the total dataset indicate a random pattern significant at the 5.0% level, though the median results indicate a tendency toward dispersion.

Global spatial autocorrelation results for all commodity-based segments show the spatial pattern of both mean and median EtA ratios at origins and destinations to be random for all commodity-based segments. For 3 out of 7 segments (non-perishable time-sensitive, manufactured, and miscellaneous products), the number of destinations (16, 29, and 28) is low enough to produce untrustworthy results. Thus destination results for these commodity categories are not presented in the synthesis below.

A synthesis of results hailing from the cluster-outlier feature selection based upon local spatial autocorrelation results is presented in Table 10 and Fig. 44. To provide the most conservative measure with brevity, median results only are synthesized here. Potential areas of focus for disproportionately high trade impedance in landside export flows from origin to port of export are presented by commodity category.

TABLE 10: Summary of potential areas of focus for disproportionately high trade impedance in landside export flows from origin to port of export, by commodity category (October 2006 containerized waterborne U.S. imports originating from South America - Data source: PIERS/UBM Global/JOC Group, Inc., 2007)

Data segment	**Potential areas of focus for disproportionately high trade impedance**
Total Dataset (All Shipment Flows)	1.Area between Asuncion, Paraguay; Western Misiones Province, Argentina, and the ports of Itajai and Santos (Brazil) 2.Area between Asuncion, Paraguay; Goya, Argentina on the Rio Parana; and port of Montevideo, Uruguay 3.Area between high trade impedance origin outlier – in the area of Montes Claros (Minas Gerais), Brazil – and the ports of Salvador and Rio de Janeiro-Sepetiba Bay pseudo-port 4.Area north of the port of Arica, Chile; bounded by Arequipa, Peru, and south of the port of Callao, Peru 5.Area south of the port of Iquique, Chile, approximately bounded by the area of Calama, Chile
1: Perishable	6.Area between the ports of Buenos Aires, Argentina, and Montevideo, Uruguay and two high trade impedance origin clusters - 205 km west of Buenos Aires and 335 km northwest of Montevideo on the Rio Uruguay
2: Non-Perishable Time-Sensitive	7.Area between Asuncion, Paraguay, and Buenos Aires, Argentina
3: Heavy/ Minimally Processed Nat. Resources	8. Area between Cartagena and pseudo-port Barranquilla-Santa-Marta 9.Area between Cartagena and high-ratio origin cluster and outlier in Colombia's northeastern Andes region (east of the Rio Magdalena, near Cucuta) 10.Area between Asuncion, Paraguay, and Santos and Itajai, Brazil 11.Area in proximity of the port of Georgetown, Guyana 12.Area in proximity of the port of Paramaribo, Suriname 13.Area in proximity to the port of Antofagasta, Chile
4: Manufactured	14.Area south of the port of Iquique, Chile, approximately bounded the area of Calama, Chile 15.Area proximate to the port of Montevideo, Uruguay
5: High-Value Manufactured	16.Area proximate to the port of Valparaiso, Chile 17.Area proximate to the port of Maracaibo, Venezuela
6: Miscellaneous/ Undetermined	18.Area proximate to the port of Manaus, Brazil

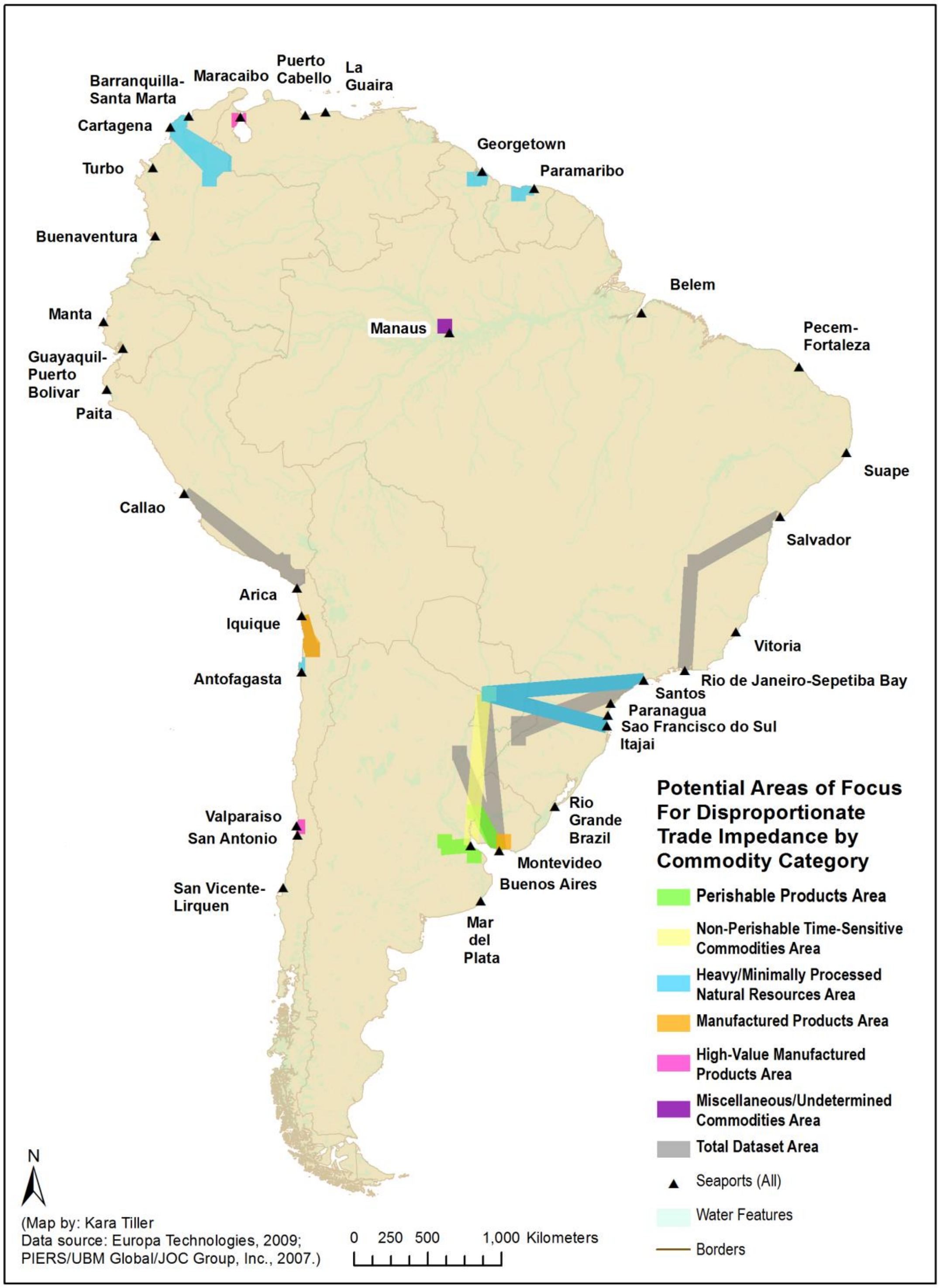


FIGURE 44: Visualization of potential areas of focus for disproportionately high trade impedance in landside export flows from origin to port of export, by commodity category (October 2006 containerized waterborne U.S. imports originating from South America - Data source: PIERS/UBM Global/JOC Group, Inc., 2007)

8.2. Conceptual Contribution: Relative Space in South American Trade

Recently, the IDB has produced an answer to its previous research – 2013's *Too Far To Export* (Mesquita Moreira et al., 2013). This research appears to bear some similarities to the present research, even in terms of phraseology. It is natural for a comparison to be discussed.

This volume states the following about the impetus for this research in Latin America and the Caribbean (LAC):

> Firms with the resources and skills to produce goods in high demand by regional or world markets face high domestic transport costs that destroy their competitive advantages, literally along the road. Or they refrain from exporting altogether. In many cases, prohibitive transport costs prevent them from even opening their business. As such, they are invisible to policymakers and researchers, but represent very real missed opportunities for regions that are usually at the bottom of the income distribution. (p. 2)

Further similarities to this research include claims of using an "unprecedented" or "novel" dataset, seemingly echoing the words used in the 2008 through 2012 presentations related to the present research (listed in Appendix G) about the locationally-verified PIERS data used in this research.

The IDB's data source is customs transaction data from each of the countries' own agencies, improved in terms of locational verification methods which varied by country, by means of using carefully chosen samples, municipal censuses, firm directories, and the like. Ad valorem costs of a good were calculated along a least-cost network route from municipality to customs. The IDB study uses a constructed estimation of transport cost (constructed from variable cost and time factors, weight, and value, taking into account elevation and adjusted for road type), which is based upon the likely least-cost route distance. This is compared to the distribution of known exports. A beneficial similarity is the IDB's focus on the landside, using a different dataset, for a selected set of countries.

Therefore, as it happens, research into similar flow type and study areas was conducted in tandem – but from opposite perspectives - which may be mutually beneficial.

First, it appears in the study's initial presentation of maps of transport cost versus exports, that these are nearly reciprocal in the study area. It appears that, generally, high transport costs are occurring in areas further from main ports. This seems to relate back to the intensive dependence upon distance used in the construction of transport cost (as both variable cost types are in the end based upon distance). Different methodologies and variables are variously used in each country in an attempt to control for the effects of origin comparative advantage and institutions, dimension and transportability of products, port specialization, and product category (manufacturing, agriculture, and mining). Other strategies involved using known effects of a Chilean earthquake upon the road network; replacing transport cost by a non-export volume correlated factor, such as presence of the Inca road network; and straight-line distance approximation to port (as in this study) in Brazil and Colombia. Effects of transport costs on export diversity were also explored. Planned infrastructure projects were part of the analysis, as well.

Yet the authors acknowledge that what they term logistics costs (route congestion and warehousing) are not included (p. 10). Particularly, the effects of congestion – which may occur in many of the areas identified as having higher shares of exports – would seem to be an essential factor to include, as the authors suggest that the dispersion that would normally relieve the negative congestion effects of urban agglomeration are prevented by the lack of quality transport infrastructure at the periphery (p. 17).

Additionally, much of the discussion of results focuses on the suggestion that there are latent export opportunities waiting for the opening of better infrastructure. This may be the case, but much hinges upon that which cannot be evidenced.

What can be evidenced is trade volumes. It would seem that to rely on what is known to reveal the structure of that which is not - trade impedance – may more directly reveal what is at the core of interest: to what degree, and where, are flows experiencing greater resistance than what would be expected? A clear delineation of the impedance that is experienced by flows may be a more reliable basis than the repeated construction of transport costs in an attempt to determine their relationship with exports.

In this way, all that is experienced from the perspective of the feature is delineated in trade space already; what remains is to determine which variables appear to be most highly associated with its occurrence.

One major difference between the studies is the data used – while the IDB's datasets use a locational verification of customs transaction data for each country, for a selection of countries, in various years – the data used here includes all South American countries, using the same time period across the dataset (the larger database being a year-long snapshot of U.S. imports from the entire globe). Yet, to its benefit, the IDB dataset is selected from a universe of all exports, not just the subset which comprises U.S.-bound shipments.

This is a potential future addition; however, the intensive process required to clean and process the data via manual correction – to locationally verify the origins –

requires time and resource budgets which are likely be unrealistic for the majority of researchers. The overarching data issue is addressed below.

