THE EFFECT OF INCOME INEQUALITY IN THE DETERMINATION OF FREE TRADE AGREEMENTS

İlkay YILMAZ
Mersin Üniversitesi

ABSTRACT

This paper tests the relationship between income inequality and the possibility of free trade agreements (FTAs). The model is based upon Baier and Bergstrand (2004). Two new explanatory variables are added to their model. These variables measure income inequality and the level of democracy in country pairs. It is found that the potential welfare gains and likelihood of an FTA between a pair of countries is higher, the more (less) egalitarian the income distribution in the relatively capital (labor) abundant country of the pair is, if the country is democratic.

Key words: Free trade agreements; International trade; Median-voter

1. Introduction

The goal of this article is to empirically investigate whether income inequalities in country pairs could be one of the key economic factors influencing the possibility of a Free Trade Agreement (FTA) between the two countries.

The median-voter approach to trade policy determination as in Mayer in the Heckscher-Ohlin framework indicates that an increase in inequality in a capital-abundant (labor-abundant) economy raises (decreases) trade barriers.[1] Dutt and Mitra find support for this prediction using cross-country data on inequality, capital-abundance and diverse measures of protection.[2] For developing countries with lower capital-labor ratios, greater inequality leads to lower tariffs. Conversely, for industrialized countries with higher capital-labor ratios, greater inequality leads to higher tariffs. This provides support for the median voter framework in the context of the

[1] Wolfgang Mayer, 'Endogenous Tariff Formation,' *American Economic Review,* No. 74, 1984, pp. 970-985.
[2] P. Dutt and D. Mitra, 'Endogenous Trade Policy Through Majority Voting,' *Journal of International Economics,* No. 58, 2002, pp.107-134.

Heckscher-Ohlin model. In addition, Dutt and Mitra find that this relationship holds better in democracies than in dictatorships.

Although there is a large literature explaining empirically tariff and non-tariff barriers between countries, the first econometric work that tries to explain empirically the determinants of FTAs is a very recent one: Baier and Bergstrand.[3] We basically follow their work. Inequality variables (GINI coefficients in the country pairs) are added to their explanatory variables to see whether they make a difference. Their econometric model is based upon a general equilibrium theoretical model of world trade with two factors of production, two monopolistically competitive product markets, and explicit intercontinental and intracontinental transportation costs among multiple countries on multiple continents. They find that trade-creating and trade-diverting economic characteristics play an important role in explaining the probability of an FTA between two governments. According to their results, two economies tend to form FTAs: (i) the closer are two countries in distance; (ii) the more remote a pair of continental trading partners is from the rest of the world; (iii) the larger and more similar in economic size are two trading partners; (iv) the greater the difference of capital–labor ratios between two partners; and (v) the smaller the difference of the members' capital–labor ratios with respect to the ROW's capital–labor ratio. In their empirical model these characteristics correctly predict 85 percent of the 286 FTAs existing in 1996 among 1431 pairs of countries and 97 percent of the remaining 1145 pairs with no FTAs.

The contribution of this article is to include income inequality and the level of democracy in the analysis of the economic determinants of the likelihood of FTAs between country pairs. Although Dutt and Mitra find support for the prediction that an increase in inequality in a capital-abundant (labor-abundant) economy raises (decreases) trade barriers, their work does not say anything about FTAs.[4] On the other hand Baier and Bergstrand did not consider income inequality levels in their attempt to explain economic determinants of FTAs.[5]

The main finding of this article is that, when we consider two countries, a negative (positive) relationship exists between the income inequality level in the relatively capital-abundant (labor-abundant) country and the possibility of the FTA, if the country is democratic. A similar relationship in non-democratic countries is not found.

2. Econometric Model and Estimation Method

As we mentioned above, our model is based on Baier and Bergstrand's best probit result.[6] Therefore it includes their explanatory variables, which we will call B&B variables from now on, as well as inequality variables (gini coefficients) of the two countries in the pair and also interaction terms between these inequality variables and dummy variables which indicate whether the countries can be regarded as democratic:

$$P(FTA=1) = P(y^*>0) =$$

$$G(\beta_0 + \mathbf{x}\boldsymbol{\beta} + \delta_1 GINIP + \delta_2 GINIR + \delta_3 GINIP.DEMP + \delta_4 GINIR.DEMR) \quad (1)$$

where y* denotes the (unobservable) difference in utility levels from the formation of an FTA and

