ISSN 2234-8867
Journal of East Asian Economic Integration Vol. 16, No. 3 (September 2012) 227-247

Does Specialization Matter for Trade Imbalance at Industry Level?

E. Young Song
Department of Economics, Sogang University
eysong@sogang.ac.kr

Chen Zhao
Department of Economics, Hong Kong University of Science and Technology
zhaochen@ust.hk

This paper investigates the source of bilateral trade imbalance at industry level. We build a simple model based on gravity theory and derive the prediction that the bilateral trade balance in an industry is increasing in the difference between trading partners in the output share of the industry. We test this prediction and find that the difference in industry share is highly significant in predicting both the sign and the magnitude of trade balance at industry level. We also find that FTAs tend to enlarge trade imbalance at industry level. However, the overall predictive power of the model is rather limited, suggesting that factors other than production specialization are important in determining trade balance at industry level. Another finding of the paper is that the influence of the difference in industry share on trade balance increases as we move to industries that produce more homogeneous products. This finding calls into question monopolistic competition as the main driver of gravity in international trade.

Keywords: Trade imbalance, Gravity theory, Specialization, Output share, Homogeneous products
JEL Classification: F11, F12, F14

I. Introduction

Bilateral trade imbalance is a sensitive issue in international politics. The gravity of the issue is magnified when repeated and large trade deficits become a central question in elections. The US trade deficits with Japan during 1980s and those with China during 2000s are notable examples. The deficit country often urges the surplus country to adopt policy measures to correct the

* This paper is an extension of Zhao (2009). The authors are grateful to an anonymous referee for the valuable comments that substantially improved the paper.

"undesirable" situation. If the surplus country does not comply, retaliatory measures can be taken by the deficit country, increasing the risk of trade wars.

More often, tensions are triggered by overall trade imbalance. However, trade deficits at industry level also become thorny issues in trade negotiations. Korea's trade surplus with the United States in automobile industry has been a perennial issue in Korea-US trade relation. The U.S. insists that the existence of large deficits in automobiles trade is evidence for the unfair trade practices of Korea. Under the U.S. pressure, Korea had to replace the already-signed FTA by a new one that is more favorable to the U.S. automobile industry. Korea's trade deficits with China in agricultural industries and its deficits with Japan in parts and machinery industries also have been barriers against a successful negotiation toward a Korea-China or a Korea-Japan FTA.

In contrast, economists consider trade imbalance at industry level as a natural phenomenon, and furthermore, they think that gains from trade originate from generating trade deficits or surpluses at industry level. Trade imbalance at industry level is the realization of comparative advantages, and it should be fostered through freer trade. Any political action trying to "correct" trade imbalance in specific industries should be viewed as welfare-deteriorating.

To justify this view of economists at an empirical level, we have to find evidence that the inter-industry specialization of trading partners indeed is a major determinant of trade imbalance at industry level. To pursue this task, this paper constructs a model based on gravity theory to predict industry-level trade balance. According to a simple version of the model, bilateral trade balance in an industry is proportional to the difference between trading partners in the output share of the industry. In a more general version that allows for trade costs, we derive the prediction that a normalized trade balance in an industry is increasing in the difference in logarithmic industry share, controlling for trade costs that may vary with country pairs and industries.

We test these predictions of the model. We find that the difference in industry share is significant in predicting the sign of trade balance at industry level. We also find that the difference in logarithmic industry share is highly significant in predicting the size of the normalized trade balance, controlling for trade costs and various fixed effects. However, the overall predictive power of the model is limited, suggesting that incomplete specialization, non-homothetic or non-identical preferences may be playing important roles in determining bilateral trade balance at industry level.

In addition, we test whether FTAs increase the degree of industry-level trade imbalance and find a positive correlation between FTAs and trade imbalance. This evidence suggests that FTAs will lead to more political tensions over industry-level trade imbalance. Finally, we explore the question whether the effect of the difference in industry share on industry-level trade balance decreases as we move to industries that produce more homogeneous products. The empirical results show that the opposite tendency exists. This evidence raises doubt about monopolistic competition as the main driver of gravity in international trade.

Davis and Weinstein (2002) conducted a study to investigate whether gravity theory can explain the variation of bilateral trade balance across county pairs and industries. They found that the variance of predicted trade balance is much smaller than the actual one, and suggested that we should look for an alternative explanation for large trade imbalances. Our approach is different from theirs in that we derive and directly estimate the equation determining trade imbalance, while they use two-stage procedures. By using a normalized trade balance and a normalized difference in industry share, we also purge our variables of any scale effect, which gravity theory owes a lot for its good performance. Despite this increased rigor, we find that our equation explains a much larger portion of the variation in trade imbalance than their approach. A number of studies fit gravity equations at industry level, and many of them investigate how the performance of gravity equations varies with industry characteristics. Harrigan (1996), Rauch (1999), Feenstra, Markusen and Rose (2001) and Evans (2003) are just a few examples. Our paper is a variation of this line of research, but none of these studies tests the performance of gravity theory in predicting trade balance. From a broader perspective, studies on the determinants of intra-industry trade can be viewed as research on industry-level trade imbalance. The Grubel-Lloyd index of intra-industry trade is measured as one minus net trade over gross trade. Thus factors that influence trade imbalance also affect the index of intra-industry trade. Caves(1981), Toh(1982) and Greenaway-Hine-Milner(1995) are studies that attempt to explain the variation of intra-industry across industries by industry characteristics such as economies of scale and product differentiation. Theoretically more related are studies by Helpman (1987) and Hummels and Levinsohn (1995). They test whether bilateral intra-industry trade are inversely related to specialization induced by the difference between trading partners in factor endowments. Song and Sohn (2012) also try to explain the variation of bilateral intra-industry trade by specialization due to the difference in labor productivity. However, none of these studies investigates the relation between intra-industry trade and specialization at industry level.

