Région et Développement
n° 52-2020

www.regionetdeveloppement.org

The Relevance of Regional Integration in the Euromed Area: Evidence from a Logistic Median-Voter Model

Javad ABEDINI[*]
Nicolas PÉRIDY[**]

Abstract - Regional integration in Europe and more generally into the Euro-Mediterranean (Euromed) area has recently been debated and increasingly questioned. The Median-Voter approach provides a methodical way to answering such questions through comparing the relative utility of the median voters before and after the integration. In this study, we suggest a conditional fixed-effect logit model, which relates such a relative utility to some observables, such as differences of countries in factor endowments and technology, to estimate the likelihood of Free Trade Agreement (FTA) creation across the entire Euromed area. Including all FTA partner countries in the world between 1981 and 2019, we estimate the likelihood of FTA creation for all country pairs as well as at country and regional level in the Euromed area. Receiver Operating Characteristic (ROC) metrics shows excellent performance of our model in truly predicting the (past) FTA and No-FTA events. Similarly, a Reverse Backtesting procedure is developed to attribute an FTA realization, from "impossible" to "highly probable", to each estimated probability. With an estimated probability equal to 82.6%, the European integration is qualified as a probable FTA, while the Non-European-Med integration (46.6%) and the overall Euromed integration (60.5%) are perceived as just "possible FTAs".

JEL Classification
F12, F14, F15, F17

Key-words
Regional integration,
Euro-Mediterranean area
European Union
Logistic Median-Voter Model
Conditional fixed effect logit model

---

[*] LinkedIn: https://www.linkedin.com/in/javadabedini ; javad.abedini@outlook.com
[**] Université de Toulon, LEAD, France ; nicolas.peridy@univ-tln.fr

INTRODUCTION

Regional integration in Europe and more generally into the Euro-Mediterranean area (hereafter, Euromed) constitutes one of the most important questions at the regional but also at the world political and economic scene. The Brexit is the most evident example but more generally, the public opinion tends to increasingly doubt about the relevance of European integration. Paradoxically, the Euromed area is the biggest regional area in the world but it seems to fail to find a permanent legitimacy.

There is an extensive literature dedicated to the assessment of the effects of regional integration in this area, especially its trade effects: Péridy and Roux (2012), Pehlivan (2014), Péridy (2015), Parra et al. (2016), Kahouli (2016). The recent DCFTA[1] agreements have also been partially addressed (ECORYS, 2013a and 2013b, Gasiorek, 2019). However, the relevance of such an enlarged area has not been really assessed. In other words, is the Euromed FTA natural and realistic from an economic point of view? Is there any inherent heterogeneity across this enlarged region for adhering to an economic integration?

Median voter and lobbying approach are two main approaches addressing the choice of Free Trade Agreement (FTA) across countries. The first one assumes that choices are guided by the majority rule, i.e., votes by the greatest part of the population (Mayer 1984). In contrast, the second approach emphasizes the role of interest groups in handling trade policies (Rodrik 1986). In other terms, the median voter approach counts on benefits that the overall economy makes from free trade while the lobbying approach takes only into account the benefits of some powerful special interest groups to decide on a free trade agreement.

Three reasons lead us to favor the median voter over the lobbying approach for a modelling purpose like ours in here. The first is the better connection of the median voter approach with the major international trade theories, e.g. the theory of comparative advantage or that of monopolistic competition, which generally back free trade based on the gains that it brings to the overall economy in partner countries. The second comes from the difficulty of modelling free trade under the lobbying approach because the true interests of special groups and the lobbying process itself are not usually publicized (or clearly known). Finally, the third reason originates from the fact that even a dictatorship cannot totally ignore public interests in its decision for free trade because of its fear of being overturned.

It is still possible to identify two kinds of approaches under the median voter stream of thinking. The first is based on a wide political-economy framework (Levy 1997) and the second relies on a more restricted Krugman-type monopolistic-competition model (Baier and Bergstrand 2004). Finally, Baier et al. (2014) fruitfully extend this approach by considering the effect of other FTAs on the probability of a pair having an FTA. This extension makes it possible to consider interdependence effects. These are both the "own FTA effect" (i.e., the impact on the likelihood of an FTA between two countries owing to either already having other FTAs) and the cross-FTA effect (the impact of other implanted FTAs in the rest of the world on the likelihood of a new FTA).