[3] S. L. Baier and J. H. Bergstrand, 'Economic Determinants of Free Trade Agreements', *Journal of International Economics,* No.64, 2004, pp. 29-63

[4] Dutt and Mitra, 'Endogenous…'

[5] Baier and Bergstrand, 'Economic…'

[6] *ibid*

$$y^* = \beta_0 + \mathbf{x\beta} + \delta_1 GINIP + \delta_2 GINIR + \delta_3 GINIP.DEMP + \delta_4 GINIR.DEMR + e. \quad (2)$$

It is assumed that e is independent of **x** (the vector of B&B variables), GINIP, GINIR, GINIP.DEMP and GINIR.DEMR and it has a standard normal distribution. Since both countries' consumers need to benefit from an FTA for their representative countries to form one, formally $y^* = \min(\Delta U_i, \Delta U_j)$.

The dependent variable FTA gets the value 1 if there exists an FTA between the two countries in 1996, which indicates $y^* > 0$, and 0 otherwise, which indicates $y^* \leq 0$. Here the standard normal cumulative distribution function G(.) ensures that P(FTA=1) is in (0,1).

Parameters $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6]'$ are the ones corresponding to the explanatory variables from Baier and Bergstrand.[7] The first one of these explanatory variables $NATURAL_{ij}$ measures the geographical closeness of i and j. It is the natural logarithm of the inverse of the distance between the economic centers i and j. The second one $REMOTE_{ij}$, on the other hand measures the remoteness of a pair of continental trading partners from the rest of the world. It takes the value 0 if the two countries are on different continents. However if they are in the same continent then $REMOTE_{ij}$ measures the simple average of the natural logarithms of the mean distance of country i from all of its trading partners except j and the mean distance of country j from all of its trading partners except i. While the third explanatory variable $RGDP_{ij}$ simply measures the sum of the logs of real GDPs of countries i and j in 1960, the fourth explanatory variable $DRGDP_{ij}$ measures the absolute value of the difference between the logs of real GDPs of countries i and j in 1960. The fifth explanatory variable DKL_{ij} measures the absolute value of the difference between the logs of the capital-labor ratios of countries i and j in 1960. The sixth explanatory variable $DROWKL_{ij}$ measures the difference between the capital-labor ratios of i and j and the rest of the world's capital-labor ratio. It is the simple average of two differences, which are between the natural logarithm of the combined capital-labor ratio of i(j)'s all trading partners and the natural logarithm of the capital-labor ratio of i(j).

The explanatory variable GINIP measures the income inequality in the relatively labor abundant country in the pair. Similarly GINIR measures the income inequality in the relatively capital abundant country in the pair. These are the averages of gini coefficients from the years 1960-69. The dummy variable DEMP takes the value 1 if the relatively labor abundant country in the pair is democratic in the years 1960-69, it takes the value 0 otherwise. Similarly DEMR takes the value 1 if the relatively capital-abundant country in the pair is democratic in the years 1960-69 and it takes the value 0 otherwise.

Since it is not possible to find reliable gini coefficients and democracy data from 1960s for all the countries in Baier and Bergstrand, the number of observations available for our model shrank to 406 from 1431 observations used in Baier and Bergstrand.[8] To see whether this shrinkage in data size causes any substantial change in the Baier and Bergstrand model we also recalculated their model, i.e.

$$P(FTA=1) = P(y^* > 0) = G(\beta_0 + \mathbf{x\beta}) \quad (3)$$

by using only the 406 observations available for our model.

As in Baier and Bergstrand, we use the maximum likelihood estimation (MLE) method to estimate the parameters of the model. As Wooldridge stated, the general theory of (conditional) MLE for random samples implies that, under very general conditions, the MLE is consistent, asymptotically normal, and asymptotically efficient.[9]

---

[7] *ibid*
[8] *ibid*
[9] J. M. Wooldridge, *Introductory Econometrics*, (Cincinnati: South-Western Publishing, 2000)

3. The data

The data for B&B variables (FTA_{ij}, $NATURAL_{ij}$, $REMOTE_{ij}$, $RGDP_{ij}$, $DRGDP_{ij}$, DKL_{ij}, $DROWKL_{ij}$) were taken from Dr. Baier. For further information about their sources, one can look at Bair and Bergstrand.[10] Here it should be emphasized once more that although FTA_{ij} shows whether the pair has an FTA in 1996, explanatory variables $RGDP_{ij}$, $DRGDP_{ij}$, DKL_{ij} and $DROWKL_{ij}$ are measurements related to 1960. This time difference between the dependent variable and these explanatory variables are due to potential endogeneity. It is explained by the following lines:

"… Since an FTA formed several years prior to 1996 likely influenced subsequent trade – which then influenced economic growth – incomes and capital stocks in 1996 may well be endogenous. To account for this, we used the earliest data on incomes and capital and labor stocks… for a wide sample, namely, 1960 data."[11]

The data for GINI coefficients are obtained from the UNU/WIDER – UNDP World Income Inequality Database (WIID) which can be downloaded from the UNU/WIDER web pages at

http://www.wider.unu.edu/wiid/wiid.htm .

To obtain a fairly reliable data subset of gini coefficients from this source, only those data points with OKIN ("*Reliable income or expenditure data referring to the entire [national] population, not affected by apparent inconsistencies*") quality rating were chosen.[12] Since very few countries had gini coefficient data for the year 1960 for this rating, averages of gini coefficients over the years from 1960 to 1969 are used. By this way, gini variables are obtained for 30 countries out of 54 in Baier and Bergstrand.[13]

The data for dummy variables DEMP and DEMR are constructed by using data from the Polity IV Project dataset. This dataset is easily available on the web at

www.cidcm.umd.edu/inscr/polity.

The indicator "POLITY" in the dataset ranges from -10 (full autocracy) to +10 (full democracy). For each country average of POLITY scores from the years 1960-69 are used. Those countries with positive average POLITY scores are regarded as democratic and others with negative average POLITY scores are regarded as nondemocratic. This method allowed us to construct democracy dummy variables for the 29 countries out of the 30 countries with gini variables. Table 1 gives GINI and POLITY averages of the 29 countries used in this article.

Subsequently GINIP, GINIR, DEMP and DEMR variables are constructed for 406 pairs out of 1431 pairs in Baier and Bergstrand, by using the data available for the 29 countries.[14] To determine the labor abundant "poor" and the capital abundant "rich" in each pair, capital-labor ratios for the year 1960 are compared and the country with a higher capital-labor ratio is labeled as "rich" and the other as "poor". If the average POLITY score for the relatively labor abundant country in the pair is positive (negative), then the variable DEMP is unity (zero). Similarly, if the average POLITY score for the relatively capital abundant country in the pair is positive (negative), the variable DEMR is unity (zero).

---

[10] Baier and Bergstrand, 'Economic…'
[11] *ibid,* p.40
[12] UNU/WIDER-UNDP, *World Income Inequality Database User Guide and Data Sources*, Version 1.0, 2000, p.10. Available from: http://www.wider.unu.edu/wiid/wiid.htm [Accessed April 2004].
[13] Baier and Bergstrand, 'Economic…'
[14] *ibid*

On Table 2 it is no surprise that minimum and maximum values of GINIP and GINIR variables are the same. This is due to the fact that 27 out of the 29 countries used in the formation of the 406 country pairs are labeled “poor” in some country pairs and “rich” in others. Since we are considering relative country pairs, a given country is labeled “poor” in a pair if the other country in the pair has higher 1960 capital-labor ratio and it is labeled “rich” if the other country has lower 1960 capital-labor ratio. The two countries that do not change their status of being “poor” or “rich” in all pairs are Thailand and Australia. Thailand has the lowest 1960 capital-labor ratio whereas Australia has the highest among the 29 countries.

Table 1:
GINI and POLITY averages for 1960s

		GINI average (1960s), %	POLITY average (1960s)
1	Argentina	42.0	-4.2
2	Australia	32.5	10.0
3	Bolivia	53.0	-3.6
4	Brazil	53.5	-2.9
5	Canada	31.8	10.0
6	Chile	37.7	5.6
7	Columbia	62.0	7.0
8	Costa Rica	50.0	10.0
9	Denmark	37.0	10.0
10	Ecuador	38.0	0.5
11	El Salvador	53.0	-1.2
12	France	48.3	5.3
13	Germany	45.0	10.0
14	Honduras	61.9	-1.0
15	Japan	35.7	10.0
16	Mexico	54.2	-6.0
17	Netherlands	42.0	10.0
18	Norway	35.0	10.0
19	Panama	48.0	1.8
20	Peru	61.0	2.3
21	Philippines	50.4	4.7
22	South Korea	32.2	1.5
23	Spain	32.0	-7.0
24	Sweden	37.9	10.0
25	Thailand	42.3	-6.0
26	Turkey	56.0	8.4
27	United Kingdom	32.8	10.0
28	United States	34.7	10.0
29	Venezuela	42.0	6.4

