Determinants of Intra-Industry Trade: An Investigation with BMA for the European Union
Krzysztof Beck
Lazarski University, Poland

ABSTRACT

The empirical literature on determinants of intra-industry trade (IIT) is vast and comprehensive, yet as the authors failed to properly account for model uncertainty it has brought inconsistent and conflicting results. To resolve this issue, Bayesian model averaging was applied to investigate the robustness of 48 potential determinants of bilateral IIT for the panel of 26 European Union countries over the 1999-2011 period. Application of BMA demonstrated that 11 of them are robust determinants of IIT, namely real GDP product, trade openness, membership in the European Union and the Euro area, corruption, and differences in factor abundance. Among the factors of production, the key role in the determination of IIT patterns can be assigned to the differences in human capital. Yet, transportation cost and cultural similarity have no impact on the IIT patterns.

1. Introduction
Intra-industry trade (IIT) is at the heart of European integration. A high degree of IIT promotes more symmetrical distribution of economic shocks and, in turn, tighter business cycle synchronization. Only when business cycles of a given group of countries are fairly synchronized, can the countries give away their independent monetary and exchange rate policy to a supranational institution in order to constitute an effectively functioning currency union. For these reasons, IIT should be one of the main concerns of the Euro area current as well as potential members. For many years the authors have been trying to establish what factors are the determinants of intra-industry trade (e.g., Sharma, 2004; Thorpe and Zhang, 2005; Zhang and Clark, 2009; Jensen and Lüthe, 2009; Sawyer et al., 2010; Dautovic et al., 2014), yet none of them took into account model uncertainty, which leads to many conflicting and inconsistent results. For this reason, this paper presents the results of the sensitivity analysis of the determinants of IIT with Bayesian model averaging (BMA). Thus far, there has been only one attempt at sensitivity analysis of determinants of IIT (Torstensson, 1996). Using a rather outdated (by current standards) methodology, the author analyzed 17 determinants of IIT on the Swedish data and found that only physical capital intensity and transportation cost are robust. The paper is organized as follows. The first two sections describe data and estimation strategy, while the other two present results and conclusions.

2. Data and Measurement
The analysis covers 26 European Union countries, namely: Austria, Belgium, Bulgaria, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxemburg, Netherland, Poland, Portugal, Romania, the Slovak Republic, Slovenia, Spain, Sweden, and the United Kingdom. All variables are in bilateral form – for 26 countries, it amounts to 325 country pairs. The time span of the research covers the period between 1999 and 2011. As most of the variables used in the research are characterized by unit root, first differences were used. Consequently, the balanced panel consists of 3900 observations.

The dependent variable is the first difference of the natural logarithm of intra-industry trade measured by the sum of imports and exports of intermediate goods within the same sector between a pair of countries in each year. IIT is measured with the division of the economy into 35 different sectors and the data for IIT comes from World Output-Input Database (WOID). The list of the independent variables along with their descriptions and data sources are given in Table 1. The set of regressors is made up of variables describing differences in absolute and relative factor endowments, gravity variables, structural similarity, as well as macroeconomic, institutional and cultural indicators. In total, the set amounts to 48 regressors.

3. Estimation Strategy
To find a set of robust determinants of intra-industry trade, Bayesian model averaging (BMA) under different prior specification was applied. A detailed description of BMA (Hoeting, et al., 1999; Beck, 2017) and prior structure (Fernández, et al., 2001; Ley and Steel, 2009 and 2012, Feldkircher and Zeugner, 2009, Eicher, et al., 2011) is left for references. The particular estimation strategy, customized for the problem at hand, along with the key BMA statistics is described in this subsection. As a high degree of
multicollinearity among the regressors is possible, an appropriate prior structure has been employed to deal with this issue.

Table 1: Data description

| Source | Description | Variable |
|--------|-------------|----------|
| A       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| B       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| C       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| D       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| E       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| F       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| G       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| H       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| I       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| J       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| K       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| L       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| M       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| N       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| O       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| P       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| Q       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| R       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| S       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| T       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| U       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |
| V       | A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities: | $P(M_j)$, $|R_j|$, $\beta_i,j$ |