The authors note that causality may be problematic in terms of interpretation, as low transport costs may beget higher flow volume, but higher flow volume lowers transport costs by producing economies of scale (p. 4). This can be set as a caution for interpretation in the present research as well. In the present dissertation research, the vast majority of the dataset consists of flows with proportionate EtA ratios. Proportionate flows have the highest average volume for the total dataset, and throughout its commodity segments. Though this is fertile ground for the emergence of such, care should be taken not to attach any assumed causality to the large proportion of proportionate flows. The association of flows' ratios with trade-related variables has not yet been explored and stands as a future extension – thus little can be said other what has been said here related to the existence of flows with particular trade impedance ratios, their location, and the spatial distribution of their central tendencies at origins and at ports. While these results may suggest particular phenomena, such ideas must be set in the framework of a future research extension.

Nonetheless, the present results are an initial step forward for research in "trade space."

FUTURE EXTENSIONS

9.1. Planned Future Extensions

In comparing Latin America's transport costs to those of other exporters to the U.S. , the IDB presents that freight rates to the U.S. are higher from countries in the LAC than from other world regions such as the Far East or Europe. IDB authors state (2010):

> As with import freights, little can be said about the determinants of these results…The general conclusion is that proximity does not always transfer into lower freight rates… The burning question is, what drives the high level of transport costs in Latin America? Answering this question involves isolating the role of a number of complex and interrelated issues, ranging from the quality of infrastructure services to distance, scale, and market structure. (p. 109)

As stated in Sects. 1 and 2.1, the overall question about ameliorating trade impedance via transportation infrastructure improvement cannot be addressed without – at minimum – the intermediate Step 4 (Association Step) suggested here, and subsequent investigation of trade impedance at a finer scale. Therefore this step stands as a future extension of this analysis, intended to be followed by sub-continental/regional-scale trade and transportation analysis in corridors identified by Step 4 (Association Step).

Multiple regression is the appropriate method for this second portion of the analysis, as the end goal is to determine the relationship between trade impedance and several previously identified independent variables which stand as potentially trade impeding factors (presented in operationalized form in Table 11). As demonstrated in Fig. 5, these independent variables are varied in nature, according to the potential trade barriers identified in the literature. Thus regression is helpful to identify the level of association that each variable has with the dependent variable of disproportionate trade impedance, and to reveal the relative contribution of each independent variable to disproportionate trade impedance.

Trade impedance (EtA ratio) for each O-D flow will comprise the dependent variable in multiple regression analysis. Again, as the actual routes of shipments cannot be determined from the data, the straight line (geodesic) from origin to port is an approximation. In order to identify sub-continental trade corridors as a finer geography for further research, dual goals will be fulfilled by representing the same width of buffer on both sides of originally identified O-D flow lines, as that used for the grid cell size for aggregating the shipments at their origins. Therefore the dependent variable will be EtA ratios for O-D flows - now represented as O-D trade flow corridors of 100 kilometers in total width, or 50 kilometers on each side of the O-D flow line.

With EtA ratios set as the dependent variable, trade impedance so defined will be investigated as to its association with the independent variables comprised of potential trade barriers alleged to be contributors to sectoral, political-economic, and transportation factors of trade impedance. The suite of independent variables identified here represent a sufficient first-cut attempt to engage this process for the initial examination of factors potentially associated; however, their ideal representation may be conditioned by availability of data. This step represents the first attempt to identify factors important in relative trade space, whereas previously such factors' association has been investigated in terms of absolute space.

Additionally, some variables which had been previously proposed for this step must be reconsidered; based upon feedback to this research proposal and the work already conducted in this area, other variables should be added after further consideration about their form and operationalization. These include:

- **Slope/Elevation**: This variable has been suggested as an addition to the suite of variables in this research proposal. Recently, elevation was included in the trade studies underlying the IDB's *Too Far to Export* (2013) as a categorical variable based upon gradient (0-3%, 3-5%, and 5-7%). An illustrative example was made of wood products producers in Peru's Selva region shipping to the port of Callao via roads winding through the very high elevations of the Andes. This impacts the suggested path forward here in two ways:

  (1) The association of variables related to the physical environment must be included as potentially associated with trade impedance, and the specification, form, and operationalization of such variables must be determined (potentially impacting model selection); and

  (2) Conceptually, the level of the physical environment must be integrated into the Multi-Level Approach to Landside Trade Impedance (Fig. 5). A review of the literature could elucidate the physical/environmental factors which most impact the movement of goods, and a complete identification of these should inform this addition to the conceptual diagram.

- **Trade organizations:** This variable was originally conceived of as a set of dummy variables representing the various trade organizations operating within South America (Mikuriya, 2011; Limão and Venables, 2001, p. 453;

Lawrence et al., 2008, p. 5). It was originally suggested that this variable identify the trade organization affiliations of the country of shipment origin. However, in terms of modeling flows as origin-destination corridors, an improved specification of this variable is under consideration.

- **State Borders Crossed:** As identified by Daumal and Zignago (2010) in terms of Brazilian trade, sub-national borders stand as an important consideration, especially in research in which the goal is to move to a finer scale of analysis.

- **Quality of transportation infrastructure:** This variable was originally conceived of as a variable representing – and attempting to parse the effects of - national transportation infrastructure quality differences (in terms of country of shipment origin), as a comparison to dummy variables indicating country of shipment origin (in that differences between the two may separately suggest transport infrastructure-related impacts) (Limão and Venables, 2001, p. 451; Clark, Micco, & Dollar, 2004, p. 423; Mesquita-Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11.). This variable was to be included in the form of an author-identified binary identification (if infrastructure is poor, or not) based upon the literature. Since originally proposed, this variable appears to require reconsideration both due to changes in the proposed structure of the dependent variable (modeling flows as origin-destination corridors), as well as the realization that one assessment for a nation may be overgeneralizing or inaccurate. The sub-continental and potentially sub-national identification of this factor may be more appropriate, and along with the explicit identification of mode type, these are both necessary improvements here. It is important to note that while country-specific assessments of overall transportation infrastructure quality assist in narrowing the area of focus, Mesquita Moreira et al. (2013) cite Behrens (2011) in presenting that variance in infrastructure quality *sub-nationally* appears to be associated with country-level economic divergence. This variable could also encapsulate such concepts as well.

TABLE 11: Independent variables for proposed research extension - Step 4/Association Step (with reference to Multi-Level Approach to Trade Impedance [Fig. 5])

Level	**Variable Type**	**Operationalization**	**Data Source**	**Concept Represented**	**Conceptual Citation**
Infrastructure	Road Density	Road length (km)/ Total area (km^2)	- Global Roads Open Access Data Set (gROADS), v1 (1980–2010) -CODATA Catalog of Roads Data Sets, Version 1	Road density; connectivity.	(Limão and Venables, 2001, p. 453; Calderón and Servén, 2004; Fay & Morrison, 2007, pp. 85-86; Mesquita Moreira, et al, 2013.)
Infrastructure	Roads: Proportion by Type	Local/Urban road length (km)/Total length road length (km)	See above.	Connectivity.	(Mesquita Moreira, et al, 2013.)
Infrastructure	Lack of modal competition (Roads vs. Rail)	Total road length (km)/ Rail buffer area in O-D corridor (km^2) (Applicable rail lines will be buffered to represent service area – 100 km each side of track and endpoints.)	See above and Global Insight Plus (v. 2009.2) (Europa Technologies, 2009).	Modal competition within O-D corridor.	(Limão and Venables, 2001, p. 453; Mesquita-Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11)
Transportation	Port of Export Landside Congestion	TEU per Gate Hour (annual)	Containerisation International Yearbook 2009 and 2007 (and other sources, such as port websites, as needed).	Reliability of transport chain.	(Notteboom, 2001, p. 14.)

Table 11 (continued)

Level	**Variable Type**	**Operationalization**	**Data Source**	**Concept Represented**	**Conceptual Citation**
Transportation	Urban Congestion	Urban sprawls total area (km^2)/Total area (km^2)	Global Insight Plus (v. 2009.2) (Europa Technologies, 2009).	Reliability of transport chain.	(Notteboom, 2001, p. 14; OECD/ECMT, 2007, p. 37; Inter-American Development Bank, 2010, p. 26; p. 31; Mesquita Moreira, et al, 2013.))
Location (With Respect To Central Places)	Lack of modal competition (Inland Waterway)	Proportion of O-D corridor in river port service area (Applicable river ports will be buffered 100 km.)	Ports data from National Geospatial Intelligence Agency (2006); other sources as needed.	Modal competition within O-D corridor.	(Limão and Venables, 2001, p. 453; Mesquita-Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11)
Location (With Respect To Central Places)	Carrier Competition	Number of direct-call liner services at port of export.	Containerisation International Yearbook 2009 and 2007.	Competition among carriers at port of export.	(Clark, Micco, & Dollar, 2004, p. 423; Mesquita-Moreira, Volpe, & Blyde, 2008, p. 127; Inter-American Development Bank, 2010, p. 11)
Location (With Respect To Central Places)	Transshipment Connectivity Index	Hard data (score) or rank (out of 118)	World Economic Forum (WEF) The Global Enabling Trade Report 2008 – UNCTAD, Transport Section, Trade Logistics Branch, 2006.	Connectivity of port of export country.	(Lawrence et al., 2008, p. 394; UNCTAD, 2006).

Table 11 (continued)

Level	**Variable Type**	**Operationalization**	**Data Source**	**Concept Represented**	**Conceptual Citation**
Location (With Respect To Central Places)	Country of Shipment Origin	11 Dummy variables representing 12 countries (1 if =Country of origin; 0 otherwise)	PIERS TI (Trade Intelligence) U.S. Waterborne Imports dataset – Latin American subset (PIERS/UBM global/JOC Group, Inc., 2007).	For comparison with Quality of Transportation Infrastructure (isolating factors).	(Clark, Micco, & Dollar, 2004, p. 429)
Location (With Respect To Central Places)	Country of Port of Export	11 Dummy variables representing 12 countries (1 if ={Port of Export Country; 0 otherwise)	See above.	For comparison with Quality of Transportation Infrastructure (isolating factors).	(Clark, Micco, & Dollar, 2004, p. 429)
Location (With Respect To Central Places)	Landlocked Country of Origin	Dummy variable, 1 if Country of origin is landlocked (0 otherwise)	See above.	Connectivity.	(Limao and Venables, 2001).
Trade Policy	Number of Memberships (country of shipment origin)	Number of trade organization membership(s)/affiliations of country of origin.	Organization of American State's Foreign Trade Information System (SICE).	Compatibility of landside portion of transport chain.	(Mikuriya, 2011; Limão and Venables, 2001, p. 453; Lawrence et al., 2008, p. 5)
Trade Policy	Number of Memberships (port of export country)	Number of trade organization membership(s)/affiliations of port of export country.	See above.	Compatibility of landside portion of transport chain.	(Mikuriya, 2011; Limão and Venables, 2001, p. 453; Lawrence et al., 2008, p. 5)

Table 11 (continued)