More recently, Larch et al. (2019) have showed that estimating based on an iterative Poisson pseudo-maximum likelihood (PPML) facilitates the inclusion of fixed effects in the gravity model for large data sets and also allows for correlated errors across countries and time.

In this original paper, we rely on the median voter approach in proposing an empirical model to identify the driving forces behind the binary choice of free trade in the Euromed area and to estimate the feasibility of such an integration at regional,

---

[1] Deep and Comprehensive Free Trade Agreement.

sub-regional and bilateral level. The model proposed here is an extension of the preliminary results published by Abedini and Péridy (2018). Our model relies on the political-economy framework, suggested by Levy (1997), while also testing for the recent developments such as the interdependence effects. In addition, using a likelihood-based logistic model allows us to reduce the errors in the estimation of fixed effects as suggested through the PPML structure by Larch et al. (2019).

Our database includes the partners of all effective free trade agreements in the world, i.e. 105 countries (10920 country-pairs), over 1981-2019, the largest possible data choice.

The presence of fixed effects leads us to use a conditional fixed effect logit model to consistently estimate the bilateral probabilities of forming an FTA between all Euromed country-pairs. Probabilities for sub-regional integration, and the Euromed area as a whole, are then obtained from income-weighted averages of the bilateral probabilities. We employ a Receiver Operating Characteristic (ROC) analysis, usually used in Medical sciences, to check the performance of our model in truly predicting FTA and No-FTA events. In addition, a Reverse Backtesting procedure is used to find a history-backed interpretation for the calculated probabilities. This allows us to relate each estimated probability to an FTA outcome, from being "impossible" to "highly probable". Accordingly, with an estimated probability equal to 60.5%, we qualify the whole Euromed area, as a possible integration. Similarly, the European (82.6%) and Non-European (46.6%) sub-regions are qualified for a probable and possible FTA respectively. In spite of these general figures, a great heterogeneity is found across the bilateral relationships.

This article proceeds as follows. Section 1 represents the median voter approach based on which our empirical model is derived later in Section 2. Section 3 discusses the estimation method and econometric tests. It also provides our calculation of FTA probabilities at bilateral and regional level within the Euromed area. Using a ROC analysis, section 4 evaluates the performance of our model in truly predicting FTA and No-FTA events. In turn, section 5 employs a reverse backtesting procedure to provide a history-backed interpretation of the results. Finally, section 6 concludes the current paper.

1. MEDIAN VOTER APPROACH[2]

In a differentiated-product framework, we assume that countries only differ in their factor endowment. In this framework, each agent *i* who also participates in the median voter system is assumed to own one unit of labor and a certain amount of capital (k_i) as follows:

$$\sum_{i=1}^{L} k_i = K \tag{1}$$

where L and K, respectively, reflect the total labor and capital available in the economy. Denoting w and r as the factor rewards (wages and interest rate), each agent receives an income equal to:

$$I_i = rk_i + w \tag{2}$$

Assuming the existence of two industries: Y (numeraire good) for homogenous products with constant returns to scale and X for differentiated products with a total number of n varieties called x and increasing returns to scale, production functions can be set as follows:

---

[2] As previously indicated, this section is inspired from Levy (1997).

$$Y = \gamma_Y K_Y^{\mu} L_Y^{1-\mu} \tag{3}$$

$$x = \gamma_x K_x^{\xi\eta} L_x^{\xi(1-\eta)} \tag{4}$$

where $0 < \mu$, $\eta < 1$ and $\xi > 1$. At equilibrium, X would be the sum of output for all n varieties in the economy.

