The Use of Proxy Variables in Economic Gravity Models: A Cautionary Note

This study examines the validity of using proxy variables in gravity model analyses. Models describing eight different trade scenarios are developed. Four different models are estimated for the land transport of refrigerated pork from the five major exporting provinces in Canada to the 48 contiguous states of the continental United States. Four equivalent gravity models are estimated for intermodal ocean container shipments of lentils from Western Canada to destinations in 97 countries in Europe, Africa, South America, and Asia. For half the scenarios, freight rates are specified as the impedance factor, while the balance of the scenarios use distance as a proxy for freight rates. In addition, this study analyzes the implications of using the dollar value of exports instead of the volume of exports as the dependent variable in half of the gravity model estimations. The results suggest that using proxy variables in gravity model analysis produces acceptable results, but the margin of error is greater.

by Erica Vido and Barry E. Prentice

Economic gravity models are used to examine the role played by transportation costs in the volume of commodity flows and international trade. Although inspired by the laws of gravity, economic gravity models assume that the Law of One Price applies. Specifically, trade between two regions depends on the economic distance between them. Under the Law of One Price, markets are at equilibrium when all prices are separated by the cost of moving goods between them.

Geographic information data requirements are less burdensome for gravity models. This is a significant advantage for quantitative analysis. Traditional trade equilibrium models require commodity prices in different regions that represent the appropriate market level (wholesale or retail) and similar value-added quality. Economic gravity models require only an accurate measure of commodity flows and transportation costs between origins and destinations. Nevertheless, data on transport costs are often unavailable or difficult to obtain. Consequently, distance has been used as a proxy for transport costs in previous studies.

This study examines the validity of proxy variables in gravity model analysis, and models describing eight different trade scenarios are developed. Four different models are estimated for the land transport of refrigerated pork from the five major exporting provinces in Canada to the 48 contiguous states of the continental United States. Four equivalent gravity models are estimated for intermodal ocean container shipments of lentils from Western Canada to destinations in 97 countries in Europe, Africa, South America, and Asia. For half the scenarios, freight rates are specified as the impedance factor, while the balance of the scenarios use distance as a proxy for freight rates. In addition, this study analyzes the implications of

Journal of the Transportation Research Forum. Published in *Transportation Quarterly*,
Vol. 57, No. 1, Winter 2003 (123–137).

using the dollar value of exports instead of the volume of exports as the dependent variable in half of the gravity model estimations.

MICRO FOUNDATIONS AND ESTIMATION ISSUES OF GRAVITY MODELS

Gravity Models were first applied to international trade by Tinbergen (1962), Poyhonen (1963) and Linneman (1966), who proposed that the volume of trade is an increasing function of the national incomes of the trading partners, and a decreasing function of the distance between them. The empirical success of gravity models was hard to deny but they were criticized by economists because they seemed to lack theoretical foundations. These foundations were subsequently developed by, among others, Anderson (1979) and Bergstrand (1985), who derived gravity models from models of monopolistic competition, and Deardorff (1998), who demonstrated that the gravity model could be derived within Ricardian and Heckscher-Ohlin frameworks.

Prentice et al. (1998) and Urbina-Olano (1996) derive the gravity model for interregional trade using the Law of One Price as a theoretical base. In a world of one commodity and two regions where price differences exist in the absence of trade, this difference would give rise to potential trade flows from where price is low to where price is high. Assuming that exchange rates are fixed and that free trade is negotiated between the two regions, the Law of One Price dictates that trade flows would increase until the prices differed exactly by the transportation and logistics costs associated with transfer. The gravity model represents the derived demand for transportation from the low price market to the high price market.

Impedance Factor Specification

Linneman (1966) uses distance between trading countries as a proxy variable for the total natural trade impediments, in their widest sense. It embodies three elements: transport costs, transport time, and psychic distance, constituting together the obstacles to trade due to the existence of space. Over the years, many other economists have adhered to this method.

Several problems exist when using distance as a proxy for transport costs. First, no distance may be representative for countries that share long borders with goods moving across many entry points. Second, problems arise when allowing for variation in the cost of alternative means of transport. Land transport is more expensive than ocean transport, while air transport is more expensive than land transport. Distance measures cannot account for these cost differences. Third, improvements in transportation technology affects the cost of producing transport services and consumer demand for goods through the level of transport services provided (Bandyopadhay 1999). A fall in transportation costs because of a more efficient distribution sector will lower the final prices of goods and increase their quantities demanded.

Geraci and Prewo (1977) and Ferguson (1972) observe other reasons that make distance measures a poor proxy for transport costs. First, freight rates are influenced by factors other than distance, such as the value, weight, and bulk of the commodity being transported, as well as the mode of transportation used. Second, the use of distance assumes that freight rates are the same in either direction. In every trade lane there is a front haul and a back haul, with higher rates on front haul moves.