To our knowledge, the only paper that directly estimates an equation for industry-level trade balance is Chung et al. (2008). As this paper, they derive a regression equation in the spirit of gravity theory. However, their equation has not been derived from a theory and is seriously misspecified. Consequently, we are not provided any clue over what signs are expected for the coefficients of their variables, and whether their estimation results support gravity theory or not. In contrast, we derive an exact functional form for our estimation equation directly from gravity theory. We prove that a normalized trade balance should be positively related to the difference in logarithmic industry share, and the coefficient on the latter should be equal to unity. Our controls for trade costs also are founded on gravity theory and are more comprehensive than those used by Chung et al. (2008). The scope of our paper also is larger in that we use data on all available country pairs, while Chung et al. (2008) restrict their estimation to country pairs involving Korea. That we estimate the effects of FTAs on trade imbalance can also be considered as value-added.

The paper is organized in the following way. Section II derives an estimation model. Section III conducts empirical tests. Section IV concludes.

II. Some Theoretical Observations

We base our estimation model on a world of N countries where complete specialization prevails in world production (each good is produced by only one country) and all countries share identical homothetic preferences. X_{ij}^h denotes the value of good h shipped from country i to country j. Let Y_i be the value of goods produced in country i, E_i be the total expenditure of country i residents, and Y_w be the value of world production. We use Y_i^h to denote the value of good h produced in country i. Suppose that there is no trade barrier and trade is completely frictionless. Then we can easily show that the following equation holds for every country pair (i , j).

$$X_{ij}^h = \frac{Y_i^h E_j}{Y_W}. \tag{1}$$

In other words, Y_i^h is distributed across importing countries in proportion to their expenditure sizes. Equation (1) is nothing but a (frictionless) gravity equation that holds at good level.

Let K_i be the set of goods that country *i* produces in industry *k*. From (1),

$$X_{ij}^{k} = \frac{Y_i^k E_j}{Y_W}, \tag{2}$$

where

$$X_{ij}^{k} = \sum_{h \in K_i} X_{ij}^{h},$$
$$Y_i^k = \sum_{h \in K_i} Y_i^h.$$

Let us define

$$\alpha_i^k \equiv \frac{Y_i^k}{Y_i}, \tag{3}$$

$$\lambda_i \equiv \frac{Y_i - E_i}{Y_i}. \tag{4}$$

α_i^k is the output share of industry *k* in country *i*, and λ_i is the trade surplus of country *i* as a percentage of total output. Using (2), (3) and (4), we can derive the following equation for trade balance in industry *k* between country *i* and country *j*.

$$X_{ij}^{k} - X_{ji}^{k} = ((1-\lambda_j)\alpha_i^k - (1-\lambda_i)\alpha_j^k)\frac{Y_i Y_j}{Y_w}. \tag{5}$$

Or

$$BAL_{ij}^{k} = DIF_{ij}^{k} \frac{Y_i Y_j}{Y_w}, \tag{6}$$

where

$$BAL_{ij}^k \equiv X_{ij}^k - X_{ji}^k,$$
$$DIF_{ij}^k \equiv SHARE_{ij}^k - SHARE_{ji}^k,$$
$$SHARE_{ij}^k \equiv (1-\lambda_j)\alpha_i^k,$$
$$SHARE_{ji}^k \equiv (1-\lambda_i)\alpha_j^k.$$

(6) is a key equation for our empirical investigation. It tells us that in a world of complete specialization and frictionless trade, bilateral trade balance at industry level is determined by the difference in industry share adjusted for overall trade imbalance (DIF_{ij}^k) and the product of total output levels ($Y_i Y_j / Y_w$). Given the levels of total output, as the degree of inter-industry specialization between two countries intensifies, the absolute values of the differences in industry share will increase toward unity and the extent of bilateral trade imbalances at industry level will intensify. Thus we can evaluate the role of inter-industry specialization in determining bilateral trade balance at industry level by estimating (6).