Next, we consider that agents can be represented by identical utility functions:

$$U = U_X^{\alpha} Y^{1-\alpha} \tag{5}$$

$$U_x = \left(\sum_{i=1}^{n} D_i^{\beta}\right)^{\frac{1}{\beta}} \tag{6}$$

$$\beta = \left(1 - \frac{1}{\sigma}\right) \qquad \sigma > 1$$

where D is the consumption of variety x (i indexes the varieties) and σ is the constant cross-price elasticity of substitution between varieties that is commonly found in the Spence-Dixit-Stiglitz framework. In this case, each firm in industry X produces an identical optimal level variety x:

$$x = \frac{\sigma}{\xi(\sigma-1)}$$

that is sold at equilibrium price p. The number of firms in industry X is then equal to the number of varieties produced, and then the total production of this industry is given by the production of each variety times the number of firms (or varieties)

$$X = nx \tag{7}$$

From this, the indirect utility of agent i can be derived by substituting the equilibrium values in (5) and (6):

$$U_i = I_i(1-\alpha)^{1-\alpha}\alpha^{\alpha} n^{\alpha/(\sigma-1)} p^{-\alpha} \tag{8}$$

Considering, finally, the choice of agent i between a free trade state and an autarkic state, its relative utility under these two states is given by (9).

$$\frac{U_i^{FTA}}{U_i^{AUT}} = \left(\frac{I_i^{FTA}}{I_i^{AUT}}\right)\left(\frac{p^{FTA}}{p^{AUT}}\right)^{-\alpha}\left(\frac{n^{FTA}}{n^{AUT}}\right)^{\alpha/(\sigma-1)} \tag{9}$$

Note that in a median voter structure as developed above, the utility function (9) would also represent the relative utility of the society to which the typical agent i belongs. From equation (9), we expect that a given country will create an FTA providing that implementing this FTA leads to an increase in its utility, i.e., $\frac{U_i^{FTA}}{U_i^{AUT}} > 1$. This will occur provided a rise in income, i.e., $\left(\frac{I_i^{FTA}}{I_i^{AUT}}\right) > 1$, a decrease in prices (terms of trade effect), i.e., $\left(\frac{p^{FTA}}{p^{AUT}}\right) < 1$ or an increase in the number of varieties available, i.e., $\left(\frac{n^{FTA}}{n^{AUT}}\right) > 1$. Taking the logarithm of equation (9), the FTA will occur if $\log U^{FTA} - \log U^{AUT} > 0$. At this stage, it is worth mentioning that this model is close to Baldwin and Venables (1995) who analyze welfare effects of FTAs in imperfect competition. In a more recent study, Facchini et al. (2016) use a comparable utility-based political economy model to identify the heterogeneity between the determinants of free trade agreements and those of currency unions.

2. EMPIRICAL MODEL AND DATA

Though Equation (9) provides a sound theoretical ground for an FTA choice by a typical country, it does not yet include the bilateralism, which is essential in free trade agreements. The bilateralism implies that an FTA is implemented only if it is beneficial to both sides of the agreement. Following our median voter approach, this would happen when the utility condition from Equation (9): $\frac{U_i^{FTA}}{U_i^{AUT}} > 1$ is simultaneously satisfied in both partner countries. To this end, the forthcoming FTA should imply significant potentials to increase income, reduce prices and/or raise the number of varieties in the two partner countries as modeled through Equation (9). Such potentials could be, in turn, identified using international trade theories which specify the conditions under which free trade would achieve those objectives.

We classify international trade theories under two wide schools of thought. The first one relates the benefits of countries from free trade to their differences (or similarities) in factor endowments, demand structure, technology and trade costs. The famous theories such as comparative advantages (Ricardo, 1817), Heckscher-Ohlin (Heckscher, 1919 and Ohlin, 1933), Linder (1961), Viner (1950) and local comparative advantages (Deardorff, 2004) are in this school. The second line of thought however includes the monopolistic competition theories of international trade (the new trade theory) which highlight the role of economies of scale (market size) and product differentiation in suggesting free trade.

Equation (10) represents our estimable model. It adds the bilateralism to equation (9) in two ways. First, the utility ratio at the left hand side of equation (9) is replaced by a binary choice variable, FTA_{ijt}, which takes the value 1 if an FTA exists between countries *i* and *j* at time *t* and zero otherwise. The bilateralism implies that $FTA_{ijt} = 1$ only if the utility-ratio condition is simultaneously satisfied in both countries, i.e., $\frac{U_i^{FTA}}{U_i^{AUT}} > 1$ and $\frac{U_j^{FTA}}{U_j^{AUT}} > 1$. In other terms, the existence of an FTA between two countries reveals their preferences to have it.