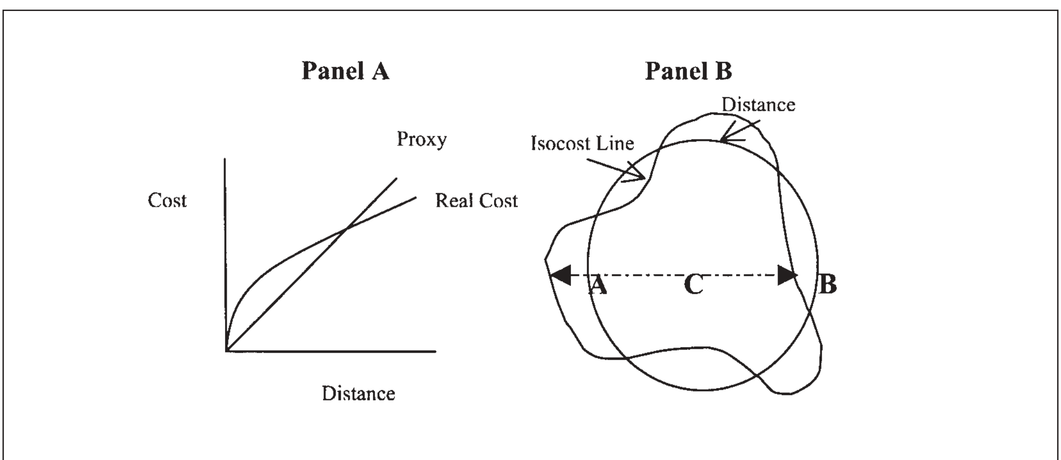
Figure 1 illustrates the relationship between transport cost and distance. In panel A, the proxy variable (distance) is assumed to exhibit a linear relationship. The real transport cost, on the other hand, does not lend itself to such simplicity. In this specific example, transportation costs rise rapidly for shorter distances and begin to taper off as the distance increases. For shorter routes, the proxy variable underestimates the effect of transfer costs on export flows and the opposite holds true for longer routes. Only on mid-range routes would the proxy variable accurately reflect the impact of transfer costs on export flows.

Panel B outlines this same relationship, but includes other factors such as the level of infrastructure and the direction of travel. The smooth circle represents distance and every point on the circle is equidistant from the center, point C. The warped circle represents an isocost line where the transportation costs from point C to any point on the isocost line are equal. Let's assume that a shipper at point C has goods to deliver to points A and B, both of which are equidistant from point C. The route to point B is through a mountainous region where the road twists and turns and travel speed is slow. On the other hand, the route to point A is through a plain where the road is generally flat and straight and travel speed is fast. For a given level of costs (the isocost line) the distances traveled toward points A and B are different. In this example, the route toward point B is more costly than the route toward point A. Again, a proxy variable like distance would not distinguish between these cost differences.

Distance measures fail to account for currency exchange rates and do not reflect technological advances that impact the shippers' decisions about mode choice. Mode choice, in turn, depends on the particular trade corridors used, the availability of equipment and infrastructure, freight rates, and the volumes and types of commodities being shipped. Finally, the estimated relationships between trade flows and static variables such as distance are not very helpful in predicting future trade levels or for policy analysis (Geraci and Prewo 1977).

Alcaly (1967) estimated a gravity model using travel cost as the impedance factor instead of distance. The inclusion of cost tends to make the equations somewhat more specific to each mode of transport and more like traditional demand equations. Depending on the research objectives, different specifications of the model may be appropriate. Aggregate models with distance as the impedance factor help to explain aggregate trade flows as a function of distance, but do not provide useful results for marketers and policy analysts. Disaggregated models with a transport cost factor as an explanatory vari-

Figure 1: The Relationship Between Transport Cost and Distance



able applies more to traditional demand analysis.

Prentice et al. (1998) and Urbina-Olano (1996) both use actual freight rates paid by shippers as the impedance factor variable. Their analyses concentrated on estimating the derived demand for transport or the freight rate elasticity for the Canadian chilled pork export market. Because they used actual transport costs in a disaggregated model, their model performed more like a traditional demand model and their results are more meaningful to marketers, transportation service providers, and policy analysts.

Alcaly (1967) points out that the relationship between distance and export volumes are likely to be different for different transportation modes. Air travel, for example, is the most expensive form of transport, reserved mainly for time sensitive and/or high valued commodities. Aggregating these commodities with low valued, bulk commodities, which generally face much lower freight rates, would tend to make the characteristics of each commodity category less distinct, and thus less useful for understanding specific trade relationships. Furthermore, aggregate equations reflect the responsiveness of the dominant mode of travel, leading Alcaly to suspect that what was being demonstrated may actually be the applicability of gravity models to the dominant mode of travel as opposed to travel by all modes and commodities.

Hillberry (2001) demonstrated that aggregating data across sectors contributes to estimation bias because the border effect induces the commodity composition of trade to change across national, provincial, and state borders. His analysis concentrated on shipments traveling by truck or rail, since only land-based trade can be reliably assigned to a specific province or state of entry. Hillberry's results suggest that borders affect the number of commodities that are traded in both level and composition. He concluded that gravity models that utilize data aggregated across either transport mode or commodity bundle are "quite likely to lead such models to overstate the welfare consequences of geographic trade frictions" (Hillberry 2001: 15).