Level	Variable Type	Operationalization	Data Source	Concept Represented	Conceptual Citation
Customs	Corruption Perceptions Index	For country of shipment origin, Score of perception of public sector corruption on a 1-10 scale (10 = least corrupt).	Transparency International – Corruption Perceptions Index 2006.	Public sector corruption in country of shipment origin.	Berríos (2010); (Lawrence et al., 2008)
Customs	Irregular Payments in Imports and Exports	For country of port of export, scoring of frequency of cash bribes made by firms to obtain import and export permits by country (executive opinion of frequency of occurrence on a 1-7 scale [1 = is common; 7 = never occurs]).	World Economic Forum (WEF) –The Global Enabling Trade Report 2008 - WEF Executive Opinion Survey 2006-2007.	Customs corruption (cash bribery)/ Customs inefficiency in country of port of export.	(Manzetti and Blake, 1996; Lane, 1998; Wei, 2000); McLinden, 2005; Thachuk, 2005; Ferreira, Engelschalk, and Mayville, 2006; Lawrence et al., 2008)
Customs	Number of National Borders Crossed	This will be determined from differences in Country of Shipment Origin and Country of Port of Export variables.	PIERS TI (Trade Intelligence) U.S. Waterborne Imports dataset – Latin American subset (PIERS/UBM global/JOC Group, Inc., 2007).	Simple administrative delay inherent in border crossing.	(Lawrence et al., 2008)

Other extensions and improvements which may be of benefit to this research include conducting a sensitivity analysis in terms of grid cell size for shipment origin aggregation; exploring the spatial pattern of O-D flows via paired-point cluster correspondence (as in Lu and Thill, 2003); eventually including exports to all trading partners; and modeling the trade space of flows over various time horizons (Ortúzar and Willumsen, 2001, p. 433; Inter-American Development Bank, 2010) as well as various seasons (Chambers, 2012). Using the dataset detailed here, it may be possible to do so from month to month, whereas extensions of this dataset may provide annual change.

Additionally, it is a goal of the extensions to this research to expand the study area to include all of Latin America - addressing the particular realities of shipments originating from Central America (such as shipment flows which are cross-continental or which pass through the Panama Canal); the Caribbean (such as transshipment, and island geography/scale issues); and Mexico and Puerto Rico (in terms addressing the overall trade context interpretation of their waterborne shipments to the U.S.) - as was discussed in Sect. 1.1.

As outlined in Fig. 10, in the context of a larger research plan, Step 4 (Association Step) is the concluding step in this research which has been proposed to further explore and specify the occurrence of trade impedance in South America at a finer spatial scale, while highlighting the particular factors with which its occurrence may be associated. In investigating its occurrence in this way, regions of interest for more detailed study may emerge, enabling a full suite of trade and transportation analysis to be performed in answer to original suggestions in the literature regarding transportation infrastructure improvement. This next step brings the analysis closer to the scale at which transportation infrastructure analysis may be approached.

Conceptually, it is hoped that this methodology can be improved upon and forwarded to model relative space, as conceptualized by Plane (1984) and Thill (2011) for various types of phenomena.

9.2. The Future of the Data: Trade Space and "Absolute Space" Data Constraints

The data constraints expressed by IDB authors in *Too Far To Export* (2013) (Mesquita Moreira et al., 2013) are legion in this kind of trade research. It seems that the IDB researchers, in conducting an analysis using local data sets, also found their work to be fraught with data constraints. As an enormous amount of data work was performed (locational verification, processing, and cleaning) in order to even make possible this analysis; it stands as a very large part of the preliminary work involved in such an analysis.

This is one reason that locational verification of the PIERS dataset was undertaken, despite the copious amount of manual data-cleaning that underlies the resolution for difficult-to-verify records in each country/continent; the cleaned data can be used then, again, by machine-learning algorithms which can then properly automate the processing of further records (likely to have the same shippers and origins occurring repeatedly in the data, to some degree). The manual data cleaning thus improves the effect of running the algorithms, so all data cleaning improves quality of the data for future use in further research. This is extremely valuable, as apparently even global economic organizations charged with such analysis do not have access to this type of locationally-verified data. Without this process of locational verification, there does not appear to be another option for door-to-door locational verification of shipments at this time.

However, with the onset of the IoT (Internet of Things), such data constraints may become a thing of the past (and considering current technological capabilities, perhaps should already be). When such technology becomes pervasive, today's data solutions for trade flow may quickly become irrelevant. (Its vestiges, however, may be useful for the purposes of developing and archiving historical trend data – for its own sake, and for the purpose of projection.) Even the use of some local agency data may eventually be replaced by the outputs of this technology, which can simultaneously provide far more standardized locational and data output with little in the way of additional processing required, especially in anticipation of more standardized protocol.

The IDB speaks with pride about its "hard-won" data (p. xiii), as does this research. Yet the truth is, at this point in time – or in the near future - data involving moving objects does not need to be so hard to win.

The technology has already arrived: IoT, smart objects, ubiquitous computing – not yet having reached their full development or prime in the market, these are still old news in terms of technological capabilities. With the vast potential for commercial benefits, it is generally accepted that pervasive use of these technologies for a world of objects is perhaps a decade away, with overwhelming financial benefits likely to trump any concern about its moral or security implications.

The IDB's data involves actual transport costs, when available, which are estimated/derived when not. When actual cost data is needed, there may be no good substitute. But when these are estimated or derived, it appears this technology can provide an automated feed for variables such as distance and route – the desired information enrichment for the data used here. These technologies foster a different reality in terms of connectivity and widespread market adoption than technologies which came before. It is expected that these technologies will create richer, more standardized, and far more connected, accessible, affordable, and voluminous data output.

While the shipment data resulting from these pervasive technologies may still be subject to protection from widespread public availability due to its commercial nature, by the mere fact of its more automated, lower-cost, connected nature, the state of the data – and associated services and products – will likely follow suit. At that time, the hand-wringing over data constraints in trade research may ease considerably. Such data will soon become easier to obtain – eventually, perhaps, a seamless result of the movement of objects, harvested automatically.

Though this may seem cast a shadow over the significant budget and time invested in the dataset development, it is important to recognize that it is an essential precursor to attempting the methodology used here, based upon a conceptualization of relative space, as advanced by Plane (1984) and Thill (2011). Just as Plane's innovative cartograms forwarded concepts predating the sophisticated software that would later be used for its progeny, the essential data verification work described here underlies crucial initial explorations of the relative space-based methods.

These investments ensure that when seamless data feeds arrive, the methodology will be ready.

REFERENCES

Anselin, L. (1988). Spatial econometrics: methods and models. Dordrecht: Kluwer.

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis, 27*(2): 93-115.

Batty, M. (1976). Urban modelling: algorithms, calibrations, predictions. Cambridge, UK: Cambridge University Press.

Batty, M., & Sidkar, P.K. (1982). Spatial aggregation in gravity models, *Environment and Planning A, 14*, 377-405; 525-583; 629-658; 795-822.

Behrens, K. (2011). International integration and regional inequalities; how important is national infrastructure? *The Manchester School*, *79*(5), 952–971. Mesquita Moreira, M., Blyde, J., Volpe, C., & Molina, D. (2013). Too far to export: domestic transportation costs and regional export disparities in Latin America and the Caribbean (Special report on integration and trade). Washington, D.C.: Inter-American Development Bank. Retrieved from: http://www10.iadb.org/intal/intalcdi/PE/2013/13227en.pdf

Bergstrand, J. (1985). The gravity model in international trade: some microeconomic foundations and empirical evidence. *Review of Economics and Statistics*, *67*, 474-481. As cited in Clark, X., Dollar, D., & Micco, A. (2004). Port efficiency, maritime transport costs, and bilateral trade. *Journal of Development Economics*, *75*, 417– 450.

Berríos, R. (2010). Corruption as a drag on development, *Latin American Research Review*, *45*(2), 245-252.

Bhagwati, J. (1995). *US trade policy: the infatuation with FTAs* (Discussion Paper No. 726). New York, NY: Columbia University. Retrieved July 03, 2011, from http://globalcitizen.net/Data/Pages/1457/papers/20090222131937814.pdf

Bulmer-Thomas, V. (2003). *The economic history of Latin America since independence* (2nd ed.). Cambridge, UK: Cambridge University Press.

Calderón, C., & Servén, L. (2002). The output cost of Latin America's infrastructure gap. In W. Easterly & L. Servén (Eds.), *The limits of stabilization: infrastructure, public deficits, and growth in Latin America.* Palo Alto, CA and Washington, DC: Stanford University Press and World Bank. As cited in Fay, M., & Morrison, M. (2007). *Infrastructure in Latin America and the Caribbean: Recent developments and key challenges.* Washington, DC: World Bank.

Calderón, C., & Servén, L. (2004). The effects of infrastructure development on growth and income distribution. (World Bank Policy Research Series, Working Paper

3400). Washington, DC: World Bank. Retrieved July 5, 2011 from http://ideas.repec.org/p/wbk/wbrwps/3400.html

Carrothers, G.A.P. (1956). DISCUSSION: Gravity and potential models of spatial interaction. *Papers in Regional Science*, *2*(1), 196-198. As cited in Wilson, A.G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning*, *3*: 1-32.

Chambers, M. (2012). Transportation trends in focus: containerships carry inventory for U.S. retailers. U.S. Department of Transportation, Bureau of Transportation Statistics, Research and Innovative Technology Administration (September 2012).

Chisholm, M.. & O'Sullivan, P. (1973). *Freight flows and spatial aspects of the British economy*. Cambridge, UK: Cambridge University Press.

Clark, X., Dollar, D., & Micco, A. (2004). Port efficiency, maritime transport costs, and bilateral trade. *Journal of Development Economics*, *75*, 417– 450.

Cliff, A.D., & Ord, J.K.. (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, *4*(3): 267-284.

Cliff, A.D., & Ord, J.K.. (1981). Spatial processes: models and applications. London: Pion.

Combes, P.-P., & Lafourcade, M. (2005). Transport costs: measures, determinants, and regional policy. implications for France. *Journal of Economic Geography, 5*(3), 319-349.

CONAIE (Confederación de las Nacionalidades Indígenas del Ecuador). (1994). "Proyecto político de la CONAIE." MS. As cited in Jameson, K.P. (2010). The indigenous movement in Ecuador: The struggle for a plurinational state. *Latin American Perspectives*, *176*(38), 63-73.

Containerisation International. (2009). Containerisation International Yearbook 2009. London, UK: Informa UK.

Daumal, M., & Zignago, S. (2010). Measure and determinants of border effects of Brazilian states. *Papers in Regional Science*, *89*(4): 735-758.

De Castro, C.F. (1996). *Trade and transport facilitation: review of current issues and experience*. (SSATP Working Paper 27). Geneva: UNCTAD and Washington, D.C.: World Bank. As cited in Ferreira, C., Engelschalk, M., & Mayville, W. (2006). The challenge of combating corruption in customs administrations. In Campos, J.E., & Pradhan, S. (Eds.), *The many faces of corruption: Tracking vulnerabilities at the sector level* (pp. 367-386). Washington, DC: World Bank.

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, *90*(1), 37-48. As cited in Lawrence, R.Z., Blanke, J., Hanouz, M.D., & Moavenzadeh, J. (2008). *The Global Enabling Trade Report 2008*. Geneva, Switzerland: World Economic Forum.