Second, we take the advantage of international trade theories to replace the relative variables at the right hand side of equation (9) by a series of proxy variables that measure the benefits of the two partner countries (from free trade) in terms of those variables. These proxies represent the differences of the two countries in factor endowments, technology and trade costs as well as their market sizes, in line with our bilateralism consideration for equation (9).

$$\begin{aligned} FTA_{ijt} = {} & \beta_o + \beta_1(ln\,DTEC_{ijt}) + \beta_2(lnDWAGE_{ijt}) \\ & + \beta_3(lnDRENT_{ijt}) + \beta_4(lnSCOMP_{ijt}) + \beta_5(lnDIST_{ij}) \\ & + \beta_6(CONT_{ij}) + \beta_7(LANG_{ij}) + \gamma_i + \delta_j + \theta_t + \varepsilon_{ijt} \end{aligned} \tag{10}$$

To avoid data scarcity on factor endowments, we use factor rewards which are assumed to be negatively related to factor endowments.[3] In this regard, $DWAGE_{ijt}$ reflects the difference in wages and $DRENT_{ijt}$ the difference in the rental rate of capital between countries *i* and *j* at *t*. We expect that the greater the difference in factor rewards between two countries, the lower their chance to create an FTA. This idea is supported by the Viner theory that states that FTA is more beneficial to similar countries.

$DTEC_{ijt}$ measures productivity differences between countries *i* and *j* at time *t*. It has been obtained from the productivity capability index initially developed by Archibugi and Coco (2004). The index has been calculated as an unweighted index

---

[3] Anyway, the sign of correlation does not matter here, as we use the absolute difference of factor rewards across i and j.

of primarily calculated three other indices, i.e., technology-creation index, technology-infrastructure index and human skill index. Each index has been calculated by taking into account several other measures described fully in Table 2. Overall, 10 variables are included in the calculation of the productivity index. We expect that the greater the productivity difference between two countries, the lower the likelihood for an FTA, according to the Viner theory again.

The market size is measured by the sum of the number of companies or alternatively by the sum of GDPs of the two countries. Indeed, following the monopolistic competition theory, we expect that the greater the size of two countries, the more likely they create an FTA, in order to benefit from a larger number of varieties that becomes available at lower prices. As a sensitivity test, we would also use the sum of GDPs of countries *i* and *j*. In addition, the differences across countries in GDP is tested.

DIST, CONT and LANG reflect distance, contiguity and difference in languages, respectively. These are traditional trade-cost variables that reflect the price variable in equation (9). In this regard, the higher the trade costs, the lower the chance for reducing prices after an FTA and thus the lower probability of creating an FTA. Bilateral distance is measured by a spatial weighted index that takes into account the geographical distribution of the population, as developed by Cepii (Dist Database). Correspondingly, we expect its coefficient to be negative. CONT and LANG are dummy variables that take the value of 1, respectively, for common border and language, and 0 in the other cases. We expect a positive sign for these last two variables.

Finally, $\gamma_i + \delta_j + \theta_t$ denote country and time specific effects dedicated to capture potentially omitted variables.

As a specification test, equation (11) provides an extended version of equation (10) by including the multilateral and FTA interdependence effects, as discussed in Baier et al. (2014).

$$\begin{aligned} FTA_{ijt} = \beta_o + \beta_1(\ln DTEC_{ijt}) + \beta_2(lnDWAGE_{ijt}) \\ + \beta_3(lnDRENT_{ijt}) + \beta_4(lnSCOMP_{ijt}) + \beta_5(lnDIST_{ij}) \\ + \beta_6(CONT_{ij}) + \beta_7(LANG_{ij}) + \beta_8(MDIST_{ij}) + \beta_9(MFTA_i) \\ + \beta_{10}(MFTA_j) + \beta_{11}(ROWFTA_{ij}) + \gamma_i + \delta_j + \theta_t + \varepsilon_{ijt} \end{aligned} \tag{11}$$

Following Baier et al. (2014), MDIST accounts for multilateral distance. This is measured as the average of the distances separating each country *i* and *j* from all their other partners:

$$MDIST_{ij} = \frac{1}{2N}\left(\sum_{k=1,k\neq j}^{N} DIST_{ik} + \sum_{k=1,k\neq i}^{N} DIST_{jk}\right)$$

In other words, it measures the remoteness of countries *i* and *j* from the rest of the world. We expect that the higher this multilateral distance, the lower the relative bilateral distance between *i* and *j*, and thus more probable the conclusion of a common FTA between them.