A problem in applying equation (6) to actual trade flows is that it does not consider the effects of trade barriers. A clue for extending equation (6) for a world with trade costs can be found in the work of Chaney (2008). From a variation of the Melitz (2003) model where monopolistic competition among heterogeneous firms and fixed costs of entry determine trade flows, he derives the following equation.

$$X_{ij}^k = \mu_k \frac{Y_i Y_j}{Y_W}\left(\frac{w_i \tau_{ij}}{\theta_j}\right)^{-\gamma}\left(f_{ij}\right)^{-[\gamma/(\sigma-1)-1]}. \quad (7)$$

μ_k is the expenditure share of goods produced in industry k in Cobb-Douglas preferences, w_i is the overall productivity level of country i, τ_{ij} is transportation cost between i and j, γ is the shape parameter for a Pareto distribution that characterizes the productivity distribution of firms (the larger γ gets, firms become more homogeneous in terms of productivity), σ is the elasticity of substitution among goods produced in an industry, and f_{ij} is the fixed cost of exporting from country i to country j. θ_j is an index that captures the remoteness of country i from the rest of the world. τ_{ij}, f_{ij}, θ_j, γ and σ can take different values for different industries. Using (7) and the accounting requirement that $\sum_j X_{ij}^k = \alpha_i^k Y_i^k$, we can show that the following equation holds.

$$X_{ij}^k = \alpha_i^k \frac{Y_i\, Y_j}{Y_W} \left(\frac{\tau_{ij}}{\pi_i\, \theta_j} \right)^{-\gamma} \left(f_{ij} \right)^{-[\gamma/(\sigma-1)-1]}, \tag{8}$$

where

$$\pi_i = \left(\sum_j \frac{Y_j}{Y_W} \left(\frac{\tau_{ij}}{\theta_j} \right)^{-\gamma} \left(f_{ij} \right)^{-[\gamma/(\sigma-1)-1]} \right)^{-\frac{1}{\gamma}},$$

$$\theta_j = \left(\sum_i \frac{Y_i}{Y_W} \left(w_i \tau_{ij} \right)^{-\gamma} \left(f_{ij} \right)^{-[\gamma/(\sigma-1)-1]} \right)^{-\frac{1}{\gamma}}.$$

One can note that (8) is similar to the gravity equation derived by Anderson and van Wincoop (2003) except that it contains the fixed cost of exporting. Though the approach of Chaney (2008) is popular in empirical studies of bilateral trade flows, it is not totally satisfactory for our purpose. The derivation of the equation relies on the assumption that the productivity distribution of firms in an industry is identical among countries, and therefore it cannot properly capture trade pattern driven by inter-industry specialization, the focus of classical trade theories. What we need is a gravity equation that captures both inter-industry specialization and monopolistic competition among heterogeneous firms[1], but from (8) we can still conjecture how (6) should be modified to accommodate trade costs.

[1] Bernard, Redding and Schott (2007) have made an important progress toward this goal, but we still lack a workable model ready for empirical applications.

To derive a regression equation from (8), we replace α_i^k by $SHARE_{ij}^k$ to allow for trade imbalance at country level, and take the logarithm of both sides.

$$\log X_{ij}^k = \log SHARE_i^k + \log \frac{Y_i \, Y_j}{Y_W} - \gamma \log \tau_{ij} + \gamma \log \pi_i + \gamma \log \theta_j - [\frac{\gamma}{\sigma - 1} - 1] \log f_{ij} \tag{9}$$

From (9),

$$BALNORM_{ij}^k = DIFNORM_{ij}^k + \delta_i + \delta_j + \delta_{ij}, \tag{10}$$

where

$$BALNORM_{ij}^k \equiv \log X_{ij}^k - \log X_{ji}^k,$$

$$DIFNORM_{ij}^k \equiv \log SHARE_{ij}^k - \log SHARE_{ji}^k,$$

$$\delta_i = \gamma \ \log(\pi_i / \theta_i),$$

$$\delta_j = \gamma \ \log(\theta_j / \pi_j),$$

$$\delta_{ij} = -\gamma \log(\tau_{ij} / \tau_{ji}) - [\frac{\gamma}{\sigma - 1} - 1] \log(f_{ij} / f_{ji}).$$

By a Taylor approximation, we can show that

$$\log y - \log z \approx \frac{y - z}{(y + z)/2}. \tag{11}$$

Thus $BALNORM_{ij}^k$ is approximately equal to trade surplus normalized by gross trade (divided by 2) in industry *k*. Likewise, $DIFNORM_{ij}^k$ is approximately equal to the difference in industry share normalized by the sum of industry shares (divided by 2). Thus (10) can be considered as a normalized version of (6), but it also incorporates the effects of trade costs.[2,3]

It is hard to construct the remoteness variables δ_i and δ_j. As becomes standard in gravity research, we will capture their effects simply by inserting exporter and importer dummies.[4] We can take two approaches regarding bilateral trade costs δ_{ij}. One is to capture them by adding 'gravity variables' that are frequently used in estimating gravity equations.

$$\delta_{ij} = \rho \log DIST + BORDER + LANG + ISLAND + ISLAND + LANDLOCK + COLONY + FTA \quad (12)$$

DIST : physical distance between *i* and *j*,
BORDER: "1" for border sharing, "0" otherwise,
LANG: "1" for a common language, "0" otherwise,
ISLAND: "2" if both countries are islands, "1" if one of them is, "0" otherwise,
LANDLOCK: "2" if both countries are landlocked, "1" if one of them is, and "0" otherwise,
COLONY: "1" if one was a colony of the other, "0" otherwise,
FTA: "1" if countries are in a free trade agreement or a customs union, "0" otherwise.