Dependent Variable Specification

A gravity model requires some measure of the goods shipped as the dependent variable. The dependent variable in Bergstrand's (1985) gravity equation was specified as PX_{ij}, where PX_{ij} is the US dollar value of the flow of goods from country *i* to country *j*. If prices have been relatively stagnant throughout the time series, PX_{ij} is merely the volume of exports multiplied by some constant. No doubt it is more difficult to get quantity data of trade flows, but price data is a questionable substitute in a gravity model if prices have great flexibility. For instance, the 1970s saw a tripling of grain prices within a couple of years. A model using the value of exports as the dependent variable would treat this as a tripling of trade volumes, but the impact of transport costs on this trade is unlikely to have changed as much. As a proportion of the delivered price, transport costs almost certainly fell during this period. The fluctuations in market prices and trade volumes could as easily move in opposite directions than move in tandem. By adding a volatile price variable into the equation, as in PX_{ij}, the results become less clear.

Both the value of goods traded or the volume of goods traded could be used as the dependent variable. Volume data are likely to be more independent of transport costs, to the extent that carriers are known to try to extract higher freight rates when commodity prices rise, and vice versa. At the same time, the burden of transport costs is greater for low value commodities, especially if they are heavy or bulky, which is most often the case.

Aggregate trade could include all goods, ranging from high valued manufactured items, such as computers, and low valued raw materials, like wheat. The valuation of goods provides a common unit of measure that is more easily analyzed and compared. The problem that arises is the relationship between some aggregate value and its sensitivity to transportation costs. A unit train of wheat (100 rail cars) has about the same value as a truckload of computers, but the absolute costs of transport are about 30 times greater for the wheat.

Alcaly's (1967) gravity model was used to understand passenger travel between cities in the state of California. The dependent variable was specified as the volume of travel in passenger miles. Prentice et al. (1998) and Urbina-Olano (1996) use volume of pork exports measured in metric tonnes in their analyses. Using this specification helps the authors to answer the fundamental question; what volume of goods (passengers) are expected to move and where are they expected to move, based on the cost of moving them?

SPECIFICATION TESTING

The purpose of this analysis is to focus on the impact of using value and distance as proxy variables for quantities shipped and transport costs, respectively. The experiments were designed to minimize other sources of variation. A land transport model and a marine transport model are estimated separately. Test commodities were chosen that move exclusively by one mode of transport, and are very uniform in quality. Finally, to test the sensitivity of location, the models are estimated with individual origins of supply and aggregate sources of supply. The land transport of pork model was first estimated aggregately and then disaggregated by Canadian province. The marine transport of lentils model was first estimated aggregately and then disaggregated into Atlantic-based trade and Pacific-based trade.

EMPIRICAL MODELS

The model as derived in Prentice et al. (1998) is used in this analysis. It is the simplest form of the gravity model.

Marine Transport Gravity Models

In the marine models, only intermodal container transport is considered. Lentils that are shipped exclusively in containers were used as a test commodity.

(1) $\ln Q_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$

(2) $\ln Q_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$

(3) $\ln V_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$

(4) $\ln V_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$

Q_{ijt} is the quantity of lentil exports, in terms of container loads, from Canada to 97 different countries worldwide, obtained from the Canada Grains Council (2001). V_{ijt} is the value of lentil exports expressed in constant US dollars, obtained from Industry Canada. F_{ijt} are actual, confidential, ocean-leg container freight rates from the ports of Montreal or Vancouver to the destination markets. These data were obtained on a confidential basis from a large international freight forwarder that is active in this trade. D_{ijt} is the ocean distance (in nautical miles) from the ports of Montreal or Vancouver to the nearest port of entry, as found at www.distances.com. Y_{jt} is defined as importing countries' income (Gross Domestic Product—GDP) in constant US dollars as obtained from the World Bank (2000).

A dummy variable, B_1, is used to identify competing lentil production in importing countries. The dummy variable was set to unity if the importing country produced lentils domestically, and zero otherwise. This model uses two dummy variables to distinguish high-volume, relational trade linkages, from low-volume spot market trade flows. Discussions with exporters revealed that they enjoy long-term relationships with some markets, while other markets may only emerge at random when some local shortage appears. The former markets are consistent importers and import disproportionately large volumes. The "spot market" regions have relatively small volumes and do not import every year. B_2 identifies those markets that consistently import Canadian lentils every year, irrespective of actual volumes; and it is set to unity if countries import every year, and zero otherwise. B_3 distinguishes large importers from small importers; and it was set to unity if imports exceed 500 container loads per year, and zero otherwise. B_4 distinguishes rich countries from poor countries; it was set to unity if the importing country is a member of the Organization for Economic Cooperation and Development (OECD), and zero otherwise. The above dummy variable approach best explains the diverse platform of the global lentil trade.

There are 97 countries (cross-sections) and eight time periods (1991–1998). Two sub-models were estimated for Atlantic-based and Pacific-based trade, as well as the aggregate model.

Land Transport Gravity Models

In these models, only refrigerated, long-haul tractor-trailer transport trucks were considered. The test commodity, fresh and frozen pork, moves exclusively in these vehicles between Canada and the US.