Table 2: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
FTA	0.1749	0.3803	0	1
NATURAL	-8.5186	0.8000	-9.6086	-5.0752
REMOTE	1.8652	3.5769	0	9.1274
RGDP	34.9953	2.3018	28.8239	41.0509
DRGDP	1.9203	1.4122	0.0071	6.9436
DKL	1.0145	0.7158	0.0076	2.8312
DROWDKL	0.8545	0.2780	0.1491	1.6893
GINIP	46.8700	9.5700	31.8	62
GINIR	41.5507	8.7985	31.8	62
DEMP	0.6650	0.4726	0	1
DEMR	0.8350	0.3717	0	1
GINIP.DEMP	30.3402	22.9238	0	62
GINIR.DEMR	33.7943	16.8427	0	62

Number of observations: 406

4. Results

Probit results indicate that the smaller data set of my model does not cause an important distortion in the calculations of B&B coefficients. In Table 3, the first column gives the results from Baier and Bergstrand.[15] The coefficient estimates of the same model calculated with the smaller data set are given in the second column (2a). For each explanatory variable coefficient estimates from both columns have the same sign and all the coefficient estimates except the one for DROWKL in the second column are statistically significant at 5% level. The coefficient estimates of the same model without DROWKL are presented in the column 2b, where all variables are statistically significant at 1% level.

The estimated coefficients of the model with gini coefficients and democracy dummies from the column 3a show that the variables GINIP and GINIR are statistically insignificant although interaction terms GINIP.DEMP and GINIR.DEMR are statistically significant at 5% level with expected signs. This indicates that income inequality has an effect on the formation of FTAs only in democratic countries. Also once again we see that the variable DROWKL is insignificant at 5% level, although all other B&B variables are statistically significant at 1% level. Therefore in column 3b the version without DROWKL is presented. Taking DROWKL out of the regression does not have any effect on the signs of the coefficient estimates. It only makes the variable DKL statistically significant at 5% level instead of at 1% level. The variables GINIP and GINIR stay statistically insignificant at 5% level.

Therefore another probit specification which includes only B&B variables and the interaction terms GINIP.DEMP and GINIR.DEMR is estimated with the 406 pairs and presented in column 4a of Table 3. The coefficient estimates of the both interaction terms have expected signs and they are statistically significant. The coefficient estimate of the interaction term GINIP.DEMP is positive and it is statistically significant at 1% level. The coefficient estimate of the other interaction term GINIR.DEMR is negative and it is statistically significant at 5% level. Also, in column 4b the version without DROWKL is presented. Taking DROWKL out of the

[15] Baier and Bergstrand, 'Economic…'

regression does not have any effect on the signs of coefficient estimates or on their statistical significances.

The fact that the estimated coefficients of GINIP.DEMP and GINIR.DEMR (interaction terms of the average gini variable and the democracy dummy variable in the "poor" and the "rich" countries respectively) are both statistically significant with the expected signs also shows its effect on the goodness-of-fit measure percent correctly predicted. The probit estimate of the model with B&B variables and the interaction terms GINIP.DEMP and GINIR.DEMR (column 4a in Table 3) correctly predicts 81.69 percent of the 71 FTAs, and 97.01 percent of the remaining 335 pairs with no FTAs while the probit estimate of the model only with B&B variables (column 2a in Table 3) correctly predicts 77.46 percent of the 71 FTAs and 96.72 percent of the remaining 335 pairs with no FTAs. Comparisons of the goodness-of-fit measure percent correctly predicted of the models without DROWKL (the variable that measures the difference between the capital-labor ratios of the two countries in the pair and the rest of the world's capital-labor ratio) also shows a similar picture. The probit estimate of the model with B&B variables without DROWKL, and with the interaction terms GINIP.DEMP and GINIR.DEMR (column 4b in Table 3) correctly predicts 83.10 percent of the 71 FTAs, and 97.01 percent of the remaining 335 pairs with no FTAs while the probit estimate of the model only with B&B variables without DROWKL (column 2b in Table 3) correctly predicts 74.65 percent of the 71 FTAs and 97.61 percent of the remaining 335 pairs with no FTAs.