2 In a small percentage of our sample, we have zero exports or imports. In this case, *BALNORM* is not defined and so we have replaced them by 1,000 dollars. This practice, used by many authors, can be justified on the ground that zero trade flows in trade data often are caused by rounding small numbers to zero. This practice essentially is ad hoc, but dropping these observations or replacing them by trade balance divided by gross trade change none of our conclusions below.

3 Using our notation, the key estimation equation used by Chung et al. (2008) can be written as:

$$\log \frac{|X_{ij}^k - X_{ji}^k|}{X_{ij}^k + X_{ji}^k} = \eta_0 + \eta_1 \log \frac{Y_i^k Y_j}{Y_W} + \eta_2 \log \frac{Y_j^k Y_i}{Y_W} + other\ controls\ .$$

We can see from (10) that this equation should instead be written as:

$$\log \frac{|X_{ij}^k - X_{ji}^k|}{X_{ij}^k + X_{ji}^k} = \eta_0 + \log \left| \log \frac{Y_i^k}{Y_i} - \log \frac{Y_j^k}{Y_j} + other\ controls \right| .$$

By comparing these equations, we can see the extent of misspecification involved and why it is difficult to give a theoretical interpretation to their results that η_1 and η_2 are positive.

4 See, for example, Baldwin and Taglioni (2006).

The alternative approach is inserting country-pair dummies to capture bilateral trade costs. This method will filter out all time-invariant fixed effects specific to country pairs.

III. Empirical Results

The data comes from Nicita and Olarreaga (2006) who compiled data on production and trade in 28 manufacturing industries (ISIC 3 digit, Revision 2) of 100 countries over the period 1976-2004. The data on production are from the UNIDO Industrial Statistics Database. The data on trade are from the UN Comtrade Database, but they reclassified them according to the 3 digit ISIC codes to match with the industry classification of the production data. This data set perfectly suits our purpose because our model requires that trade balances and the differences in industry share should use an identical industry classification. Note that the data on production and trade flows are from manufacturing industries, and agricultural and service products are not present in our data. We use gross output to calculate Y_i^k, and mirrored exports (*i*'s exports as reported by *j*) to calculate X_{ij}^k as they are known to be more reliable. Data on overall trade surplus were obtained from the IMF International Financial Statistics and all gravity variables other than *FTA* were imported from Rose (2004). FTA dummies were constructed from the WTO website.[5]

From the data sets, we construct a panel of bilateral trade flows by industry (X_{ij}^k) and the industry shares of trading partners adjusted for trade surpluses ($SHARE_{ij}^k$ and $SHARE_{ji}^k$). To calculate the industry shares of a country in a given year, we need to observe gross output for *all* industries. A large number of missing values in gross output data seriously constrain the data availability of industry shares. To increase the size of our panel, we dropped four industries that seem to contain too many missing values.[6] In benchmark regressions, we use data from every fourth year starting from 1976 through 2002. The main reason for using quadrennial data instead of annual series is to keep the dimension of dummy variables within the computational limit of the program. However, we also have other considerations: the differences in industry share

[5] We thank Jung Hur for kindly providing the data.

[6] These industries are tobacco (314), petro refineries (353), other petro and coal products (354) and pottery and chinaware (361).

are slowly changing, and annual trade balances contain a high degree of serial correlation. However, none of our results seems to be sensitive to data frequency. The resulting panel contains observations over 56 countries and 1,193 country pairs. The total number of observations is 55,912 and the number of exporter-importer-industry triples is 22,420. The panel is severely unbalanced and the average number of observations per exporter-importer-industry is 2.5. Table 1 summarizes the statistics for key variables.

Table 1. Summary of key variables

Variables	Mean	Std. Dev.	Min	Max	Obs
BAL_{ij}^{k}	3,306.5	268,347.5	-6,716,647	2.92e+07	55,912
DIF_{ij}^{k}	-.0002	.05	-.40	.60	55,912
$BALNORM_{ij}^{k}$	-.0246	3.90	-13.79	12.36	55,912
$DIFNORM_{ij}^{k}$	.0010	1.11	-8.16	9.50	55,912

Trade balances are in thousand US dollars.

We base the first two of our tests on (6).

$$BAL_{ijt}^{k} = DIF_{ijt}^{k} \frac{Y_{it} Y_{jt}}{Y_{wt}}, \tag{13}$$

We added time subscripts to emphasize that variables change over time. (13) ignores the existence of trade barriers, and thus instead of examining its quantitative performance, we evaluate its power on predicting the sign of trade balance. According to (13), trade balance should be positive (negative) when the difference in industry share is positive (negative). Thus we conduct a sign test and Table 2 reports the results.