(5) $\ln Q_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

(6) $\ln Q_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

(7) $\ln V_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

(8) $\ln V_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

Q_{ijt} is the quantity of pork exports, in terms of truckloads, from province *i* to state j in year t, as obtained from Industry Canada. V_{ijt} is the value of chilled/frozen pork exports expressed in constant US dollars, also obtained from Industry Canada. F_{ijt} is the truckload freight rate of transporting pork from province *i* to state *j* in year *t*. These freight rates were obtained on a confidential basis from a large refrigerated truckload carrier that is an active participant in Canada-US trade. D_{ijt} is the truck distance (in miles), obtained from Rand McNally and Company (2000), from the five major pork-exporting provinces in Canada to the 48 states in the continental US, measured from the city with the largest population. Y_{jt} is defined as the income (GDP) of importing states taken directly from the US Department of Commerce, *Survey of Current Business* (1996). There are 48x5 = 240 trade flows (cross sections) covering a five-year period (1992–1996).

An additional variable was included to identify the effects of competing supply sources (B_1 in Marine Transport Gravity Model). This variable is called the production specialization index, INX_{jt}, which is calculated as the ratio of hog inventory to human population in importing regions. The larger the value of INX_{jt} for a particular state, the more self sufficient in hog production it is. For an importing region, import quantity is expected to increase as income goes up or the specialization index declines. This data was obtained from the US Department of Agriculture, National Agricultural Statistics Service (1992).

The dummy variable (B_{jt}) is specified to distinguish high-volume, relational trade linkages, from low-volume spot market trade flows. The dummy variable is set to unity when a state's annual pork imports are more than 500 metric tonnes and zero otherwise.

RESULTS AND DISCUSSION

Marine Transport Gravity Models

The regression results for the marine transport gravity models are presented in Tables 1 and 2. Standard criteria were used in evaluating the performance of the different gravity models such as agreement of the variable signs with theoretical expectations, the magnitude of the coefficients, the statistical significance of the coefficients and the explanatory power of the entire relationship. Given the consistency of the data, we would expect a similar level of performance from each of the different gravity models.

Impedance Factor Elasticity. The impedance factors used in the estimate are ocean-leg container freight rates and distances to ports. A negative sign is expected for these variables since trade would tend to decrease as the cost of transport or the distance to market increases. All but one of the estimates of the freight rate variable has the expected sign; however, this variable is statistically significant only on the Atlantic trade routes. This indicates that there may be some inherent differences in the structure of ocean-leg container rates on Atlantic versus Pacific trade corridors.[1]

Since transportation costs make up a significant component of the final cost of the delivered product for agricultural commodities, we would expect the freight rate elasticity for lentils to be elastic (i.e., >-1). This means a 1% decline in freight rates on a given trade lane would induce exports to increase by more than 1%. The two statistically significant freight rate elasticity estimates (appearing in the Atlantic trade) are both greater than unity (-1.2 and -1.1), suggesting that lentils exports are quite sensitive to the cost of freight.

The distance elasticities are more robust. All have the expected sign and only one regression produced an insignificant estimate. Of the significant distance elasticities, all but one produced estimates that are greater than unity. The Aggregate Trade regression of Model 2 produced an estimate of less than unity. This is likely due to the Atlantic Trade whose distance elasticity produced an insignificant estimate.

Income Effect. Total disposable income reflects both the size of the human population and per capita income and is a natural candidate to represent potential market size. The income coefficients represent the percentage change in quantity of lentils exported in response to a 1% change in importers' income. Thus, a positive sign is expected, which occurs in 11 of the 12 regressions. Of the 12 regressions, five produced an insignificant estimate of income elasticity. The seven significant income elasticity estimates are all positive and less than unity. These results suggest that lentils exports are inelastic to changes in foreign incomes. This may be because lentils are a protein-rich, staple food source for many of the world's poor, and are time sensitive with a high demand near Easter. This could make lentil consumption more inelastic to income changes. A second problem may lie in the mix of very high-income European countries, with very low-income African and Asian countries. The OECD Member dummy was intended to capture this phenomenon but was statistically significant in only three of the regressions.

Dummy Variables. The Competitive dummy variable (B_1) was intended to capture the effects of competing domestic supply sources. It is statistically significant and positive in all but two of the regressions. The positive result may seem surprising at first, but upon further examination, the positive

Table 1: Estimation Results for the Marine Transport of Lentils (Volume as Dependent Variable)

Model 1: Volume = f(Freight Rates)

$\ln Q_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 1n INX_{jt} + \delta B_{jt} + e_{ijt}$

Elasticity	Aggregate Trade	Atlantic Trade	Pacific Trade
Intercept *(t-ratio)*	0.36928 *(0.1556)*	7.4544 *(2.703)**	7.5732 *(0.9395)*
Ln(Freight Rate) *(t-ratio)*	-0.15718 *(-0.6567)*	-1.2156 *(-4.078)**	-0.69965 *(-0.8212)*
Ln(Income) *(t-ratio)*	0.057064 *(1.523)*	0.11562 *(2.941)**	-0.097373 *(-0.8904)*
B1 (Competitive) *(t-ratio)*	0.7675 *(4.083)**	0.29566 *(1.559)*	1.6783 *(2.529)**
B2 (Consistent) *(t-ratio)*	4.9416 *(23.42)**	3.4192 *(14.39)**	6.2502 *(12.34)**
B3 (Large Importer) *(t-ratio)*	3.1328 *(13.12)**	3.6074 *(14.26)**	2.9061 *(4.575)**
B4 (OECD Member) *(t-ratio)*	-0.065657 *(-0.2654)*	-0.64908 *(-2.106)**	-.89595 *(-1.255)*
Observations	776	584	192
R-square	0.7629	0.7929	0.7842