Dorigo, G., & Tobler, W. (1983). Push-pull migration laws. *Annals of the Association of American Geographers*, *73*, 1-17. As cited in Ledent, J. (1985). The doubly constrained model of spatial interaction: A more general formulation. *Environment and Planning A, 17*, 253-262.

Environmental Systems Research Institute, Inc. (ESRI). (2008). ArcGIS 9.3 - Build 1770 (ArcView license) [Mapping and spatial analysis software]. Redlands, CA: Environmental Systems Research Institute.

Environmental Systems Research Institute, Inc. (ESRI). (2011). ArcGIS 10.0 – Build 2414 (ArcInfo license) [Mapping and spatial analysis software]. Redlands, CA: Environmental Systems Research Institute.

Environmental Systems Research Institute, Inc. (ESRI). (2012). ArcGIS 10.1 – Build 3035 (Advanced license) [Mapping and spatial analysis software]. Redlands, CA: Environmental Systems Research Institute.

Europa Technologies Ltd. (2009). Global Insight Plus – Global Data Products (Version 2009.2) [Spatial dataset for GIS applications]. Surrey, UK: Europa Technologies Ltd.

Fay, M., & Morrison, M. (2007). *Infrastructure in Latin America and the Caribbean: Recent developments and key challenges*. Washington, DC: World Bank.

Ferreira, C., Engelschalk, M., & Mayville, W. (2006). The challenge of combating corruption in customs administrations. In Campos, J.E., & Pradhan, S. (Eds.), *The many faces of corruption: Tracking vulnerabilities at the sector level* (pp. 367-386). Washington, DC: World Bank. Retrieved September 16, 2011 from: http://books.google.com/books?id=Wy-oNmjqqQC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Fischer, M.M., & Griffith, D.A. (2008). Modeling spatial autocorrelation in spatial interaction data: an application to patent citation data in the European Union. *Journal of Regional Science, 48*(5), 969–989.

Fotheringham, A.S., & M.E. O'Kelly. (1989). *Spatial interaction models: Formulations and applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fotheringham, A.S., Brunsdon, C., & Charlton, M.E. (2000). *Quantitative geography*. London, UK: Sage Publications.

Fotheringham, A.S. & Wong, D. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A, 23*, 1025-1044.

Friedlaender, A.F. & Spady, R.H. (1980). A derived demand function for freight transportation. *The Review of Economics and Statistics, 62*(3), 432-441.

Fuller, L. (2011). *What goes around... How Brazil has rediscovered the merits of the roundabout: Background story.* [Case study on International Road Assessment Programme (IRAP) website]. Retrieved October 21, 2011 from http://www.irap.net/library/cat_view/5-case-studies.html?start=5.

Gentleman, R. & Ihaka, R. (2011) "R" [Spatial statistical software/language]. Available at: http://www.r-project.org/

Getis, A. (2010). Spatial Autocorrelation. In M.M. Fischer and A. Getis (Eds.), *Handbook of applied spatial analysis: software tools, methods and applications* (pp. 255-277). Heidelberg: Springer-Verlag Berlin Heidelberg.

Green, D. (2003) *Silent revolution: The rise and crisis of market economics in Latin America* (2nd ed.). New York: Cassell Press.

Grossman, G. (1998). "Comment." The regionalization of the world economy, Frankel, J.A. [Ed.] Chicago, IL: University of Chicago for the NBER, 29-31. As cited in Lawrence, R.Z., Blanke, J., Hanouz, M.D., & Moavenzadeh, J. (2008). *The Global Enabling Trade Report 2008*. Geneva, Switzerland: World Economic Forum.

Guerrero, P., Lucenti, K., & Galarza, S. (2009). Trade logistic and regional integration in Latin America & the Caribbean. [IDB (Inter-American Development Bank) Working Paper # IDB-WP-148 (December 2009)]. Washington DC: IDB. Retrieved September 20, 2011 from http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=35031936

Gwynne, R.N. (1990). *New horizons? Third world industrialization in an international framework*. Harlow: Longman. As cited in Gwynne, R.N., & Kay, C. (2004). *Latin America transformed: Globalization and modernity* (2nd ed.). London, UK: Edward Arnold Publishers.

Gwynne, R.N., & Kay, C. (2004). *Latin America transformed: Globalization and modernity* (2nd ed.). London, UK: Edward Arnold Publishers.

Hay, A.M. (1979). The geographical explanation of commodity flow. *Progress in Human Geography, 3*, 1-12.

Haynes, K.E., & Fotheringham, A.S. (1984). *Gravity and spatial interaction models.* Beverly Hills, CA: Sage Publications.

Held, D., McGrew, A., Goldblatt, D., & Perraton, J. (1999). *Global transformations:*

politics, economics, and culture. Cambridge: Polity. As cited in Gwynne, R.N., & Kay, C. (2004). *Latin America transformed: Globalization and modernity* (2nd ed.). London, UK: Edward Arnold Publishers.

Hill, T.H. (2008). Introduction to law and economic development in Latin America: a comparative approach to legal reform. *Chicago-Kent Law Review, 83*(1), 3-23.

Holguín-Veras, J., & Patil, G.R. (2008). A multicommodity integrated freight origin–destination synthesis model. *Networks and Spatial Economics, 8*, 309–326.

Hornbeck, J. F. (2010). U.S.-Latin America trade: Recent trends and policy issues. (Congressional Research Service Publication - 7-5700 - 98-840). Washington, D.C.: Congressional Research Service. Retrieved October 8, 2010, from http://www.fas.org/sgp/crs/row/98-840.pdf

Hummels, D. (2001). *Toward a geography of trade costs*. Unpublished. Department of Agricultural Economics, Purdue University. As cited in Inter-American Development Bank. (2010). Trade and productivity: A route to reallocation with a high transport toll [Electronic version]. In Páges, C. (Ed.), *The age of productivity: Transforming economies from the bottom up* (pp. 97-122). New York, NY: Palgrave Macmillan. Retrieved October 9, 2010, from http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=2270519

Hummels, D. (2007). Transportation costs and international trade in the second era of globalization. Journal of Economic Perspectives, 21(3), 131-154. As cited in Lawrence, R.Z., Blanke, J., Hanouz, M.D., & Moavenzadeh, J. (2008). *The Global Enabling Trade Report 2008*. Geneva, Switzerland: World Economic Forum.

Hummels, D. (2009). Globalisation and freight transport costs in maritime shipping and aviation. (Forum Paper 2009-3 Produced for 2009 International Transport Forum.) Paris, France: Organisation for Economic Co-Operation and Development (OECD)/International Transport Forum (ITF). Retrieved November 26, 2011 from http://www.internationaltransportforum.org/Pub/pdf/09FP03.pdf

IBM Corp. (2011). IBM SPSS Statistics 21.0 (Release 21.0.0.0). [Computer software].

Inter-American Development Bank. (2009, November 27). Latin American economies stymied by high transportation costs, IDB study shows. [IDB website News/Webstories]. Retrieved October 9, 2010, from http://www.iadb.org/features-and-web-stories/2009-11/english/latin-american-economies-stymied-by-high-transportation-costs-idb-study-shows-5979.html

Inter-American Development Bank. (2010). Trade and productivity: A route to reallocation with a high transport toll [Electronic version]. In Páges, C. (Ed.), *The age of productivity: Transforming economies from the bottom up* (pp. 97-122). New York, NY: Palgrave Macmillan. Retrieved October 9, 2010, from http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=2270519

Inter-American Development Bank (IDB), World Bank, and United Nations Economic Commission for Latin America and the Caribbean (UNECLAC). (2010). Bridging integration gaps: Scenarios and policy recommendations to promote physical infrastructure and reduce intra-regional trade costs (Policy Discussion Brief prepared in response to Third Meeting of the Finance Ministers of the Americas and the Caribbean - Lima, Peru, May 28, 2010). Retrieved from http://issuu.com/idb_publications/docs/policybri_en_7766?e=0

International Road Assessment Programme (IRAP). (2011). What goes around... How Brazil has rediscovered the merits of the roundabout. [Case study on IRAP website.] Retrieved October 21, 2011 from http://www.irap.net/library/cat_view/5-case-studies.html?start=5

Jameson, K.P. (2010). The indigenous movement in Ecuador: The struggle for a plurinational state. *Latin American Perspectives*, 176(38), 63-73.

Johnston, R.J., Gregory, D., Pratt, G. & Watts, M. (Eds.). (2000). *The dictionary of human geography* (4th ed.). Oxford, UK: Blackwell Publishing.

Kali, R., Méndez, F., & Reyes, J. (2007). Trade structure and economic growth. *Journal of International Trade & Economic Development*, *16*(2), 245-270. As cited in Reyes, J., Schiavo, S., & Fagiolo, G. (2010). Using complex networks analysis to assess the evolution of international economic integration: The cases of East Asia and Latin America. *The Journal of International Trade & Economic Development, 19*(2), 215-239.

Kashiha, M., Tiller, K.C., & Thill, J.-C. (2013, September 27). *Brazilian ports: a trade context baseline analysis for containerized U.S.-bound exports*. Presentation at RSAmericas/SOCHER Conference 2013 (Arica, Chile).

Kay, C. (2002). Why East Asia overtook Latin America: agrarian reform, industrialization and development. *Third World Quarterly, 23*(6), 1073-1102. As cited in Gwynne, R.N., & Kay, C. (2004). Latin America transformed: *Globalization and modernity* (2nd ed.). London, UK: Edward Arnold Publishers.

Kockelman, K.M., & Ruiz-Juri, N. (2003, November). *Extending a random-utility-based multiregional input-output model for trade: incorporating land-use constraints, domestic demand and network congestion*. Paper presented at the 50th Annual North American Meetings of the Regional Science Association International, Philadelphia, PA. Abstract retrieved July 1, 2011 from: http://209.200.67.135/narsc/2003 Abstracts.pdf

Lane, M. (1998). Customs and corruption. Working paper, Transparency International, Berlin. Ferreira, C., Engelschalk, M., & Mayville, W. (2006). The challenge of combating corruption in customs administrations. In Campos, J.E., & Pradhan, S. (Eds.), *The many faces of corruption: Tracking vulnerabilities at the sector level* (pp. 367-386). Washington, DC: World Bank.

Latin American, Caribbean exports to jump 21 per cent this year, finds UN study. (2010, September 2). *UN News Centre*, Retrieved October 4, 2010, from http://www.un.org/apps/news/story.asp?NewsID=35815&Cr=latin+america&Cr1=

Lawrence, R.Z., Blanke, J., Hanouz, M.D., & Moavenzadeh, J. (2008). *The Global Enabling Trade Report 2008*. Geneva, Switzerland: World Economic Forum. Retrieved September 16, 2011 from: http://books.google.com/books?id=PRq2n50PqSYC&pg=PT5&lpg=PT5&dq=global+enabling+trade+report+2008&source=bl&ots=j1p9OM_0ZO&sig=2geVslQw6po9DEcLyBHtv9O2glM&hl=en&ei=cUV5TvTeFLCEsAKhzMHGAw&sa=X&oi=book_result&ct=result&resnum=10&ved=0CFkQ6AEwCTgK#v=onepage&q&f=false

Lawrence, R.Z., Hanouz, M.D., and Moavenzadeh, J. (2009). *The Global Enabling Trade Report 2009*. World Economic Forum. Geneva, Switzerland: World Economic Forum. Retrieved September 16, 2011 from Google Books: http://books.google.com/books?id=MuMSrtqkApUC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

LeSage, J. P., & Pace, R.K. (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science, 48*(5), 941–967.