Table 2. Sign test: Prob[$BAL_{ijk} \times DIF_{ijk} > 0$] = 0.5

N	Expected Matches	Observed Matches	Assumed Probability	Observed Probability
55,912	27,956	35,843	0.50	0.64

The null hypothesis that the probability of *BAL* and *DIF* having the same sign is 0.5 is strongly rejected in favor of the alternative that the probability is greater than 0.5. Out of 55,912 observations, 64% have matched signs, and the chance for a random tossing to get matching higher than 64% is almost zero. Thus (13) passes the sign test, but its predictive power may not be seen as high as we should expect.

Next, we test our prediction using Probit. As we saw in the previous section, trade balance and the difference in industry share may have different signs if trade costs are not symmetric, even when we have complete specialization. Thus we ask this time whether the probability that trade balance is positive increases as the difference in industry share increases. Table 3 shows that this tendency strongly is present in our data. The coefficient on *DIF* in regression (1) is 7.23 and is significant. The estimation result implies that the probability that trade is in surplus increases from 50% to 64% when the difference in industry share increases from zero by one standard deviation (0.05). Regression (2) estimates the effects of $SHARE_{ij}^{k}$ and $SHARE_{ji}^{k}$ separately. As expected, the former has a positive effect on the probability of trade surplus, while the latter has a negative effect. Both of them are highly significant and have sizable effects on the sign of trade balance.

Table 3. Probit analysis

$\Pr[BAL_{ij}^{k} > 0]$	(1)	(2)
DIF_{ij}^{k}	7.23 (59.86)	
$SHARE_{ij}^{k}$		6.96 (53.87)
$SHARE_{ij}^{k}$		-7.61 (-54.75)
Constant	0.01 (1.57)	0.04 (4.97)
Obs.	55,912	55,912

The numbers in the parentheses are z-ratios.

We evaluate the quantitative performance of gravity theory in predicting industry-level trade balance using (10).

$$BALNORM_{ijt}^{k} = DIFNORM_{ijt}^{k} + \delta_k + \delta_{it} + \delta_{jt} + \delta_{ij}\,, \tag{14}$$

where

$$BALNORM_{ijt}^{k} \equiv \log X_{ijt}^{k} - \log X_{jit}^{k}\,,$$
$$DIFNORM_{ijt}^{k} \equiv \log SHARE_{ijt}^{k} - \log SHARE_{jit}^{k}\,.$$

We added time subscripts on variables to clarify that they are time-dependent. Note that the remoteness variables δ_{it} and δ_{jt} are treated as time-dependent as the remoteness of countries can change over time as trade costs change. These dummies also can depend on industries as trade costs vary over industries. However, allowing them to vary over industries pushes us off the computational limit of the program so in benchmark regressions, we capture the dependence of trade costs on industries simply by including industry dummies δ_k. As noted above, bilateral trade costs can be captured by gravity variables or country-pair dummies. In the latter case, we could allow δ_{ij} to vary over time (and drop δ_{it} and δ_{jt} to avoid collinearity), but this method is not computationally feasible. In some of our regressions, we will allow δ_{ij} to change across industries. Using exporter-importer-industry dummies δ_{ijk} has the advantage of eliminating bias caused by industry characteristics that make trade costs (or any omitted variables) and the difference in industry share co-move across industries. When we use fixed effects δ_{ijk}, we should drop δ_k from (14).

Table 4 reports the estimation results for a simple version of equation (14) where we drop all trade costs variables. In regressions below, we report *t*-ratios calculated using robust standard errors clustered by exporter-importer pairs, unless otherwise stated. Here, we would like to check how much the single variable of *DIFNORM* can explain the variation of trade balance across country pairs and industries. In regression (3), we can see that *DIFNORM* alone explains about 14% of trade balance variations. The coefficient on *DIFNORM* is 1.31 and highly significant, but it is well above the theoretically correct value of 1.

In regression (4), we include $\log SHARE_{ij}^{k}$ and $\log SHARE_{ji}^{k}$ as separate regressors. We can note that R^2 hardly increases and the effects of $\log SHARE_{ij}^{k}$ and $\log SHARE_{ji}^{k}$ are almost symmetric, even though they are different in a statistical sense. Thus, in the following, we will concentrate on the performance of the single variable *DIFNORM*. In regressions (5), (6) and (7), we estimate the same model for three different samples. Here *North* means countries whose income per capita is in the high or mid-high range according to the World Bank classification, and *South* means countries whose income per capita is in the mid-low or low range. We do this exercise to see whether the bite of the gravity model weakens as the sample includes more developing countries where monopolistic competition is not likely to be the right model for their trade patterns. We indeed find that the explanatory power (R^2) of the model decreases as we move from the *North-North* sample to the *South-South* sample. However, even in the *South-South* sample, the simple model still explains 12% of trade balance changes and the significance of *DIFNORM* cannot be questioned.