We include three other multilateral variables in order to test the *own-FTA* and *cross-FTA* effects as described in Bair et al. (2014). The *own-FTA* effects are captured by

$$MFTA_{i,t-5} = \textstyle\sum_{k\neq j}^{N} FTA_{ik,t-5} \text{ and } MFTA_{j,t-5} = \textstyle\sum_{k\neq i}^{N} FTA_{jk,t-5}$$

where $MFTA_{i,t-5}$ is a multilateral index of country *i*'s FTA with all other countries (non *j*) lagged 5 years (to avoid endogeneity). This corresponds to the number of country i's FTAs with all countries except its current partner *j*, given that $FTA_{ik,t-5}$

is a dummy variable that is equal to 1 if *i* and *k* have an FTA in year *t-5*, and 0 otherwise. A similar definition can be given for $MFTA_{j,t-5}$. We expect that the more a country is involved in other FTAs, the more likely it will conclude a new FTA with a third party (Bair et al., 2014). Finally, the *cross-FTA* effects are shown by

$$ROWFTA_{ij,t-5} = \sum_{k \neq i,j}^{N} \sum_{l \neq i,j}^{N} FTA_{kl,t-5}$$

This last is going to capture the effect of all FTAs concluded elsewhere in the world on the probability of creating an FTA between countries *i* and *j*.

Table 1. Expected sign and sources for the variables

Variable	Name	Expected sign	Source
Difference in technology	DTEC	-	refer to Table 2
Difference in wages	DWAGE	-	Global Wage Database (ILO)
Difference in rental rate of capital	DRENT	-	WDI
Sum of GDP	SGDP	+	WDI (constant 2005 US$)
Sum of the number of companies	SCOMP	+	WDI
Distance	DIST	-	CEPII
Contiguity	CONT	+	Dummy
Common Language	LANG	+	Dummy
Multilateral distance	MDIST	+	CEPII
Multilateral own FTA for i	MFTAi	+	own calculation
Multilateral own FTA for j	MFTAj	+	own calculation
cross-FTA	ROWFTA	+	own calculation
Difference in GDP	DGDP	-	WDI (constant 2005 US$)

The list of variables including their expected signs and data sources has been provided in Table 1. Table 2 also details the components of the productivity capability index used to construct the variable $DTEC_{ijt}$ for the model.

Table 2. The Productivity Capability Index

Technology-Creation Index	1- Number of patent grants per 1 million people 2- Number of publications in scientific and technical journals per 1 million people	World Intellectual Property Organization, World Bank (WDI)
Technology-Infrastructure Index	3- Fixed broadband Internet subscribers per 100 people 4- Telephone fixed-lines per 100 people 5- Mobile cellular subscriptions per 100 people 6- Electric power consumption (kWh per capita)	World Bank (WDI)
Human-Skill Index	7- Literacy rate, adult total (% of people ages 15 and above) 8- Enrolment in tertiary education per 100,000 inhabitants 9- Mean years of schooling of adults	World Bank (WDI) UNESCO United Nations Development Programme (UNDP)
Productivity Capability Index	- Technology Creation Index (DCRE) - Technology Infrastructure Index (DINF) - Human Skill Index (DHSK)	--

In all time-series, missing observations have been filled using the average of their nearby observations. In addition, all indicators included in the Productivity Capability Index are in non-monetary measures and have been expressed in normalized values in order to preserve the comparable feature of data across countries and over time. We have achieved the normalization using the following formula:

$$\frac{\text{Observed value of observation} - \text{Minimum value of observations}}{\text{Maximum value of observations} - \text{Minimum value of observations}}$$

The above formula has been applied in logarithmic scale with regard to the indicators 4 through 7 in Table 2. This allows us to alleviate large differences (between the observed and minimum values) that may not be meaningful once a country reaches the effective level in those indicators. That is, a country with 3 fixed telephone lines per person does not necessarily have a telecommunication infrastructure three times stronger than that of a country with 1 fixed telephone line per person. Using logarithmic measures reduces the impact of those insignificant differences.

Finally, each index has been calculated using the simple average of its component indicators (as listed by Table 2).