Source: Author’s own elaboration

A uniform model prior (Ley and Steel, 2009) is supplemented with a function accounting for multicollinearity (George, 2010) to obtain prior model probabilities:

$$P(M_j) \propto |R_j|^0.5 \cdot \left( \frac{1}{2} \right)^K,$$

where $K = (48)$ is the number of covariates, while $|R_j|$ is the determinant of the correlation matrix for all the regressors in the model $j$. The uniform model prior implies equal probabilities assigned to all the models ($|M_j| \propto 1$), so the $|R_j|$ component of (2) determines the distribution of the prior probability mass. The higher the multicollinearity between the variables, the closer the value of $|R_j|$ to 0, and the lower the prior ascribed to a given model. Model space is reduced with MC³ (Markov Chain Monte Carlo model composition) sampler (Madigan et al., 1995). The convergence of the chain is assessed by the correlation coefficient between the analytical and MC³ posterior model probabilities for the best 10000 models. The application of BMA requires the specification of $g$ prior. Benchmark prior rule (Fernández, et al., 2001) dictated the choice of unit information prior (UIP) (Kass and Wasserman, 1995) for the dataset at hand. Additionally, risk inflation criterion (RIC) put forward by Foster and George (1994) was employed in the main results. The combination of prior model probabilities with the values of likelihood function allows to calculate posterior model probability as:

$$P(M_j) = \frac{P(M_j) \cdot p(M_j)}{\sum_j \left( P(M_j) \cdot p(M_j) \right)} (2)$$

Where $(y|M_j)$ denotes model specific marginal likelihood, $y$ given data set, and because $p(y) = \sum_j P(y|M_j) = p(M_j)$, model weights can be treated as probabilities. Then the posterior mean $(PM)$ of the coefficient $\beta_i$ independent of the space of the models is given by:

$$PM = E(\beta_i | y) = \sum_j P(M_j | y) \cdot \beta_i,j \quad (3)$$

is the value of the estimated with OLS for the model $M_j$. The posterior standard deviation (PSD) is equal to:

$$PSD = \sqrt{\sum_{j=1}^{2K} P(M_j | y) \cdot \sum_{j=1}^{2K} P(M_j | y) \cdot \beta_i,j \cdot \beta_i,j} \quad (4)$$

where $V(\beta_i | y, M_j)$ denotes the conditional variance of the parameter for the model $M_j$. To better capture the relative impact of the determinants on the intra-industry trade standardized coefficients were calculated and BMA statistics based on their values. SPM denotes the standardized posterior mean, while SPDS denotes standardized posterior standard deviation (Doppelhofer and Weeks, 2009). The posterior probability of including the variable in the model – posterior inclusion probability (PIP) – is calculated as:

$$PIP = P(x_i | y) = \sum_{j=1}^{2K} (1)$$

where $\varphi_i = 1$ signifies including the variable $x_i$ in the model. In all applications of BMA here, prior inclusion probability is 0.5, and a variable is classified as robust if PIP is above that value. The posterior probability of a positive sign of the coefficient in the model – $(+)$ – is calculated in the following way:

$$P(+) = P(\varphi_i = 1 | y, M_j) \cdot P(M_j | y) \cdot (5)$$

where $\varphi_i = 1$ signifies including the variable $x_i$ in the model.
where \( CDF \) denotes cumulative distribution function, while 
\[

t_\alpha = (\beta / SD) | M \).
\]