Table 3: Probit Results for the Probability of an FTA

Variable	1 (B&B)	2a	2b	3a	3b	4a	4b
CONSTANT	7.90	6.66	4.46	3.62	2.36	6.52	5.32
	(4.92)*	(2.68)*	(2.15)**	(1.07)	(0.75)	(2.43)**	(2.21)**
	[5.40]*	[3.43]*	[3.04]*	[1.48]	[1.01]	[3.00]*	[2.85]*
NATURAL	1.76	1.53	1.52	1.79	1.82	1.74	1.76
	(13.43)*	(6.41)*	(6.57)*	(6.22)*	(6.39)*	(6.38)*	(6.54)*
	[12.05]*	[6.63]*	[7.02]*	[7.34]*	[7.51]*	[6.92]*	[7.12]*
REMOTE	0.18	0.18	0.18	0.21	0.21	0.20	0.20
	(10.03)*	(5.77)*	(5.80)*	(5.78)*	(5.80)*	(5.74)*	(5.76)*
	[10.04]*	[5.60]*	[5.53]*	[5.50]*	[5.75]*	[5.50]*	[5.67]*
RGDP	0.17	0.15	0.19	0.25	0.28	0.20	0.22
	(3.67)*	(2.20)**	(3.06)*	(2.89)*	(3.42)*	(2.66)*	(3.23)*
	[4.53]*	[2.58]*	[3.68]*	[3.76]*	[4.44]*	[3.11]*	[3.88]*
DRGDP	-0.34	-0.45	-0.43	-0.59	-0.59	-0.57	-0.56
	(-5.45)*	(-4.15)*	(-4.19)*	(-4.47)*	(-4.54)*	(-4.46)*	(-4.52)*
	[-5.46]*	[-3.70]*	[-3.79]*	[-4.55]*	[-4.48]*	[-4.67]*	[-4.92]*
DKL	0.85	0.59	0.44	0.70	0.64	0.65	0.59
	(7.37)*	(2.73)*	(2.35)*	(2.69)*	(2.58)**	(2.84)*	(2.70)*
	[6.74]*	[2.61]*	[2.20]**	[3.11]*	[3.03]*	[2.96]*	[2.82]*
DROWKL	-1.29	-1.00		-0.68		-0.66	
	(-5.53)*	(-1.83)		(-1.16)		(-1.14)	
	[-4.91]*	[-2.22]**		[-1.24]		[-1.23]	
GINIP				0.02	0.02		
				(1.00)	(0.93)		
				[1.10]	[1.02]		
GINIR				0.02	0.02		
				(1.07)	(1.12)		
				[1.07]	[1.13]		
GINIP.DEMP				0.02	0.02	0.02	0.02
				(2.42)**	(2.50)**	(2.70)*	(2.80)*
				[2.61]*	[2.66]*	[3.02]*	[3.08]*
GINIR.DEMR				-0.02	-0.02	-0.02	-0.02
				(-2.12)**	(-2.42)**	(-2.09)**	(-2.38)**
				[-2.11]**	[-2.41]**	(-2.05)**	[-2.34]**
Pseudo Rsq	0.728	0.665	0.655	0.707	0.703	0.700	0.697
Log likelihood	-194.4	-63.12	-64.86	-55.21	-55.89	-56.39	-57.04
# of observations	1431	406	406	406	406	406	406

Notes: The quantities in parentheses below the estimates are z-statistics. The quantities in brackets are robust z-statistics. * and ** denote statistically significant z-statistics at 1% and 5% levels in two-tailed test, respectively.

5. Conclusion

According to the median voter approach to trade policy determination within a Heckscher-Ohlin framework, while an increase in inequality in a capital-abundant country raises trade barriers, an increase in inequality in a labor-abundant country decreases them. The purpose of this study was to test the median voter approach to trade policy determination within the context of Free Trade Agreements (FTAs). We tried to find evidence of the effect of income inequalities on the formation of FTAs. The data show that the potential welfare gains and likelihood of an FTA between a pair of countries is higher, the more egalitarian the income distribution in the relatively capital abundant country of the pair is, if the country is democratic. Similarly the potential welfare gains and likelihood of an FTA between a pair of countries is higher, the less egalitarian the income distribution in the relatively labor abundant country of the pair is, if the country is democratic. We failed to find similar results in nondemocratic countries, which might be due to the fact that majoritarian concerns are much less important in dictatorships than in democracies. Our results about democracies are completely in line with the median voter approach.