Ledent, J. (1985). The doubly constrained model of spatial interaction: a more general formulation. *Environment and Planning A*, 17, 253-262.

Lee, N. (2009). Now more than ever: The case for a new integration strategy for the Americas. (Paper produced in support of 2009 Summit of the Americas as joint project of Canadian Foundation for the Americas [FOCAL], Inter-American Dialogue [IAD], and Center for Global Development.) Retrieved October 8, 2010, from http://www.focal.ca/pdf/summitV_Lee_CGD_case%20for%20new%20integration%20strategy_January%2013%202009_Washington.pdf

Limão, N., & Venables, A.J. (2001). Infrastructure, geographical disadvantage, transport costs, and trade. *World Bank Economic Review, 15*(3), 451-479.

Lopez-Claros, A., Altinger, L., Blanke, J. , Drzeniek, M., & Mia, I. (2006). The Global Competitiveness Index: Identifying the Key Elements of Sustainable Growth. *The Global Competitiveness Report 2006-2007*. Hampshire: Palgrave Macmillan. 3–50, as cited in Mia, I., Estrada, J., & Geiger, T. (2007). Benchmarking national attractiveness for private investment in Latin American infrastructure [Electronic version]. Geneva, Switzerland: World Economic Forum. Retrieved September 19, 2011 from: http://www3.weforum.org/docs/WEF_GCR_BenchmarkingNationalAttractiveness_PrivateInvestmentLA_2007.pdf

Lu, & Thill, J.-C. (2003). Assessing the cluster correspondence between paired point locations. *Geographical Analysis, 35*(4), 290-309.

McLinden, G. (2005). Integrity in customs. In Customs modernization handbook, De Wulf, L., & Sokol, J.B. (Eds.). Washington, DC: World Bank. As cited in Ferreira, C., Engelschalk, M., & Mayville, W. (2006). The challenge of combating corruption in customs administrations. In Campos, J.E., & Pradhan, S. (Eds.), *The many faces of corruption: Tracking vulnerabilities at the sector level* (pp. 367-386). Washington, DC: World Bank.

Manzetti, L., & Blake, C.H. (1996). Market reforms and corruption in Latin America: new means for old ways. *Review of International Political Economy, 3*(4), 662-697.

Mesquita Moreira, M., Volpe, C., & Blyde, J.S. (2008). Unclogging the arteries: The impact of transport costs on Latin American and Caribbean trade. *Globalization, Competitiveness, & Governability, 2*(3), 126-141, Retrieved from: http://gcg.universia.net/pdfs_revistas/articulo_107_1227718940236.pdf

Mesquita Moreira, M., Blyde, J., Volpe, C., & Molina, D. (2013). Too far to export: domestic transportation costs and regional export disparities in Latin America and the Caribbean (Special report on integration and trade). Washington, D.C.: Inter-American Development Bank. Retrieved from: http://www10.iadb.org/intal/intalcdi/PE/2013/13227en.pdf

Mia, I., Estrada, J., & Geiger, T. (2007). Benchmarking national attractiveness for private investment in Latin American infrastructure. Geneva, Switzerland: World Economic Forum. Retrieved September 19, 2011 from: http://www3.weforum.org/docs/WEF_GCR_BenchmarkingNationalAttractiveness_PrivateInvestmentLA_2007.pdf

Microsoft Corporation. (2005, 2008). Microsoft SQL Server (Versions 2005;2008). [Computer software].

Mikuriya, K. (2011, June). Customs and business, equally reaping rewards from trade facilitation. *WCO (World Customs Organization) Magazine, 65*, 10-15. Retrieved July 1, 2011from: http://www.wcoomd.org/home_press_wco_news.htm.

National Geospatial Intelligence Agency (NGA) (2006). Seaports data file. Available at: nga.mil.

Navajas, E., Mongelli, I., Kritzinger, S., Helmreich, S., Morgenstern, C., & Dennisen, T. (2010). Deliverable 3 of LogMan project: Carbon Footprint (Preliminary analysis funded by the European Commission 7th RTD Programme, Basle, Switzerland.) Retrieved July 2, 2011 from: http://logman-footprint.eu/fileadmin/Documents/Deliverable_3_FINAL_VERSION.pdf

Notteboom, T.E. (2001). *Spatial and functional integration of container port systems and* *hinterland networks in Europe* (Economic Research Centre/European Conference of Ministers of Transport – Report of 113th Round Table on Transport Economics: Land Access to Sea Ports). Paris, France: OECD Publications Service.

O'Neill, J., Wilson, D., Purushothaman, R., & Stupnytska, A. (2005, December 1). How solid are the BRICs? (Global economics paper No. 134) Retrieved December 20, 2007, from http://www2.goldmansachs.com/ideas/brics/how-solid-doc.pdf. As cited in Wu, Y.-C., & Lin, C.-W. (2008). National port competitiveness: implications for India. *Management Decision, 46*(10), 1482-1507.

Openshaw, S. (1984). The modifiable areal unit problem, CATMOG 38. Norwich: GeoAbstracts. As cited in Fotheringham, A.S., Brunsdon, C., & Charlton, M.E. (2000). Quantitative geography. London, UK: Sage Publications.

Ord, J.K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis, 27*(4): 286-306.

Organisation for Economic Co-Operation and Development (OECD). (2009). Latin American Economic Outlook 2010. Paris, France: OECD Publishing.

Organisation for Economic Co-Operation and Development (OECD) and European Conference on Ministers of Transport (ECMT). (2007). Managing urban traffic congestion. Paris, France: OECD Publications.

PIERS/UBM Global/JOC Group, Inc. (2007). PIERS TI U.S. Waterborne Imports dataset, July 1, 2006 - June 30, 2007 (Latin American exports portion) [Data file]. Newark, NJ: UBM Global Trade.

Plane, D.A. (1984). Migration space: Doubly constrained gravity model mapping of relative interstate separation [Electronic version]. *Annals of the Association of American Geographers, 74*(2), 244-256.

Reyes, J., Schiavo, S., & Fagiolo, G. (2010). Using complex networks analysis to assess the evolution of international economic integration: The cases of East Asia and Latin America. *The Journal of International Trade & Economic Development, 19*(2), 215-239.

Roy, J.R., & Thill, J.-C. (2004). Spatial interaction modelling. *Papers in Regional Science, 83*, 339-361.

Schwartz, J., Guasch, J.L., Wilmsmeier, G., & Aiga, S. (2009). Transport and food prices in the LAC: Policy guidance for improving efficiency and reducing costs (August 2009, Number 2 - LCSSD Economics Unit, World Bank and the Inter-American Development Bank). Retrieved October 10, 2010 from: www.worldbank.org/lcssdpaperseries

SELA (Latin American and Caribbean Economic System). (2011). Physical infrastructure for integration in Latin America and the Caribbean (SP/DiN°09-11). Caracas, Venezuela: Permanent Secretariat of SELA. Retrieved October 21, 2011 from: http://www.sela.org/attach/258/EDOCS/SRed/2011/06/T023600004772-0-Physical_Infrastructure_for_Integration_in_LAC.pdf

Sokal, R.R., Oden, N.L., & Thomsen, B.A. (1998). Local spatial autocorrelation in a biological model. *Geographical Analysis, 30*(4): 331-356.

Stocker, W. (2011). Comparative study on preferential rules of origin. Brussels, Belgium: World Customs Organization. Retrieved July 1, 2011 from: http://www.wcoomd.org/origin/01_study/pro_study.html

Street, J.H., & James, D.D. (1982). Institutionalism, Structuralism and Dependency in Latin America. *Journal of Economic Issues, 16*, 673-698.

Tagashira, N., & Okabe, A. (2002). The modifiable areal unit problem in a regression model whose independent variable is a distance from a predetermined point. *Geographical Analysis, 34*(1), 1-20.

Tchoukanski, I. (2010). ET GeoWizards (Version 11.0 - for ArcGIS version 10.1 – Built December 8, 2013) [Computer software]. Pretoria, South Africa: ET Spatial Techniques.

Thachuk, K. (2005). Corruption and international security. *SAIS Review, 25*(1), 143-152.

Thill, J.-C. (2011). Is spatial really that special? A tale of spaces. In V. Popovich, C. Claramunt, T. Devogele, M. Schrenk, & K. Korolenko (Eds.), *Information fusion and geographic information systems: Towards the digital ocean (Lecture notes in geoinformation and cartography)* (pp. 3-11). Heidelberg: Springer.

Thill, J.-C., Tiller, K.C., & Kashiha, M. (2010, March 11). *The Brazilian situation: investigating spatial patterns of containerized maritime export trade with the U.S.* Presentation at the 2010 Annual Forum of the Transportation Research Forum (Arlington, VA).

Thill, J.-C., Tiller, K.C., & Tang, S. (2008, April 17). *A spatial database for U.S. waterborne imports*. Presentation at 2008 Annual Meeting of the Association of American Geographers (Boston, MA).

Tiller, K.C. & Thill, J.-C. (2011a, April 12). *Functional separation as domestic freight transportation cost in containerized U.S.-bound Latin American export trade*. Presentation at 2011 Annual Meeting of the Association of American Geographers (Seattle, WA).

Tiller, K.C. & Thill, J.-C. (2011b, November 12). *Spatial patterns and determinants of landside trade impedance in containerized Latin American exports*. Paper presented at the 2011 Annual Conference of the North American Regional Science Council (Miami, FL).

Tobler, W.R. (1981). A model of geographic movement. *Geographical Analysis, 13*(1), 1-20.

Tobler, W.R. (1983). An alternative formulation for spatial-interaction modeling. *Environment and Planning A*, *15*, 693-703.

Tobler, W.R., Mielke, H.W., & Detwyler, T.R. (1970). Geobotanical distance between New Zealand and neighboring islands. *BioScience, 20*(9), 537-542.

Ullman, E.L. (1980a). Geography as spatial interaction. In Boyce, R.R. (Ed.), *Geography as spatial interaction* (pp. 13-27). Seattle, WA: University of Washington Press.

Ullman, E.L. (1980b). Human geography and area research. In Boyce, R.R. (Ed.), *Geography as spatial interaction* (pp. 28-40). Seattle, WA: University of Washington Press.

United Nations Economic Commission for Latin America and the Caribbean (UNECLAC). (2010). Series 102/103: Exports to the United States, Latin America, European Union and Asia / Imports from the United States, Latin America, European Union and Asia. Retrieved October 4, 2010 from http://www.eclac.org/ cgibin/getprod.asp?xml=/comercio/noticias/paginas/9/33539/P33539.xml&xsl=/ comercio/tpl-i/p18f.xsl&base=/comercio/tpl-i/top-bottom.xsl

United Nations Conference on Trade and Development (UNCTAD) Transport Section Trade Logistics Branch (2006). Transshipment Connectivity Index. As cited in Lawrence, R.Z., Blanke, J., Hanouz, M.D., & Moavenzadeh, J. (2008). *The Global Enabling Trade Report 2008*. Geneva, Switzerland: World Economic Forum.