3. ESTIMATION

The empirical equation (10) and the extended version (11) should be estimated using limited dependent variable models such as panel probit or panel logit that are based on maximum likelihood (ML) procedure. In addition, the presence of country and time fixed effects in the model implies the application of fixed-effect estimators. The model, however, contains a much larger number of country pairs (panel ID: 105*105=11025) compared to time periods (39 years: 1981-2019) and, in this case (large N versus small T), the ML estimation of fixed effects would be subject to serious inconsistency unless we use it with conditions on the fixed effects (conditional ML method). Such conditioning works only under logistic regression (for details, see Maddala, 1987, or Verbeek 2004: section 10.7). That is why we rely on the panel fixed-effects logit regression for the estimation of equations (10) and (11). In addition, the ML-based model allows for the correlated errors across countries and time (Larch et al. 2019) which is another advantage here for our logit estimator. However, we will also provide the estimates from the panel fixed-effects probit model for the purpose of comparison.

In order to address the multicollinearity bias in the variables, we use them in an "orthogonalized" scale. This allows us to make the variables mathematically independent and so to obtain the net effect of each one on the FTA binary-choice variable. However, the method does not significantly affect the fitted measures (estimated probabilities) from the model.

Results are displayed in Table 3. Column (1) represents our main estimates, from the panel fixed-logistic model (10). Accordingly, the probability of forming a bilateral FTA is negatively related to the differences of partners in factor rewards and in technology, positively to their market sizes and negatively to their bilateral trade costs. The first estimates have been provided including the country and time effects. For the purpose of comparison, Column (3) estimates the model without these effects. Interestingly, the estimated coefficients remain stable in sign after this last exclusion.

In a binary dependent-variable model, the marginal effect of a k^{th}, x^k, explanatory variable on the dependent variable, can be obtained from the partial derivative of the dependent variable with respect to that regressor:

$$\frac{\partial L(\acute{W}\beta)}{\partial \mathrm{x}^{\mathrm{k}}} = \frac{e^{\acute{W}\beta}}{\left(1+e^{\acute{W}\beta}\right)^{2}}\beta_k$$

where $L(\acute{W}\beta)$ is the likelihood function calculated using the averaged values of the regressors, W, and the vector of the estimated parameters, β. β_k is then the estimated parameter for x^k. Column (2) displays the calculated marginal effects, based on the estimates already provided by Column (1). To interpret these results, we can, for example, say that a one-percent increase in technological differences across two countries reduces the probability of a bilateral FTA between them by 0.54 percent, according to the marginal effect of DTEC.

Our results are in line with the trade theories of economic integration indicating that the gross trade creation of an FTA is greater if partners are similar in factor endowments and technology (Viner theory), if the FTA has a significant size (New Trade Theory) and if the bilateral relative trade costs are low (local CA).