4. Results

The results of the application of BMA to the data set are depicted in Table 2. BMA specification included 0.1 million burn-ins and 1 million iterations, which resulted in correlation coefficient between the analytical and MC posterior model probabilities for the best 10000 models above 0.999 in both cases, ergo the convergence of the chain was achieved. The first variable classified as robust is the natural logarithm of the real GDP product (\( \ln \text{RGDPPROD} \)) with PIP equal to one for both g prior specifications. The posterior mean indicates that one percent increase in \( \ln \text{RGDPPROD} \) is associated with an increase in IIT by approximately 0.69 percent. This confirms that gravity works in case of the IIT, just as in the instance of the total trade. On the other hand, geographical distance (DGE), border dummy (B), and common language dummy (L) are fragile, which suggests that transportation costs and cultural similarities are not important for the determination of the intra-industry trade. Change in the degree of the openness (OPEN) is the second variable with posterior inclusion probability higher than prior. OPEN is characterized by a positive posterior mean, which indicates that more open countries are generally more involved in IIT. This could be explained by the fact that more open countries are more integrated into global value chains, which can account for the significant part of the intra-industry trade in intermediate goods. The product of Bayesian Corruption Index (BCIPROD) is characterized by posterior inclusion probability equal to 1 under both g priors. The regressor is characterized by a negative posterior mean, which suggests that the higher degree of corruption is associated with lower IIT. This result should not come as a surprise, as companies moving part of their production abroad will try to avoid risk and additional costs associated with corrupt administration. Membership in the European Union (EU) as well as in the Euro area (MU) are classified as robust regardless of the used g prior specification. In a rather unexpected turn, values of the posterior mean of both EU and MU are negative, suggesting that countries in either of these associations are trading less with one another. In other words membership in the EU and the Eurozone is associated with lower intra-industry trade growth by around 5% a year. This result is the consequence of the analysis of the growth rates. Old members of the EU and the Euro area are characterized by higher levels of IIT. Application of OLS or BMA at level results in the positive values of coefficients or posterior mean respectively. This in turn indicates that intra-industry is growing between old EU members and the new member states, and it shows that these countries are consecutively more interlinked in European value chains. It should be underlined that dummy variable EU and MU takes the value of 1 only if both countries are members of the European Union or the Euro area respectively. The last six robust variables are differences in absolute or relative factor endowments. Four of them involve human capital, which indicates that this is the key factor of production for intra-industry trade. The absolute value of the difference in human capital to employment ratio (\( \text{HUMANEMPL} \)), human capital to land ratio (\( \text{HUMANLAND} \)), and human capital to physical capital ratio (\( \text{HUMANCAP} \)) are all characterized by the negative value of the posterior mean. The negative posterior mean indicates that countries characterized by a similar level of human capital to other factors ratios are engaged in more IIT. This suggests that traded goods are characterized by similar human capital intensity, and they must compete with the foreign counterparts. Accordingly, one can expect that human capital intensive goods account for the sizable part of intra-industry trade (product of human capital in two countries is positively and significantly correlated with IIT). Additionally, these results give support to theories associated with horizontal IIT (e.g., Krugman, 1981). On the other hand, the absolute value of the difference in human capital to arable land (\( \text{HUMANARABLE} \)) is characterized by a positive posterior mean. This result points to vertical integration, where human capital abundant countries are moving low skilled parts of the value chains to countries with abundant natural resources and arable land. Alternatively, the negative posterior mean can be explained by theories associated with vertical IIT (e.g., Flam and Helpman, 1987). Differences in absolute levels of employment (\( \text{EMPL} \)) are characterized by posterior inclusion probability of 0.94 and 0.96 for UIP and RIC respectively. A positive posterior mean indicates that higher differences in the levels of employment are related to higher growth rates of IIT. This result can be attributed to vertical integration, as countries’ scare labor force might search for the location of labor intensive parts of the value chain in the countries that are labor abundant. Finally, differences in arable land per worker (\( \text{ARABLE} \)) are the last variable classified as a robust determinant of IIT. It is characterized by a negative posterior mean, which implies that differences in arable land to labor ratio are deteriorating intra-industry trade. Accordingly, one can expect that arable land abundant countries will engage in trade in agrarian products, but depending on their particular location the exact nature of the products will be different.

Table 2: BMA statistics under UIP and RIC

| \( g \) prior | Unit Information Score | Risk Inflation Criterion |
|-------------|------------------------|-------------------------|
| \( \text{HUMANEMPL} \) | 1.00 | 1.00 |
| \( \text{BCIPROD} \) | 1.00 | 1.00 |
| \( \text{HUMANLAND} \) | 0.00 | 0.00 |
| \( \text{MU} \) | 0.00 | 0.00 |
| \( \text{DGE} \) | 0.00 | 0.00 |
| \( \text{ARABLE} \) | 1.00 | 1.00 |
| \( \text{BCIPROD} \) | 0.00 | 0.00 |
| \( \text{HUMANLAND} \) | 0.00 | 0.00 |
| \( \text{MU} \) | 0.00 | 0.00 |
| \( \text{DGE} \) | 0.00 | 0.00 |
| \( \text{ARABLE} \) | 1.00 | 1.00 |
| \( \text{BCIPROD} \) | 0.00 | 0.00 |
| \( \text{HUMANLAND} \) | 0.00 | 0.00 |
| \( \text{MU} \) | 0.00 | 0.00 |
| \( \text{DGE} \) | 0.00 | 0.00 |
| \( \text{ARABLE} \) | 1.00 | 1.00 |

Source: Author’s own elaboration
Turning to standardized posterior means, the product of GDPs and differences in human capital to employment ratios have the strongest impact on the growth rate of IIT. Second in line are the degree of openness and differences in human capital to land ratios followed by product of Bayesian Corruption Index. Next in line are memberships in the Euro area and the European Union succeeded by differences in employment, the differences in human capital to arable land and human capital to physical capital ratios. Finally, differences in arable land per worker have the lowest impact on intra-industry trade among all the robust variables. All the above-mentioned results turned out to be robust to manipulation in g prior and model prior specification. Additional robustness checks are not reported here for brevity, but are available upon request.