United Nations Conference on Trade and Development (UNCTAD). (2008, July 28). *UNCTAD Handbook of Statistics 2008*. Retrieved from: http://unctad.org/en/pages/PublicationArchive.aspx?publicationid=2371

United Nations Conference on Trade and Development (UNCTAD). (2010, January 3). *UNCTAD Handbook of Statistics 2009*. Retrieved December 19, 2010 from http://www.unctad.org/Templates/webflyer.asp?docid=12930&intItemID =1397&lang=1&mode=downloads

U.S. Army Corps of Engineers. (2006). Schedule K: Classification of Foreign Ports by Geographic Trade Area and Country [Code list]. Alexandria, VA: U.S. Army Corps of Engineers.

United States International Trade Commission (U.S. ITC). (2007). HTS :2007-07-02 – Revision 2, Official Harmonized Tariff Schedule of the United States Annotated. [Data file] Available from: http://www.usitc.gov/tata/hts/bychapter/_0702.htm

U.S. Department of Transportation (DOT) Bureau of Transportation Statistics (BTS)/U.S. Census Bureau. (2007). Commodity Flow Survey. Retrieved November 26, 2011, from: http://www.bts.gov/publications/commodity_flow_survey/index.html

Van Bergeijk, P.A.G., & Brakman, S. (2010). Introduction: the comeback of the gravity model. In P.A.G. van Bergeijk & S. Brakman (Eds.), *The gravity model in international trade: advances and applications* (pp. 1-26). New York: Cambridge University Press.

Wei, Shang-Jin. (2000). Local Corruption and Global Capital Flows Comment and Discussion. Brookings Papers on Economic Activity, 2, 303-346.

Wilson, A.G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning, 3*: 1-32.

World Bank. (2009). *Logistics, transport and food prices in LAC: Policy guidance for improving efficiency and reducing costs.* Unpublished. World Bank, Washington, DC. As cited in Inter-American Development Bank. (2010). Trade and productivity: A route to reallocation with a high transport toll [Electronic version]. In Páges, C. (Ed.), *The age of productivity: Transforming economies from the bottom up* (pp. 97-122). New York, NY: Palgrave Macmillan. Retrieved October 9, 2010, from http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=2270519

World Bank. (2010, September). Latin America hurls towards recovery, but must improve productivity (Article summarizing speeches made by various World Bank staff at the Americas Conference 2010, Miami, FL). Retrieved October 2, 2010, from http://go.worldbank.org/5GHZWWT4B0

Wu, Y.-C., & Lin, C.-W. (2008). National port competitiveness: implications for India. [Electronic version]. *Management Decision, 46*(10), 1482-1507.

APPENDIX A: COMMODITY CATEGORIES (BASED UPON AGGREGATION OF TWO-DIGIT HS CODES)

(Source: United States International Trade Commission Harmonized Tariff Schedule, 2007.)

Category Number	Category Title	Category Contents	HS Codes Included:
1	Perishable products	-Live animals; -Meat and edible meat offal; -Fish and crustaceans, molluscs and other aquatic invertebrates; -Dairy produce; birds eggs; natural honey; edible products of animal origin, not elsewhere specified or included; -Products of animal origin, not elsewhere specified or included; -Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage; -Edible vegetables and certain roots and tubers; -Edible fruit and nuts; peel of citrus fruit or melons; -Coffee, tea, maté and spices; -Cereals; -Products of the milling industry; malt; starches; inulin; wheat gluten; -Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruits; industrial or medicinal plants; straw and fodder; -Lac; gums, resins and other vegetable saps and extracts; -Vegetable plaiting materials; vegetable products not elsewhere specified or included; -Animal or vegetable fats and oils and their cleavage products prepared edible fats; animal or vegetable waxes; -Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates; -Sugars and sugar confectionery;	HS 01-22; HS 24.

Appendix A (continued)

Category Number	Category Title	Category Contents	HS Codes Included:
1	Perishable products (continued)	-Cocoa and cocoa preparations; -Preparations of cereals, flour, starch or milk; bakers' wares; -Preparations of vegetables, fruit, nuts or other parts of plants; -Miscellaneous edible preparations; -Beverages, spirits and vinegar; -Tobacco and manufactured tobacco substitutes.	
2	Non-perishable time-sensitive products	-Articles of leather; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut); -Furskins and artificial fur; manufactures thereof; -Wool, fine or coarse animal hair; horsehair yarn and woven fabric; -Man-made filaments; -Knitted or crocheted fabrics; -Headgear and parts thereof; (-Umbrellas, sun umbrellas, walking sticks, seatsticks, whips, riding-crops and parts thereof;) (-Prepared feathers and down and articles made of feathers or of down; artificial flowers; articles of human hair;).	HS 42-43; HS 51; HS 54; HS 60; HS 65-67.
3	Heavy commodities/ Minimally processed natural resources	-Salt; sulfur; earths and stone; plastering materials, lime and cement; -Ores, slag and ash; -Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes; -Wood and articles of wood; wood charcoal; -Cork and articles of cork;	HS 25-27; HS 44-46; HS 68-70; HS 72-75; HS 77-79; HS 83; HS 86-87;

Appendix A (continued)

Category Number	Category Title	Category Contents	HS Codes Included:
3	Heavy commodities/ Minimally processed natural resources (continued)	-Manufactures of straw, of esparto or of other plaiting materials; basketware and wickerwork; -Articles of stone, plaster, cement, asbestos, mica or similar materials; -Ceramic products; -Glass and glassware; -Iron and steel; -Articles of iron or steel; -Copper and articles thereof; -Nickel and articles thereof; -Lead and articles thereof; -Zinc and articles thereof; -Miscellaneous articles of base metal; -Railway or tramway locomotives, rolling-stock and parts thereof; railway or tramway track fixtures and fittings and parts thereof; mechanical (including electro-mechanical) traffic signalling equipment of all kinds; -Vehicles other than railway or tramway rolling stock, and parts and accessories thereof; -Arms and ammunition; parts and accessories thereof.	HS 93.
4	Manufactured products	-Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes; -Organic chemicals; -Fertilizers; -Essential oils and resinoids; perfumery, cosmetic or toilet preparations; -Soap, organic surface-active agents, washing preparations, lubricating	HS 28-29; HS 31; HS 33-40; HS 47-49; HS 71.

Appendix A (continued)

Category Number	Category Title	Category Contents	HS Codes Included:
4	Manufactured products (continued)	preparations, artificial waxes, prepared waxes, polishing or scouring preparations, candles and similar articles, modeling pastes, "dental waxes" and dental preparations with a basis of plaster; -Albuminoidal substances; modified starches; glues; enzymes; -Explosives; pyrotechnic products; matches; pyrophoric alloys; certain combustible preparations; -Photographic or cinematographic goods; -Miscellaneous chemical products; -Plastics and articles thereof; -Rubber and articles thereof; -Pulp of wood or of other fibrous cellulosic material; waste and scrap of paper or paperboard; -Paper and paperboard; articles of paper pulp, of paper or of paperboard; -Printed books, newspapers, pictures and other products of the printing industry; manuscripts, typescripts and plans; -Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal and articles thereof; imitation jewelry; coin.	
5	High-value manufactured products	-Pharmaceutical products; -Tanning or dyeing extracts; dyes, pigments, paints, varnishes, putty and mastics; -Raw hides and skins (other than furskins) and leather; -Residues and waste from the food industries; prepared animal feed; (-Silk;) -Cotton;	HS 23; HS 30; HS 32; HS 41; HS 50; HS 52-53; HS 55-59;

Appendix A (continued)

Category Number	Category Title	Category Contents	HS Codes Included:
5	High-value manufactured products (continued)	-Other vegetable textile fibers; paper yarn and woven fabric of paper yarn; -Man-made staple fibers; -Wadding, felt and nonwovens; special yarns, twine, cordage, ropes and cables and articles thereof; -Carpets and other textile floor coverings; -Special woven fabrics; tufted textile fabrics; lace, tapestries; trimmings; embroidery; -Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use; - Articles of apparel and clothing accessories, knitted or crocheted*; - Articles of apparel and clothing accessories, not knitted or crocheted*; - Other made up textile articles; sets; worn clothing and worn textile articles; rags; -Footwear, gaiters and the like; parts of such articles; -Aluminum and articles thereof; -Tin and articles thereof; -Other base metals; cermets; articles thereof; -Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal; -Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof; -Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles; -Aircraft, spacecraft, and parts thereof; -Ships, boats and floating structures;	HS 61-64; HS 76; HS 80-82; HS 84-85; HS 88--92; HS 94-97.

Appendix A (continued)

Category Number	Category Title	Category Contents	HS Codes Included:
5	High-value manufactured products (continued)	-Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof; -Clocks and watches and parts thereof; -Musical instruments; parts and accessories of such articles; -Furniture; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings; lamps and lighting fittings, not elsewhere specified or included; illuminated sign illuminated nameplates and the like; prefabricated buildings; -Toys, games and sports requisites; parts and accessories thereof; -Miscellaneous manufactured articles; (-Works of art, collectors' pieces and antiques).	
6	Miscellaneous/ unable to be determined	(-Special classification provisions); (-Temporary legislation; temporary modifications proclaimed pursuant to trade agreements legislation; additional import restrictions proclaimed pursuant to section 22 of the Agricultural Adjustment Act, as amended); -Unknown commodity/commodity type unable to be determined; -Multiple-commodity/mixed freight†.	HS 98-99 and miscellaneous HS codes/ unknown.

Note: ( ) – Parentheses indicate a commodity type which is not present in the data as a single-commodity shipment in the dataset; upon splitting multiple-commodity/mixed-freight records into single commodity records, these commodity types may be identified as present within the dataset. If present, their values and proper classification by commodity category will be based upon the commodity type's average value to weight ratio among shipments. Currently these commodity types have at been assigned to commodity categories in which they could be reasonably expected to be included. If the presence of shipment records specific to these commodity types are identified from unknown-commodity or multiple-commodity shipment records, these commodity types will be valued using the hard data and assigned accordingly to the proper commodity category.

† - Multiple-commodity/mixed-freight shipment records will be split into single-commodity records; however, for the purposes of commodity valuation, commodities within these mixed-content records cannot be used for valuation/accurately split from the original shipment record in terms of value.

APPENDIX B: DISTRIBUTION OF A SAMPLE OF SHIPMENTS IN THE DATASET BY COMMODITY TYPE (SHARE OF TOTAL TEU, BASED UPON EARLY STAGE/PRELIMINARY DATASET EXPLORATION)
(Source: PIERS/UBM Global, 2007; U.S. ITC, 2006.)