Model 2: Volume = f(Distance)

$\ln Q_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$

Elasticity	Aggregate Trade	Atlantic Trade	Pacific Trade
Intercept (t-ratio)	6.937 (3.645)*	-0.99843 (-0.3495)	11.438 (3.524)*
Ln(Distance) *(t-ratio)*	-0.98434 *(-4.568)**	-0.30101 *(-0.8962)*	-1.0559 *(-3.657)**
Ln(Income) *(t-ratio)*	0.096662 *(3.016)**	0.19114 *(4.864)**	-0.07267 *(0.9885)*
B1 (Competitive) (t-ratio)	0.82989 (4.5)*	0.48943 (2.432)*	1.8747 (3.954)*
B2 (Consistent) *(t-ratio)*	4.5226 *(20.05)**	3.7376 *(15.08)**	4.8422 *(9.17)**
B3 (Large Importer) *(t-ratio)*	2.9797 *(12.86)**	3.4098 *(13.34)**	2.5385 *(4.83)**
B4 (OECD Member) *(t-ratio)*	-0.3004 *(-1.274)*	-0.25678 *(-0.8645)*	-1.194 *(-2.061)**
Observations	776	584	192
R-square	0.7733	0.7823	0.8409

**Denotes the coefficient estimate is significant at the 5% level*

Table 2: Estimation Results for the Marine Transport of Lentils (Value as Dependent Variable)

Model 3: Value = f(Freight Rates)

$$\ln V_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$$

Elasticity	Aggregate Trade	Atlantic Trade	Pacific Trade
Intercept *(t-ratio)*	0.5569 *(0.1513)*	6.0153 *(1.756)**	5.3546 *(0.254)*
Ln(Freight Rate) *(t-ratio)*	0.13656 *(0.3867)*	-1.0604 *(-3.202)**	-0.63697 *(-0.2688)*
Ln(Income) *(t-ratio)*	0.1477 *(2.218)**	0.33473 *(5.258)**	0.097294 *(0.3373)*
B1 (Competitive) *(t-ratio)*	1.3279 *(6.936)**	0.79994 *(4.13)**	1.1118 *(0.9239)*
B2 (Consistent) *(t-ratio)*	10.854 *(21.59)**	9.1509 *(16.71)**	15.798 *(13.68)**
B3 (Large Importer) *(t-ratio)*	2.4004 *(12.64)**	1.8188 *(6.309)**	1.9031 *(2.281)**
B4 (OECD Member) *(t-ratio)*	1.1274 *(3.295)**	0.41076 *(0.9386)*	-0.38022 *(-0.3211)*
Observations	776	584	192
R-square	0.5944	0.6909	0.6998

Model 4: Value = f(Distance)

$$\ln V_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \delta_1 B_{1jt} + \delta_2 B_{2jt} + \delta_3 B_{3jt} + \delta_4 B_{4jt} + e_{ijt}$$

Elasticity	Aggregate Trade	Atlantic Trade	Pacific Trade
Intercept *(t-ratio)*	17.317 *(6.96)**	8.2712 *(1.311)*	20.878 *(2.953)*
Ln(Distance) *(t-ratio)*	-1.9859 *(-8.335)**	-1.2384 *(-1.866)**	-2.0218 *(-4.153)**
Ln(Income) *(t-ratio)*	0.23173 *(4.214)**	0.35563 *(6.234)**	0.017702 *(0.07215)*
B1 (Competitive) *(t-ratio)*	1.3477 *(6.995)**	1.2459 *(4.694)**	2.9791 *(3.409)**
B2 (Consistent) *(t-ratio)*	9.7777 *(19.99)**	9.1082 *(16.1)**	12.061 *(9.562)**
B3 (Large Importer) (t-ratio)	2.4268 (11.8)*	2.0514 (6.998)*	1.3394 (1.916)*
B4 (OECD Member) (t-ratio)	-0.069818 (-0.1999)	0.20797 (0.5339)	-1.641 (-1.467)
Observations	776	584	192
R-square	0.6384	0.6692	0.7195

** Denotes the coefficient estimate is significant at the 5% level* elasticity suggests that lentil-producing countries have a higher propensity to consume lentils than countries with no local production. When local production does not satisfy domestic demand, lentils must be purchased from international markets.

The dummy variables Consistent (B_2) and Large Importer (B_3) were intended to distinguish the larger import markets from the smaller ones. As expected, all the estimates are positive and highly significant. This suggests that countries that import consistently and/or in large quantities have a higher propensity to import Canadian lentils, than do small and/or inconsistent importers.

Explanatory Power. Finally, an examination of the explanatory power (R-square) of the entire relationship reveals a similar level of performance of each model. However, the Model 1 and Model 2 regressions, where volume of exports is used as the dependent variable, have higher R-squares than the Model 3 and Model 4 regressions, where value of exports is specified as the dependent variable.

Land Transport Gravity Models

The regression results for the land transport gravity model are presented in Tables 3 and 4.

Impedance Factor Elasticity. The impedance factors used to estimate the land transport gravity models are truck-load freight rates and distances between major cities. The coefficients of the impedance factor variables all have the expected negative sign, their magnitude is reasonable and all are statistically significant.