5. Conclusion
Application of BMA to the panel data for 26 European Union countries over the 1999-2011 period allows for the identification of 11 robust determinants of intra-industry trade. Out of traditional gravity model variables, only product of real GDP turned out to be robust, while transportation cost and cultural similarity proxies are classified as fragile. Corruption seems to provide strong impairment on IIT as additional risk of product of real GDP. Memberships in the EU and the Eurozone have a positive impact on the level of IIT; but negative on the growth rate of intra-industry trade. This result suggests that new member states are getting more entangled in old EU value chains. Finally, the analysis showed the crucial role of the differences in the factor abundance of the trading countries. Out of all the analyzed factors of production, the most important part in the determination of intra-industry trade patterns is played by human capital.

References

- Beck, K. (2017). Bayesian Model Averaging and Jointness Measures: Theoretical Framework and Application to the Gravity Model of Trade. STATISTICS IN TRANSITION new series, 18(3), 393-412. Crossref
- Dautovic E., Orszaghova L., Schudel W. (2014). Intra-industry Trade between CESEE countries and the EU15. ECB Working Paper.
- Doppelhofer, , Weeks, M. (2009). Jointness of Growth Determinants. Journal of Applied Econometrics, 24(2), 209-244. Crossref
- Eicher, T., Papageorgiou, C., Raftery, A. E. (2011). Determining Growth Determinants: Default Priors and Predictive Performance in Bayesian Model Averaging. Journal of Applied Econometrics, 26(1), 30-55. Crossref
- Feenstra, R.C., Inklaar, R., Timmer, M.P. (2015). The Next Generation of the Penn World Table. American Economic Review, 105(10), 3150-3182. Crossref
- Feldkircher, M., Zeugner, S. (2009). Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging. IMF Working Paper. Crossref
- Fernández, C., Ley, E., Steel, M. (2001). Benchmark priors for Bayesian model Journal of Econometrics, 100(2), 381-427. Crossref
- Flam, H., Helpman, E. (1987), Vertical Product Differentiation and North-South Trade. American Economic Review, 77(5), 810-822.
- Foster, D., George, E. (1994). The Risk Inflation Criterion for Multiple Regression. The Annals of Statistics, 22(4), 1947-1975. Crossref
- George, , Foster, D. (2000). Calibration and empirical Bayes variable selection. Biometrika, 87(4), 731-747. Crossref
- Hoeting, J., Madigan, D., Raftery, A., Volinsky, C. (1999). Bayesian Model Averaging: A Tutorial. Statistical Science, 14(4), 382-417. Crossref
- Jensen, , Lüthje, T. (2009). Driving forces of vertical intra-industry trade in Europe 1996- 2005. Review of World Economics, 145, 469-488. Crossref
- Kass, , Wasserman, L. (1995). A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwarz Criterion. Journal of the American Statistical Association, 90(431), 928-934. Crossref
- Krugman, P. (1981). Intraindustry Specialization and the Gains from Trade. Journal of Political Economy, 89(5), 959-973. Crossref
- Ley, , Steel, M. (2009). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. Journal of Applied Econometrics, 24(4), 651-674. Crossref
- Ley, , Steel, M. (2012). Mixtures of g-priors for Bayesian model averaging with economic applications. Journal of Econometrics, 171(2), 251-266. Crossref
- Sayer, W.C., Sprinkle R.L., Tochkov K. (2010). Pattern and determinants of intra-industry trade in Asia. Journal of Asian Economics, 21(5), 485-493. Crossref
- Sharma, (2004). Horizontal and vertical intra-industry trade in Australian manufacturing: does trade liberalization have any impact? Applied Economics, 36(15), 1723-1730. Crossref
- Standaert, (2015). Divining the level of corruption: a bayesian state-space approach. Journal of Comparative Economics, 43(3), 782-803. Crossref
- Thorpe, M., Zhang, Z. (2005). Study of the Measurement and Determinants of Intra- industry Trade in East Asia. Asian Economic Journal, 19(2), 231-247. Crossref
- Torstensson, J. (1996). Determinants of Intra-industry Trade: A Sensitivity Analysis. Oxford Bulletin of Economics and Statistics, 58(3), 507-524.
- Zhang, Y., Clark D. P. (2009). Pattern and Determinants of United States’ Intra-industry trade. International Trade Journal, 23(3), 325-356. Crossref