Table 4. Fitness of the simple model

$BALNORM_{ij}^{k}$	(3)	(4)	(5)	(6)	(7)
Sample			*North-North*	*North-South*	*South-South*
$DIFNORM_{ij}^{k}$	1.31 (51.56)		1.45 (35.89)	1.23 (35.25)	1.21 (14.95)
$\log SHARE_{ij}^{k}$		1.37 (46.49)			
$\log SHARE_{ji}^{k}$		-1.27 (-42.95)			
R^2	0.14	0.14	0.17	0.13	0.12
Obs.	55,912	55,912	29,002	23,423	3,487

The numbers in the parentheses are *t*-ratios calculated using robust standard errors (clustering by exporter-importer pairs). Constants are not reported.

We now turn to the full version of (14) and ask whether the difference in industry share remains significant after controlling for the effects of trade costs. Regression (8) control for bilateral trade costs by gravity variables. We also add dummy variables for catching the fixed effects for industry (δ_k), exporter-year (δ_{it}) and importer-year (δ_{jt}). The estimated coefficients on these variables are not reported. Now the coefficient on *DIFNORM* is reduced to 1.06, which is much closer to the theoretical value of 1. Its significance is beyond doubt. In regression (9), we capture bilateral trade costs through time-invariant exporter-importer dummies. The coefficient on *DIFNORM* changes little and it remains extremely significant. In regressions (10) and (11), we replace exporter-importer dummies by exporter-importer-industry dummies. Here we allow transportation costs to vary over industries and test the exploratory power of

DIFNORM on the time-series variation of trade balance within an industry of a given country-pair. Time-varying country fixed effects remain in the regressions. Reported R^2s are from within regressions and should not be compared with those in regression (8) and (9). In regression (10), we find that the coefficient is reduced to 0.13, but it is still significant at 1%. In regression (11), we report *t*-ratios calculated using robust standard errors clustered by exporter-importer-industry triples (i, j, k). The coefficient on *DIFNORM* still is significant at 1% level.

Table 5. The effects of the difference in industry share after controlling for trade costs

$BALNORM_{ij}^{k}$	(8)	(9)	(10)	(11)
$DIFNORM_{ij}^{k}$	1.06 (40.77)	1.06 (40.31)	.13 (4.32)	.13 (4.49)[b]
Gravity Vars.	Yes	No	No	No
Fixed Effects	*industry* *exporter-year* *importer-year*	*industry* *exporter-year* *importer-year* *exporter-importer*	*exporter-year* *importer-year* *exporter-importer-industry*	*exporter-year* *importer-year* *exporter-importer-industry*
R^2	0.37	0.41	0.11[a]	0.11[a]
Obs.	55,912	55,912	55,912	55,912

The numbers in the parentheses are t-ratios calculated using robust standard errors (clustering by exporter-importer pairs). Constants are not reported.

[a] : R^2s for the regressions using exporter-importer-industry dummies are from within-regression for the panel where a group consists of an exporter-importer-industry triple.

[b] : *t*-ratios calculated using robust standard errors (clustering by exporter-importer-industry triples)

We have confirmed that the difference in industry share has a significant explanatory power in predicting the sign and the magnitude of trade balance at industry level. Our final exercise is to explore two questions that might be answered by the estimation model that we have developed. The first question is whether the influence of the difference in industry share on industry-level trade balance weakens as we move to industries that are less likely to be governed by monopolistic competition. Popular gravity models are based on specialization due to monopolistic competition, and monopolistic competition derives from product differentiation. We would not expect that monopolistic competition prevails in industries where standardized raw materials like food, industrial chemicals and metals are produced. To test this hypothesis, we use the index of product homogeneity developed by Rauch (1999). Table 6 shows Rauch indexes for 3-digit ISIC industries. He calculated these indexes based

Table 6. Rauch index of product homogeneity

ISIC	Description	Rauch Index
311	Food products	0.69
313	Beverages	0.54
(314)	Tobacco	0.92
321	Textiles	0.26
322	Wearing apparel, except footwear	0.00
323	Leather products	0.00
324	Footwear, except rubber or plastic	0.02
331	Wood products, except furniture	0.49
332	Furniture, except metal	0.00
341	Paper and products	0.51
342	Printing and publishing	0.00
351	Industrial chemicals	0.53
352	Other chemicals	0.07
(353)	Petroleum refineries	0.97
(354)	Miscellaneous petroleum and coal products	0.95
355	Rubber products	0.00
356	Plastic products	0.00
(361)	Pottery, china, earthenware	0.00
362	Glass and products	0.08
369	Other non-metallic mineral products	0.54
371	Iron and steel	0.47
372	Non-ferrous metals	0.66
381	Fabricated metal products	0.15
382	Machinery, except electrical	0.00
383	Machinery, electric	0.03
384	Transport equipment	0.00
385	Professional and scientific equipment	0.00
390	Other manufactured products	0.00

Industries in the parenthesis are not included in regressions due to data availability.

on the percentage of products in an industry whose market prices are internationally accessible. More accessibility means more homogeneity. The table shows that in an industry like furniture where products are highly differentiated, the Rauch index is equal to zero, while the index for food products is as high as 0.69.