TEU	Share of TEU Total	2-digit HS Code (Chapter)	HS Chapter Description
8.00	0.0%	01	Live animals
892.64	0.8%	02	Meat and edible meat offal
1,402.07	1.3%	03	Fish and crustaceans, molluscs and other aquatic invertebrates
278.85	0.3%	04	Dairy produce; birds eggs; natural honey; edible products of animal origin, not elsewhere specified or included
22.00	0.0%	05	Products of animal origin, not elsewhere specified or included
277.88	0.3%	06	Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage
2,476.67	2.3%	07	Edible vegetables and certain roots and tubers
22,190.16	20.7%	08	Edible fruit and nuts; peel of citrus fruit or melons
3,330.55	3.1%	09	Coffee, tea, maté and spices
49.00	0.0%	10	Cereals
150.20	0.1%	11	Products of the milling industry; malt; starches; inulin; wheat gluten
249.60	0.2%	12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruits; industrial or medicinal plants; straw and fodder
12.00	0.0%	13	Lac; gums, resins and other vegetable saps and extracts
2.00	0.0%	14	Vegetable plaiting materials; vegetable products not elsewhere specified or included
94.00	0.1%	15	Animal or vegetable fats and oils and their cleavage products prepared edible fats; animal or vegetable waxes
738.70	0.7%	16	Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates
1,113.98	1.0%	17	Sugars and sugar confectionery

Appendix B (continued)

TEU	Share of TEU Total	2-digit HS Code (Chapter)	HS Chapter Description
369.90	0.3%	18	Cocoa and cocoa preparations
369.02	0.3%	19	Preparations of cereals, flour, starch or milk; bakers' wares
1,752.92	1.6%	20	Preparations of vegetables, fruit, nuts or other parts of plants
721.42	0.7%	21	Miscellaneous edible preparations
2,544.30	2.4%	22	Beverages, spirits and vinegar
19.06	0.0%	23	Residues and waste from the food industries; prepared animal feed
365.35	0.3%	24	Tobacco and manufactured tobacco substitutes
1,938.19	1.8%	25	Salt; sulfur; earths and stone; plastering materials, lime and cement
122.82	0.1%	26	Ores, slag and ash
173.71	0.2%	27	Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes
569.34	0.5%	28	Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes
569.61	0.5%	29	Organic chemicals
82.30	0.1%	30	Pharmaceutical products
2.15	0.0%	31	Fertilizers
98.27	0.1%	32	Tanning or dyeing extracts; dyes, pigments, paints, varnishes, putty and mastics
235.59	0.2%	33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations
98.80	0.1%	34	Soap, organic surface-active agents, washing preparations, lubricating preparations, artificial waxes, prepared waxes, polishing or scouring preparations, candles and similar articles, modeling pastes, "dental waxes" and dental preparations with a basis of plaster
314.73	0.3%	35	Albuminoidal substances; modified starches; glues; enzymes
6.00	0.0%	36	Explosives; pyrotechnic products; matches; pyrophoric alloys; certain combustible preparations

Appendix B (continued)

TEU	**Share of TEU Total**	**2-digit HS Code (Chapter)**	**HS Chapter Description**
65.24	0.1%	37	Photographic or cinematographic goods
251.83	0.2%	38	Miscellaneous chemical products
2,077.64	1.9%	39	Plastics and articles thereof
1,587.82	1.5%	40	Rubber and articles thereof
310.91	0.3%	41	Raw hides and skins (other than furskins) and leather
552.16	0.5%	42	Articles of leather; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)
12.56	0.0%	43	Furskins and artificial fur; manufactures thereof
13,119.79	12.2%	44	Wood and articles of wood; wood charcoal
1.00	0.0%	45	Cork and articles of cork
8.00	0.0%	46	Manufactures of straw, of esparto or of other plaiting materials; basketware and wickerwork
6.06	0.0%	47	Pulp of wood or of other fibrous cellulosic material; waste and scrap of paper or paperboard
1,303.69	1.2%	48	Paper and paperboard; articles of paper pulp, of paper or of paperboard
248.73	0.2%	49	Printed books, newspapers, pictures and other products of the printing industry; manuscripts, typescripts and plans
46.42	0.0%	51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric
583.86	0.5%	52	Cotton
15.65	0.0%	53	Other vegetable textile fibers; paper yarn and woven fabric of paper yarn
186.42	0.2%	54	Man-made filaments
44.80	0.0%	55	Man-made staple fibers
117.62	0.1%	56	Wadding, felt and nonwovens; special yarns, twine, cordage, ropes and cables and articles thereof
30.04	0.0%	57	Carpets and other textile floor coverings
82.64	0.1%	58	Special woven fabrics; tufted textile fabrics; lace, tapestries; trimmings; embroidery

Appendix B (continued)

TEU	Share of TEU Total	2-digit HS Code (Chapter)	HS Chapter Description
30.16	0.0%	59	Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use
2.00	0.0%	60	Knitted or crocheted fabrics
1,559.25	1.5%	61	Articles of apparel and clothing accessories, knitted or crocheted
6,349.04	5.9%	62	Articles of apparel and clothing accessories, not knitted or crocheted
421.37	0.4%	63	Other made up textile articles; sets; worn clothing and worn textile articles; rags
1,118.38	1.0%	64	Footwear, gaiters and the like; parts of such articles
11.48	0.0%	65	Headgear and parts thereof
2,809.06	2.6%	68	Articles of stone, plaster, cement, asbestos, mica or similar materials
4,440.19	4.1%	69	Ceramic products
703.17	0.7%	70	Glass and glassware
83.61	0.1%	71	Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal and articles thereof; imitation jewelry; coin
996.22	0.9%	72	Iron and steel
1,838.25	1.7%	73	Articles of iron or steel
303.79	0.3%	74	Copper and articles thereof
3.00	0.0%	75	Nickel and articles thereof
821.22	0.8%	76	Aluminum and articles thereof
18.43	0.0%	78	Lead and articles thereof
20.95	0.0%	79	Zinc and articles thereof
25.34	0.0%	80	Tin and articles thereof
67.60	0.1%	81	Other base metals; cermets; articles thereof
100.60	0.1%	82	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal

Appendix B (continued)

TEU	**Share of TEU Total**	**2-digit HS Code (Chapter)**	**HS Chapter Description**
252.00	0.2%	83	Miscellaneous articles of base metal
2,820.42	2.6%	84	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof
2,097.75	2.0%	85	Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles
39.97	0.0%	86	Railway or tramway locomotives, rolling-stock and parts thereof; railway or tramway track fixtures and fittings and parts thereof; mechanical (including electro-mechanical) traffic signalling equipment of all kinds
1,973.60	1.8%	87	Vehicles other than railway or tramway rolling stock, and parts and accessories thereof
10.00	0.0%	88	Aircraft, spacecraft, and parts thereof
18.00	0.0%	89	Ships, boats and floating structures
621.32	0.6%	90	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof
0.90	0.0%	91	Clocks and watches and parts thereof
2.80	0.0%	92	Musical instruments; parts and accessories of such articles
8.92	0.0%	93	Arms and ammunition; parts and accessories thereof
2,436.53	2.3%	94	Furniture; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings; lamps and lighting fittings, not elsewhere specified or included; illuminated sign illuminated nameplates and the like; prefabricated buildings
137.15	0.1%	95	Toys, games and sports requisites; parts and accessories thereof
25.16	0.0%	96	Miscellaneous manufactured articles
8,148.85	7.6%	Multiple	Multiple HS codes (within individual shipment records)
2,525.01	2.4%	Unknown	Unknown HS code
107,136.15	**100.0%**	**Total TEU**	

Appendix B (continued)

Two-digit Harmonized Tariff Schedule Chapters not represented by shipment volume in the data: -50: Silk -66: Umbrellas, sun umbrellas, walking sticks, seatsticks, whips, riding-crops and parts thereof -67: Prepared feathers and down and articles made of feathers or of down; artificial flowers; articles of human hair -77: (This chapter number reserved by U.S. ITC for possible future use.) -97: Works of art, collectors' pieces and antiques -98: Special classification provisions -99: Temporary legislation; temporary modifications proclaimed pursuant to trade agreements legislation; additional import restrictions proclaimed pursuant to section 22 of the Agricultural Adjustment Act, as amended.

APPENDIX C: AVERAGE OF THE RATIOS FOR EACH SHIPMENT: VALUE PER KILOGRAM BY TWO-DIGIT HS CODE (COMMODITY TYPE) (FOR EARLY STAGE/PRELIMINARY ANALYSIS DATA SAMPLE.)
(In descending order by value ratio: high values, corresponding to the 75% percentile, are in bold.)
(Source: PIERS/UBM Global, 2007; U.S. ITC, 2006.)

Value per kilogram by 2-digit HS code (Average of the ratios for each shipment)	**2-digit HS Code (Chapter)**	**HS Chapter Description**
33.79	**88**	Aircraft, spacecraft, and parts thereof
33.79	**89**	Ships, boats and floating structures
26.78	**23**	Residues and waste from the food industries; prepared animal feed
20.91	**90**	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof
19.17	**59**	Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use
18.58	**76**	Aluminum and articles thereof
17.28	**62**	Articles of apparel and clothing accessories, not knitted or crocheted
16.61	**41**	Raw hides and skins (other than furskins) and leather
15.67	**58**	Special woven fabrics; tufted textile fabrics; lace, tapestries; trimmings; embroidery
14.70	**84**	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof
14.50	**80**	Tin and articles thereof
14.24	**64**	Footwear, gaiters and the like; parts of such articles
14.21	**56**	Wadding, felt and nonwovens; special yarns, twine, cordage, ropes and cables and articles thereof
13.75	**61**	Articles of apparel and clothing accessories, knitted or crocheted
13.71	**82**	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal

Appendix C (continued)

Value per kilogram by 2-digit HS code (Average of the ratios for each shipment)	**2-digit HS Code (Chapter)**	**HS Chapter Description**
12.76	**85**	Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles
12.32	**30**	Pharmaceutical products
11.62	**63**	Other made up textile articles; sets; worn clothing and worn textile articles; rags
11.51	**32**	Tanning or dyeing extracts; dyes, pigments, paints, varnishes, putty and mastics
11.47	**57**	Carpets and other textile floor coverings
11.32	**81**	Other base metals; cermets; articles thereof
11.26	**53**	Other vegetable textile fibers; paper yarn and woven fabric of paper yarn
10.67	**55**	Man-made staple fibers
10.62	**52**	Cotton
9.83	60	Knitted or crocheted fabrics
9.77	87	Vehicles other than railway or tramway rolling stock, and parts and accessories thereof
9.35	39	Plastics and articles thereof
9.32	79	Zinc and articles thereof
9.28	24	Tobacco and manufactured tobacco substitutes
9.11	29	Organic chemicals
8.20	31	Fertilizers
8.07	05	Products of animal origin, not elsewhere specified or included
7.94	83	Miscellaneous articles of base metal
7.76	42	Articles of leather; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)
7.02	78	Lead and articles thereof
6.84	95	Toys, games and sports requisites; parts and accessories thereof

Appendix C (continued)