The magnitude of the freight rate elasticities in Models 5 and 7 are significantly larger than those in Models 1 and 3. Since water transport is much cheaper than other modes of transport, it makes sense that the demand elasticities are lower for the ocean-borne lentil trade. The elasticities estimated for Models 5 and 7 refer to refrigerated truck movements of fresh pork between Canada and the United States. The freight rates are higher, the distances are shorter, and service is a much more important component of carrier choice.

The impedance factor elasticities for the land transport gravity models are more robust than that of the marine transport gravity models. This may be due to the more diverse platform of the global lentil trade as compared to the relatively homogeneous North American chilled pork market. The transportation data used in these analyses could contribute to some variations. Ocean-leg container freight rates are different in structure than truckload freight rates. Container rates are set by conferences, while truck-load freight rates are highly competitive and generally quoted on a per mile basis. Consequently, marine freight rates may not reflect the true costs of shipping to a particular region as closely as door-to-door, North American truck freight rates.

Income Effect. The income coefficient represents the percentage change in quantity of pork exported by a Canadian province in response to a 1% change in importing states' income. In the land transport gravity model all the income elasticities are positive and statistically significant and most are greater than unity. This suggests that Canadian pork exports are highly sensitive to changes in US income. This is not surprising since chilled pork is a higher valued food item.

Production Specialization Index. The Production Specialization Index (INX) is calculated as the ratio of hog inventory to human population. It measures the degree of self-sufficiency in hog production for a particular state, with larger values for INX_{jt} indicating greater degrees of self-sufficiency. For an importing region, import quantity is expected to increase as income goes up or special-

Table 3: Estimation Results for the Land Transport of Pork (Volume as Dependent Variable)

Model 5: Volume = f(Freight Rates)

$\ln Q_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

Elasticity	All Provinces	Alberta	Saskatchewan	Manitoba	Ontario	Quebec
Intercept *(t-ratio)*	10.403 *(7.821)**	47.098 *(12.32)**	6.981 *(1.759)**	6.3891 *(1.775)**	-4.1663 *(-1.278)*	13.285 *(6.550)**
Ln(Freight) (t-ratio)	-3.3705 -20.76)*	-9.3132 (-22.89)*	-1.1576 (-2.443)*	-3.3307 (-7.046)*	-1.7512 (-6.565)*	-2.3745 (-7.748)*
Ln(Income) (t-ratio)	1.9489 (23.6)*	3.1876 (30.93)*	0.25961 (2.193)*	2.2628 (9.697)*	2.4088 (10.060)*	1.2723 (7.052)*
Ln(INX) (t-ratio)	-0.026011 (-0.962)	-0.03784 (-1.724)*	-0.00988 (-0.313)	0.065213 (0.859)	0.0703 (1.659)*	-0.00775 (-0.106)
Dummy (t-ratio)	8.3979 (19.69)*	0.89804 (1.501)	20.037 (37.16)*	8.2416 (7.611)*	2.8547 (4.059)*	5.4536 (7.464)*
Observations	980	195	195	195	195	200
R-square	0.9244	0.9676	0.9406	0.8592	0.6547	0.574

Model 6: Volume = f(Distance)

$\ln Q_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

Elasticity	All Provinces	Alberta	Saskatchewan	Manitoba	Ontario	Quebec
Intercept (t-ratio)	5.2296 (4.555)*	33.337 (8.156)*	4.2895 (1.422)	3.5974 (1.125)	-7.7458 (-2.474)*	11.109 (6.151)*
Ln(Distance) (t-ratio)	-2.6672 (-19.85)*	-7.8478 (-17.77)*	-0.86877 (2.431)*	-2.94 (-7.097)*	-1.1179 (-5.315)*	-2.1255 (-8.543)*
Ln(Income) (t-ratio)	1.8825 (22.37)*	3.3289 (28.52)*	0.28202 (2.318)*	2.2107 (9.526)*	2.2935 (8.965)*	1.2503 (6.691)*
Ln(INX) (t-ratio)	-0.0239 (-0.842)	-0.0236 (-1.011)	-0.00988 (-0.327)	0.06677 (0.894)	0.11101 (2.318)*	-0.06885 (-0.914)
Dummy (t-ratio)	8.7996 (20.52)*	0.34502 (0.459)	20.002 (38.060)*	8.5324 (7.874)*	3.2648 (4.657)*	5.3201 (7.162)*
Observations	980	195	195	195	195	200
R-square	0.8932	0.9547	0.9424	0.8661	0.6069	0.568

**Denotes the coefficient estimate is significant at the 5% level*

ization index declines. The specialization index coefficient was significant in only five of the regressions and seems to hover around –0.1 and 0.1, providing an ambiguous result. When all Canadian provincial pork exports are aggregated and value is specified as the dependent variable, the specialization index is –0.14 for both the freight rate and distance estimations (Models 7 and 8). This is comparable to previous results by Prentice et al. (1998).