Using these indexes, we test the hypothesis that the influence of the difference in industry share on trade balance decreases as we move toward industries with undifferentiated products. We do this by including in regressions an interacted variable *RAUCH*DIFNORM*. The hypothesis is plainly rejected. In regression (12), we control for bilateral trade costs by country-pair dummies, and in regression (13), by exporter-importer-industry dummies. In both, we find that the coefficient on *RAUCH*DIFNORM* is positive and significant at 1%. The difference in industry share has a stronger effect on trade balance in more homogeneous industries. This puzzling result suggests that the predictive power of our model does not really stem from monopolistic competition. Indirectly, it also suggests that the good performance of gravity equations in explaining trade flows is not based on monopolistic competition. Our finding is closely related to the debate among

Table 7. The influence of product homogeneity and FTAs

$BALNORM_{ij}^{k}$	(12)	(13)	(14)	(15)
$DIFNORM_{ij}^{k}$	0.94 (31.80)	.08 (2.49)	1.03 (39.31)	.12 (3.92)
$RAUCH^{k} * DIFNORM_{ij}^{k}$	0.67 (7.10)	0.37 (2.59)		
$Sign\ (BALNORM_{ij}^{k}) * FTA_{ij}$			0.97 (16.35)	0.39 (6.59)
Fixed Effects	*industry exporter-year importer-year exporter-importer*	*exporter-year importer-year exporter-importer-industry*	*industry exporter-year importer-year exporter-importer*	*exporter-year importer-year exporter-importer-industry*
R^2	0.41	0.11[a]	0.41	0.11[a]
Obs.	55,912	55,912	55,912	55,912

The numbers in the parentheses are t-ratios calculated using robust standard errors (clustering by exporter-importer pairs). Constants are not reported.

[a] : R^2s for the regressions using exporter-importer-industry dummies are from within-regression for the panel where a group consists of an exporter-importer-industry triple.

Helpman (1987), Hummels and Levinsohn (1995) and Debaere (2005) on the relevance of monopolistic competition for the fitness of gravity theory, and our result supports the view of Hummels and Levinsohn (1995). Our result also supports Song (2011) who demonstrates that the simple gravity equation like (2) can be derived from models based on Cournot competition and other models of competition among firms producing homogeneous goods.

The second question that we explore is more policy-oriented. We ask whether an FTA would enlarge trade imbalance at industry level. An FTA lowers trade barriers and it stimulates both exports and imports in an industry. We would like to check whether trade imbalance as the percentage of total trade (the absolute value of *BALNORM*) will increase with an FTA, given the difference in industry share.[7] We cannot test this hypothesis simply by adding an FTA dummy in a regression. If an FTA increases the normalized trade balance when it is in surplus and decreases it (increases its absolute value) when it is in deficit, the effect of FTAs will be cancelled out in our previous regression equation. Thus, to check whether FTAs increases the absolute value of trade balance, we include in a regression a new variable: sign(*BALNORM*)**FTA*. The variable is equal to *FTA* when trade is in surplus and is equal to - *FTA* when trade is in deficit. In regression (14) where we control for trade costs by industry, country-year and country-pair dummies, we find that the coefficient on this variable is equal to 0.97 and highly significant. The result implies that the absolute value of the normalized trade balance will increase almost by 100% with an FTA. In regression (15), we control for trade costs by exporter-year, importer-year and exporter-importer-industry dummies. We consider this case to be a more relevant exercise because what matters for policy makers is whether trade imbalance at industry level will increase after an *FTA*, not whether trading partners with an FTA run larger trade imbalance than those without an agreement. Here again, we find that an FTA tends to increase trade imbalance, now by 39%. This result suggests that FTAs will work toward increasing tensions and disputes over trade imbalance at industry level.

[7] Note that this question is different from the one whether an FTA will increase trade imbalance by increasing the difference in industry share. We will pursue this question in a future research.

IV. Conclusion

A contribution of this paper is to show that gravity theory implies that the ratio of trade balance to gross trade in an industry is increasing in the difference in logarithmic industry share. Studies that take a similar approach as ours, such as Davis and Weinstein (2002) or Chung et al. (2008), do not utilize this prediction of gravity theory. We also test the empirical performance of this relationship, and find that it is supported by data on production and trade flows.

It is well known that gravity theory performs well in predicting gross trade flows at industry level. Thus it may not be surprising that it also performs well in predicting net trade at industry level. However, as Debaere (2005) emphasizes, a large part of the good empirical performance of gravity equations comes from an accounting relationship that total export must be equal to total output. Another advantage in our approach is that it eliminates from the estimation equation the influence of this accounting relationship. We regress the ratio of net trade over gross trade on a normalized difference in industry share, and both variables are free from the influence of production scale. Under this stringent specification, we find that industrial structure is significant in predicting trade balance. However, we also find that industrial structure explains only a small portion of total variation in trade balance. Therefore our finding is somewhat mixed. The result lends support to the view of economists that trade imbalance at industry level is the realization of comparative advantages and hence it should be fostered by free trade. However, the result also suggests that the view is partly valid and trade imbalance is largely determined by factors other than production specialization. To draw a reliable conclusion, we need further research that shows what these other variables are and how these variables are affected by policy changes like FTAs.