Value per kilogram by 2-digit HS code (Average of the ratios for each shipment)	**2-digit HS Code (Chapter)**	**HS Chapter Description**
6.81	40	Rubber and articles thereof
6.42	48	Paper and paperboard; articles of paper pulp, of paper or of paperboard
6.14	33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations
5.97	91	Clocks and watches and parts thereof
5.97	92	Musical instruments; parts and accessories of such articles
5.80	93	Arms and ammunition; parts and accessories thereof
5.67	03	Fish and crustaceans, molluscs and other aquatic invertebrates
5.10	73	Articles of iron or steel
4.97	86	Railway or tramway locomotives, rolling-stock and parts thereof; railway or tramway track fixtures and fittings and parts thereof; mechanical (including electro-mechanical) traffic signalling equipment of all kinds
4.72	Multiple	Multiple HS codes (within individual shipment records)
4.57	51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric
4.36	70	Glass and glassware
4.10	47	Pulp of wood or of other fibrous cellulosic material; waste and scrap of paper or paperboard
3.98	94	Furniture; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings; lamps and lighting fittings, not elsewhere specified or included; illuminated sign illuminated nameplates and the like; prefabricated buildings
3.69	74	Copper and articles thereof
3.64	34	Soap, organic surface-active agents, washing preparations, lubricating preparations, artificial waxes, prepared waxes, polishing or scouring preparations, candles and similar articles, modeling pastes, "dental waxes" and dental preparations with a basis of plaster
3.13	28	Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes

Appendix C (continued)

Value per kilogram by 2-digit HS code (Average of the ratios for each shipment)	**2-digit HS Code (Chapter)**	**HS Chapter Description**
2.95	54	Man-made filaments
2.87	75	Nickel and articles thereof
2.72	27	Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes
2.61	02	Meat and edible meat offal
2.56	35	Albuminoidal substances; modified starches; glues; enzymes
2.56	49	Printed books, newspapers, pictures and other products of the printing industry; manuscripts, typescripts and plans
2.51	71	Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal and articles thereof; imitation jewelry; coin
2.50	72	Iron and steel
2.44	12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruits; industrial or medicinal plants; straw and fodder
2.39	38	Miscellaneous chemical products
2.33	18	Cocoa and cocoa preparations
2.32	Unknown	Unknown HS code
2.31	22	Beverages, spirits and vinegar
2.10	11	Products of the milling industry; malt; starches; inulin; wheat gluten
2.10	46	Manufactures of straw, of esparto or of other plaiting materials; basketware and wickerwork
2.08	01	Live animals
2.07	25	Salt; sulfur; earths and stone; plastering materials, lime and cement
1.88	21	Miscellaneous edible preparations
1.82	15	Animal or vegetable fats and oils and their cleavage products prepared edible fats; animal or vegetable waxes

Appendix C (continued)

Value per kilogram by 2-digit HS code (Average of the ratios for each shipment)	**2-digit HS Code (Chapter)**	**HS Chapter Description**
1.67	09	Coffee, tea, maté and spices
1.66	13	Lac; gums, resins and other vegetable saps and extracts
1.63	14	Vegetable plaiting materials; vegetable products not elsewhere specified or included
1.61	04	Dairy produce; birds eggs; natural honey; edible products of animal origin, not elsewhere specified or included
1.57	45	Cork and articles of cork
1.52	20	Preparations of vegetables, fruit, nuts or other parts of plants
1.47	17	Sugars and sugar confectionery
1.41	44	Wood and articles of wood; wood charcoal
1.39	16	Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates
1.37	19	Preparations of cereals, flour, starch or milk; bakers' wares
1.28	37	Photographic or cinematographic goods
1.16	69	Ceramic products
1.10	07	Edible vegetables and certain roots and tubers
1.10	10	Cereals
0.85	68	Articles of stone, plaster, cement, asbestos, mica or similar materials
0.83	36	Explosives; pyrotechnic products; matches; pyrophoric alloys; certain combustible preparations
0.79	08	Edible fruit and nuts; peel of citrus fruit or melons
0.78	65	Headgear and parts thereof
0.71	06	Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage
0.70	96	Miscellaneous manufactured articles
0.67	26	Ores, slag and ash
0.53	43	Furskins and artificial fur; manufactures thereof

Appendix C (continued)

Two-digit Harmonized Tariff Schedule Chapters not represented by shipment volume in the data:
-50: Silk
-66: Umbrellas, sun umbrellas, walking sticks, seatsticks, whips, riding-crops and parts thereof
-67: Prepared feathers and down and articles made of feathers or of down; artificial flowers; articles of human hair
-77: (This chapter number reserved by U.S. ITC for possible future use.)
-97: Works of art, collectors' pieces and antiques
-98: Special classification provisions
-99: Temporary legislation; temporary modifications proclaimed pursuant to trade agreements legislation; additional import restrictions proclaimed pursuant to section 22 of the Agricultural Adjustment Act, as amended.

APPENDIX D: INFORMATION AVAILABLE IN PIERS DATASET

- Physical location of origin, including:
    - Verified country of origin;
    - Verified locality within the country of origin;
- Port of export information: port name and port code (U.S. Army Corps of Engineers Schedule K code) where shipment is placed on an oceangoing vessel for eventual import to the U.S.;
- Volume of shipment in TEUs;

  Harmonized tariff code (6-digit) (later generalized to commodity category, which in a small percentage of cases is coded as miscellaneous/unknown);

Reliable attributes associated with these shipments, used only indirectly (or for database management) in this dissertation, include:

- Bill of lading number, as assigned by U.S. Customs and Border Protection – a unique identifier of each shipment (the first three characters of which tend to identify the carrier);
- Date of shipment arrival in the U.S. port of import;
- An indication (in the field related to volume) of whether the shipment is containerized or non-containerized;

Other reliable information associated with these shipments, but not used in this dissertation, include:

- North American (usually U.S./Puerto Rican, but sometimes Canadian) port of import information: port name and port code (U.S. Census Bureau code);
- U.S. destination city (physical destination);

Other information that has not been verified, for which reliability level within the original dataset is unknown:

- Various shipment movement/location point information: "Precarrier" location where carrier takes control of shipment, may or may not be present or reliable;
- Inbond code: Port code for cargo moving under Customs control where duty has not yet been paid. This indicates a payment for bypassing Customs at the landing port in order to clear through Customs at another port. Inbond codes may be used, for example, to bypass a congested port via vessel, truck, or rail. Foreign inbond codes are also possible, when goods from one nation are transferred from port to port in the U.S. for routing to another nation - no duty is paid to the U.S. because the goods never "enter" the U.S.
- Foreign destination (if shipment is being transited through the U.S.);
- Carrier information: Carrier name and SCAC code for at least one of the carriers involved in the transport of the shipment;
- Vessel information: Vessel name, Lloyd's code, and voyage number, for at least one of the vessels involved in the transport of the shipment;
- Container information: Container number for one or more containers which comprise the shipment in full or in part (the first three characters of which identify the owner of the container);

- Various commodity information, which may or may not be present, including: Long and short text descriptions of shipment composition; various units of quantity and weight; various markings which may or may not have been made on the shipment; data vendor-added commodity coding; data vendor-estimated value;
- Various shipment information, which may or may not be present: Shipper name and address information, which may be for physical production site, legal/corporate headquarters, or some other satellite location; consignee name and address, with similar stipulations; "also notified" parties' names and addresses, with similar stipulations.

APPENDIX E: COUNTRY EXPORT TRADE STRUCTURE BY PRODUCT GROUP (NOTES)

(Source: UNCTAD Handbook of Statistics 2008 – Section 3.1. Country trade structure by product group – Exports.)

As provided in UNCTAD Handbook of Statistics 2008 for Country Export Trade Structure by Product Group (Table 2) – by main SITC Revision 3 product group:

Product groups	**SITC Codes**
All food items	0 + 1 + 22 + 4
Agricultural raw materials	2 - (22 + 27 + 28)
Ores, metals, precious stones and nonmonetary gold	27 + 28 + 68 + 667 + 971
Fuels	3
Manufactured goods	5 + 6 + 7 + 8 - (667 + 68)
Chemical products	5
Machinery and transport equipment	7
Other manufactured goods	6 + 8 - (667 + 68)

APPENDIX F: REVERSE SPATIAL INTERACTION REGRESSION ANALYSIS RESULTS (MODEL SUMMARIES)

All Shipment Flows Segment

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.992	.985	.984	0.4174	.985	1,992.380	457	13,958	0.000	2.016

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	158,662.226	457	347.182	1,992.380	.000
	Residual	2,432.251	13,958	.174		
	Total	161,094.477	14,415			

Commodity Category 1 Segment: Perishable Products

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.992	.984	.984	0.3160	.984	1,737.257	314	8,741	0.000	1.750

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	54,469.545	314	173.470	1,737.257	.000
	Residual	872.813	8,741	.100		
	Total	55,342.358	9,055			

Commodity Category 2 Segment: Non-Perishable Time-Sensitive Products

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.995	.990	.989	0.1530	.990	1,018.271	56	599	0.000	.739

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1,334.834	56	23.836	1,018.271	.000
	Residual	14.022	599	.023		
	Total	1,348.855	655			

Commodity Category 3 Segment: Heavy Commodities/ Minimally Processed Natural Resources

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.993	.987	.986	0.3766	.987	1,932.324	253	6,659	0.000	1.848

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	69,350.843	253	274.114	1,932.324	.000
	Residual	944.627	6,659	.142		
	Total	70,295.470	6,912			

Commodity Category 4 Segment: Manufactured Products

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.994	0.988	.988	0.2073	.988	1,944.645	180	4,227	0.000	1.122

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15,038.787	180	83.549	1,944.645	.000
	Residual	181.607	4,227	.043		
	Total	15,220.394	4,407			

Commodity Category 5 Segment: High-Value Manufactured Products

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.983	0.966	.965	0.2925	.966	761.912	227	6,044	0.000	2.042

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14,794.675	227	65.175	761.912	.000
	Residual	517.011	6,044	.086		
	Total	15,311.686	6,271			

Commodity Category 6 Segment: Miscellaneous/Unable to be Determined Commodities

Model Summary:

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
				R Square Change	F Change	df1	df2	Sig. F Change	
.981	0.962	.960	0.2164	.962	568.883	172	3,887	0.000	1.619

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4,583.433	172	26.648	568.883	.000
	Residual	182.077	3,887	.047		
	Total	4,765.510	4,059			

APPENDIX G: RELATED PRESENTATIONS

Kashiha, M., Tiller, K.C., & Thill, J.-C. (2013, September 27). *Brazilian ports: a trade context baseline analysis for containerized U.S.-bound exports.* Presentation at RSAmericas/SOCHER Conference 2013 (Arica, Chile).

Thill, J.-C., Tiller, K.C., & Kashiha, M. (2010, March 11). *The Brazilian situation: investigating spatial patterns of containerized maritime export trade with the U.S.* Presentation at the 2010 Annual Forum of the Transportation Research Forum (Arlington, VA).

Thill, J.-C., Tiller, K.C., & Tang, S. (2008, April 17). *A spatial database for U.S. waterborne imports.* Presentation at 2008 Annual Meeting of the Association of American Geographers (Boston, MA).

Tiller, K.C. & Thill, J.-C. (2011a, April 12). *Functional separation as domestic freight transportation cost in containerized U.S.-bound Latin American export trade.* Presentation at 2011 Annual Meeting of the Association of American Geographers (Seattle, WA).

Tiller, K.C. & Thill, J.-C. (2011b, November 12). *Spatial patterns and determinants of landside trade impedance in containerized Latin American exports*. Paper presented at the 2011 Annual Conference of the North American Regional Science Council (Miami, FL).