Table 4: Estimation Results for the Land Transport of Pork (Value as Dependent Variable)

Model 7: Value = f(Freight Rates)

$\ln V_{ijt} = \alpha + \beta_1 \ln F_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

Elasticity	All Provinces	Alberta	Saskatchewan	Manitoba	Ontario	Quebec
Intercept (t-ratio)	9.8152 (6.911)*	46.52 (8.813)*	5.3158 (1.132)	4.9138 (1.183)	-6.8451 (-1.781)*	14.306 (7.029)*
Ln(Freight) (t-ratio)	-3.4616 (-18.23)*	-9.2634 (-13.81)*	-1.1293 (-2.062)*	-3.5222 (-5.962)*	-1.7713 (-4.126)*	-2.4507 (-8.223)*
Ln(Income) (t-ratio)	2.0834 (19.81)*	3.1942 (17.23)*	0.40067 (2.586)*	2.5231 (6.548)*	2.6965 (10.62)*	1.3001 (6.863)*
Ln(INX) (t-ratio)	-0.1445 (-3.405)*	-0.14712 (-1.255)	-0.032925 (-0.675)	0.10593 (0.873)	0.073921 (1.593)	-0.05803 (-0.707)
Dummy (t-ratio)	8.4204 (18.98)*	1.3084 (1.365)	20.414 (35.46)*	8.7178 (7.354)*	2.8025 (3.444)*	4.8309 (6.688)*
Obsrvation	980	195	195	195	195	200
R-square	0.7890	0.8524	0.9251	0.6704	0.6679	0.5467

Model 8: Value = f(Distance)

$\ln V_{ijt} = \alpha + \beta_1 \ln D_{ijt} + \beta_2 \ln Y_{jt} + \beta_3 \ln INX_{jt} + \delta B_{jt} + e_{ijt}$

Elasticity	All Provinces	Alberta	Saskatchewan	Manitoba	Ontario	Quebec
Intercept (t-ratio)	5.2962 (4.152)*	26.352 (5.094)*	2.7971 (0.769)	1.7817 (0.48)	-11.754 (-3.356)*	11.675 (6.077)*
Ln(Distance) (t-ratio)	-2.7716 (-17.43)*	-6.9551 (-10.41)*	-0.882 (-2.096)*	-3.1151 (-6.122)*	-0.8966 (-2.691)*	-2.3643 (-9.644)*
Ln(Income) (t-ratio)	1.9702 (18.530)*	3.2903 (15.39)*	0.43803 (2.743)*	2.493 (6.46)*	2.5417 (9.344)*	1.4279 (7.286)*
Ln(INX) (t-ratio)	-0.14201 (-3.299)*	-0.14543 (-1.111)	-0.034733 (-0.696)	0.11582 (0.954)	0.1123 (1.986)*	-0.12848 (-1.512)
Dummy (t-ratio)	8.8703 (20.10)*	1.6913 (1.488)	20.369 (36.24)*	8.9383 (7.506)*	3.551 (4.375)*	4.8592 (6.567)*
Obsrvations	980	195	195	195	195	200
R-square	0.7723	0.7983	0.9279	0.6831	0.6263	0.5858

** Denotes the coefficient estimate is significant at the 5% level*

Dummy Variable. The dummy variable in the land transport gravity model was intended to distinguish high-volume, relational trade linkages from low-volume, spot-market trade flows. It is positive and significant for all provinces but Alberta, where the dummy variable resulted in an insignificant estimate. This suggests that US importers who regularly import chilled pork from Canadian exporters tend to import larger volumes than those states that do not consistently import Canadian pork.

Explanatory Power. The R-square value is an imperfect measure of performance, but is useful to compare alternative specifications. The data in Tables 3 and 4 indicate that the models with volume of exports as the dependent variable provide a closer fit than the models using value of exports. The R-square data does not suggest that using distance as a proxy for transport costs in a gravity model reduces its explanatory power.

Aggregate versus Disaggregate Model Performance. Estimating each of the models by Canadian province of exit and then aggregating the data to produce estimates of total Canadian pork exports tested the sensitivity of location. As cautioned in the literature, the aggregate model results appear to be merely an average of all trade flows combined. For Models 5 to 8, the aggregated results (all provinces) consistently underestimates the influence of transport costs/distance on Alberta pork exports and overestimates these influences on Saskatchewan, Ontario, and Quebec pork trade. Similar results are evident in the other variables.

CONCLUSION

Four variations of gravity models were examined for land and marine transport, giving a total of eight scenarios. For each mode of transport, one variation of the model used volumes and actual freight rates, another used volumes and distances, a third used dollar values and distances and a fourth used market values and freight rates. Comparisons of the results of the four model specifications were discussed with regard to the performance of the proxy variables and quality of estimation.

Aggregation is the soft underbelly of the gravity model. These analyses show that even aggregating the shipping locations of an exporter can produce inferior results. The specification of a more defined gravity model to a single shipping point improves the accuracy and utility of the results.

Commodity specific gravity models with a single transport mode (i.e., marine container, refrigerated truck) are likely to be more accurate, regardless of data specification, because the market under analysis is more consistent. Freight rate data that applies to a specific mode is more representative of the actual transportation costs. When the model includes actual transport costs instead of distance as the impedance factor, it more closely resembles traditional demand equations, and is thus more useful for marketers, transportation service providers, and policy analysts. An elasticity with regard to distance is difficult, if not impossible to interpret.