The question about the endogeneity of production structure raises another caveat necessary for interpreting our paper. Industrial structure is treated as an exogenous variable in many empirical studies on trade. However, from a broader perspective, we should acknowledge that industrial structure and trade pattern are jointly determined through the long-term influence of resources, geography and policies. It is difficult to find good instruments for the difference in industry share. It also is tricky to adopt a dynamic estimation strategy because our data contain a lot of discontinuities. Nevertheless, we will have to find out a way to tackle these problems in a future study.

References

Anderson, J. E. and E. van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, vol. 93, no. 1, pp. 170-192.

Baldwin, R. and D. Taglioni. 2006. "Gravity for Dummies and Dummies for Gravity Equations," Working Paper 12516, National Bureau of Economic Research.

Bernard, A. B., S. J. Redding and P. K. Schott. 2007. "Comparative Advantages and Heterogeneous Firms," *Review of Economic Studies,* vol. 74, no. 1, pp. 31-66.

Caves, R. E. 1981. "Intra-Industry Trade and Market Structure in the Industrial Countries," *Oxford Economic Papers,* vol. 33, issue 2, pp. 203-223.

Chaney, T. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review*, vol. 98, no. 4, pp. 1707-1721.

Chung, et. al. 2008. *Empirical Analysis of Factors Affecting Korea's Trade Structure and Current Account*, in Korean, Research Report 08-07, Korea Institute for International Economic Policy: Seoul. (in Korean)

Davis, D. R. and D. E. Weinstein. 2002. "The Mystery of the Excess Trade (Balances)," *American Economic Review, vol.* 92, no. 2, pp. 170-174.

Debaere, P. 2005. "Monopolistic Competition and Trade, Revisited: Testing the Model without Testing for Gravity," *Journal of International Economics,* vol. 66, issue 1, pp. 249-266.

Evans, C. L. 2003. "The Economic Significance of National Border Effects," *American Economic Review,* vol. 93, no. 4, pp. 1291-1312.

Feenstra, R. C., J. A. Markusen, and Andrew K. Rose. 2001. "Using the Gravity Equation to Differentiate among the Alternative. Theories of Trade," *Canadian Journal of Economics*, vol. 34, no. 2, pp. 430-447.

Greenaway, D., R. Hine, and C. Milner. 1995. "Vertical and Horizontal Intra-Industry Trade: A Cross Industry Analysis for the United Kingdom," *Economic Journal,* vol. 105, no. 433, pp. 1505-1518.

Harrigan, J. 1996. "Openness to Trade in Manufactures in the OECD," *Journal of International Economics,* vol. 40, issue 1-2, pp. 23-39.

Helpman, E. 1987. "Imperfect Competition and International Trade: Evidence from fourteen Industrial Countries," *Journal of the Japanese and International Economies,* vol. 1, issue 1, pp. 62-81.

Hummels, D. and J. Levinsohn. 1995. "Monopolistic Competition and International Trade: Reconsidering the Evidence," *Quarterly Journal of Economics,* vol. 110, no. 3, pp. 799-836.

Melitz, M. J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica,* vol. 71, no. 6, pp. 1695-1725.

Nicita, A. and M. Olarreaga. 2006. *Trade and Production 1976-2004*, www.worldbank.org/trade. (Accessed on February 15, 2012)

Rauch, J. E. 1999. 'Networks versus Markets in International Trade', *Journal of International Economics,* vol. 48, issue 1, pp. 7-35.

Rose, A. K. 2004. "Do WTO Members Have More Liberal Trade Policy?," *Journal of International Economics,* vol. 63, issue 2, pp. 209-235.
Song, E. Y. 2011. "On Gravity, Specialization and Intra-industry Trade," *Review of International Economics*, vol. 19, issue 3, pp. 494-508.
Song, E. Y. and C. H. Sohn. 2012. "Intra-industry Trade and Industry Distribution of Productivity: A Cournot-Ricardo Approach," *World Economy*, vol. 35, issue 4, pp. 461-482.
Toh, K. 1982. "A Cross-Section Analysis of Intra-Industry Trade in U.S. Manufacturing Industries," *Review of World Economics*(*Weltwirtschaftliches Archiv)*, vol. 118, issue 2, pp. 281-301.
Zhao, C. 2009. "Does Specialization Matter in Determining Trade Imbalance at Industry Level?" Master Thesis, Sogang University.

About the Author

E. Young Song is Professor at the Department of Economics, Sogang University. He holds a Ph.D. in Economics from Harvard University. He was an assistant professor at Vanderbilt University, USA. He specializes in trade theory, especially dynamic trade models and trade pattern theory, and has published papers in International Economic Review, Journal of International Economics, Canadian Journal of Economics, Journal of Economic Dynamics and Control.

Chen Zhao is a Ph.D. student at the Department of Economics, Hong Kong University of Science and Technology. She did her master work at the Department of Economics, Sogang University. Her major research interests are in trade theory and empirical studies of international trade. She currently is working on the effects of credit allocation on exports in China.

First version received on 7 September 2012
Peer-reviewed version received on 19 September 2012
Final version accepted on 27 September 2012