Ultimately, for accurate, useful results, it is preferred to specify export volumes and transport costs in a gravity model. If only a single commodity and mode of transport are under analysis, it would appear that proxies (distance and value) would still yield acceptable results, but the margin of error is greater. This study does not suggest that proxy variables are acceptable for aggregate flows. In fact, the results cast doubt on the usefulness of aggregate product gravity models in general.

Endnotes

1. Global containerized trade imbalances are a significant influence on freight rate structures. In the westbound trade to Asia, the equivalent of one empty container is returned for every two full containers coming across the Pacific to North America. Whereas eastbound to Europe, empty returns account for roughly only 20% of all slots on containerships (*The Economist* 2002). International terminal and transshipment charges vary considerably from port to port. Consequently, Pacific freight rates may not reflect the true costs of shipping to that region as closely as Atlantic freight rates do. Goods shipped through highly modern and efficient European ports may face lower handling costs (i.e., reduced transit times and costs). Moreover, land costs in Asian countries tend to be prohibitively high. These imbalances, combined with intermodal equipment shortages on the Canadian prairies create an ambiguous outcome.

References

Alcaly, R.E. "Aggregation and Gravity Models: Some Empirical Evidence." *The Journal of Regional Science* 7(1) (1967): 61-73.

Anderson, J.E. "A Theoretical Foundation for the Gravity Equation." *The American Economic Review* 69(1) (1979): 106-116.

Bandyopadhyay, U. "Trade and the Distribution Sector: Some Evidence from OECD Countries." *Canadian Journal of Economics* 32(5) (1999): 1299-1312.

Bergstrand, J.H. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *The Review of Economics and Statistics* 67(3) (1985): 474-481.

Canada Grains Council, *Statistical Handbook,* 2000.

Deardorff, A.V. "Determinants of Bilateral Trade Flows: Does Gravity Work in a Neoclassical World?" *The Regionalization of the World Economy,* University of Chicago Press, 1998.

Economist, The. "Drawing the Line," 363(8269) (April 4, 2002): 62.

Ferguson, D.C. "Joint Products and Road Transport Rates in Transport Models." *Journal of Transportation Economics and Policy* 6(1) (1972): 69-76.

Geraci, V., and W. Prewo. "Bilateral Trade Flows and Transport Costs." *The Review of Economics and Statistics* 59(1) (1977): 67-74.

Hillberry, R. "Aggregation Bias, Compositional Change, and the Border Effect." *Office of Economics Working Paper,* US International Trade Commission, Washington, DC, 2001.

Industry Canada, "Trade Data Online," (2001) http://strategis.ic.gc.ca/sc_mrkti/tdst/engdoc/tr_homep.html

Linneman, H. "An Econometric Study of International Trade Flows." North Holland Publishing Company, Amsterdam, 1966.

Poyhonen, P."A Tentative Model for the Volume of Trade Between Countries." *Weltwirtschaftliches Archive* 90 (1963): 93-100.

Prentice, B.E., Z. Wang, and H. Urbina. "Derived Demand for Refrigerated Truck Transport: A Gravity Model Analysis of Canadian Pork Exports to the United States." *Canadian Journal of Agricultural Economics* 46 (1998): 317-328.

Rand McNally and Company. *Rand McNally Road Atlas 2000: United States, Canada and Mexico, 2000 Millennium Atlas,* 2000.

Tinbergen, J. *Shaping the World Economy—Suggestions for an International Economic Policy,* The Twentieth Century Fund, 1962.

United States Department of Agriculture, National Agricultural Statistics Service, http://www.usda.gov/nass, 1992.

United States Department of Commerce, *Survey of Current Business* 74(1) (1996): 44-47.

Urbina-Olano, H.J. "Elasticity of Demand for Red Meat Transportation: A Gravity Model Analysis of Western Canadian Pork Exports." Masters Thesis, University of Manitoba, 1996.

World Bank. *World Development Indicators CD-Rom,* 2000.

Erica Vido *is a graduate student in the Department of Agricultural Economics at the University of Manitoba. She also works full-time as a research affiliate with the University of Manitoba Transport Institute. Her interests and areas of research include international trade, agricultural transportation, intermodal transportation, and port operations. She is a member of the Transport Institute team examining the economics and logistics of shipping Western Canadian Identity Preserved Grains in containers, and she co-authored an award winning paper on this topic presented at the 2000 Canadian Transportation Research Forum. Vido has been involved in several other research projects including economic impact analyses, transport policy analyses, and tourism.*

Barry E. Prentice *is director of the Transport Institute, and associate professor, I.H. Asper School of Business at the University of Manitoba. He holds degrees from the University of Western Ontario (B.A.), University of Guelph (M.Sc.), and University of Manitoba (Ph.D.). His major research and teaching interests include logistics, transportation economics, agribusiness marketing, and trade policy. In 1999, Prentice was named Manitoba Transportation Person of the Year and received a University of Manitoba outreach award. In 1998, he won the TRF Rail Tex Paper Award. He serves on the Boards of Directors of several transportation organizations: Winnipeg Airports Authority, Inc., National Transportation Week (president, 2001 and 2003), and the Canadian Transportation Research Forum (past president, 1997). Prentice is honorary president of the Canadian Institute for Traffic and Transportation (2001-2003). In addition, he serves on task forces, expert committees, and is frequently asked to speak on the topics of trade and transportation.*