Table 3. Estimation results

	(1)	***(2)***	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
Dependent variable	**FTA**	***Marginal effects***	**FTA**	**FTA-Liu**	**FTA**	**FTA**	**FTA**	**FTA**	**FTA**	**FTA**	**FTA**
DTEC	-0.902^{***}	*-0.0054*	-0.602^{***}	-0.103^{***}	-0.503^{***}	-	-	-	-0.605^{***}	-0.650^{***}	-0.481^{***}
DWAGE	-0.312^{***}	*-0.0050*	-0.845^{***}	-0.903^{**}	-0.756^{**}	-0.787^{***}	-0.581^{***}	-0.472^{***}	-0.506^{***}	-0.451^{***}	-0.139^{***}
DRENT	-0.301^{***}	*-0.0036*	-0.701^{***}	-0.551^{***}	-0.704^{***}	-0.386^{***}	-0.398^{***}	-0.389^{***}	-0.555^{***}	-0.405^{***}	-0.209^{***}
SCOMP	0.101^{**}	*0.0019*	0.501^{***}	0.850^{***}	-	0.405^{***}	0.350^{***}	0.474^{***}	0.426^{***}	0.085^{***}	0.015^{*}
DIST	-1.651^{***}	*-0.0082*	-1.077^{***}	-0.503^{***}	-1.171^{***}	-1.491^{***}	-1.506^{***}	-1.600^{***}	-1.257^{***}	-1.514^{***}	-0.935^{***}
CONT	0.442^{***}	*0.0031*	0.235^{***}	0.165^{***}	0.302^{***}	0.465^{***}	0.442^{***}	0.436^{***}	0.321^{***}	0.402^{***}	0.233^{***}
LANG	0.468^{**}	*0.0045*	0.254^{***}	0.156^{***}	0.198^{***}	0.509^{***}	0.535^{***}	0.525^{***}	0.280^{***}	0.356^{***}	0.262^{***}
SGDP					0.749^{***}						
DGDP					-0.281^{***}						
DHSK						-0.423^{***}					
DINF							-0.506^{***}				
DCRE								-0.419^{***}			
MDIST									1.905^{***}	3.850^{***}	
MFTAi									0.991^{***}	0.135^{***}	
MFTAj									0.761^{***}	0.030^{***}	
ROWFTA									0.0017^{*}	0.0005^{*}	
Constant	-11.28^{***}	-	-2.584^{***}	-5.034^{***}	-14.55^{***}	-10.52^{***}	-11.39^{***}	-11.56^{***}	-12.19^{***}	-42.17^{***}	-5.58^{***}
Number of obs.	254,139	*254,139*	310,363	215,201	307,256	254,139	254,139	254,139	269,347	269,347	254,139
Pseudo R²	0.591	-	0.381	0.203	0.307	0.553	0.570	0.582	0.481	0.569	0.581
Count. fixed effects	yes	*yes*	no	yes	no	yes	yes	yes	no	yes	yes
Time fixed effects	yes	*yes*	no	yes	no	yes	yes	yes	no	yes	yes

Definition of variables: see Table 1 and 2.

The next 8 columns in Table 3 check the sensitivity of our main estimates to a modification in data, model specification or estimator. The first sensitivity test (Column 4) verifies the correction for the temporal-dependence bias as suggested by Liu (2008). This bias is because once a country pair concludes an FTA, the dependent variable moves from the value 0 to 1, but can never go back to the value 0. Such a bias can be corrected by removing the observations the year after the conclusion of the FTA (Liu 2008). In this case, if an FTA is concluded, only the first observation with FTA=1 is kept. This method also addresses the endogeneity problem due to the impact of post-FTA trade on the factor rewards in our model. However, one drawback of this method is that it strongly increases the proportion of the FTA values equal to 0, as compared with values equal to 1, while significantly reducing the number of observations. Column (4) provides the Liu-based estimates. They are more or less different from our previous results in column (1), but still the same in sign and significance level.

Table 4. Bilateral probability matrix (in %)

	.	23.0	90.3	18.0	40.3	19.8	64.3	51.5	59.8	52.7	31.6	63.5	40.3	49.7	62.1	27.3	77.2	29.2	44.4	ALB
ALB	ALB	ISR	94.1	29.7	77.9	34.4	91.6	77.0	98.6	82.4	37.3	98.2	39.1	53.4	98.6	51.8	99.7	17.0	83.1	DZA
DZA	42.6	DZA	ITA	81.7	51.3	87.7	99.1	84.0	72.6	98.0	96.5	96.8	98.2	99.1	97.7	94.0	84.6	26.6	95.8	AUT
AUT	61.2	68.8	AUT	LVA	85.4	88.2	99.9	86.1	97.8	100.0	90.6	99.0	88.8	94.3	99.2	98.3	93.3	31.7	99.9	BEL
BEL	57.4	95.2	99.5	BEL	LBN	22.8	65.6	25.7	22.3	48.6	37.3	61.5	57.0	74.9	50.1	23.4	59.1	18.5	39.9	BIH
BIH	55.1	30.0	74.6	50.0	BIH	LTU	70.3	44.3	43.6	67.9	66.0	72.4	74.1	61.7	77.8	49.5	61.3	58.6	54.4	BGR
BGR	76.8	30.5	74.4	73.5	37.4	BGR	LUX	48.0	45.1	83.5	69.7	80.6	84.7	95.3	84.9	55.9	66.5	13.0	70.8	HRV
HRV	55.9	35.3	96.6	89.1	82.2	75.3	HRV	MLT	47.3	69.5	53.6	89.1	60.0	52.0	84.8	62.3	67.5	90.3	65.2	CYP
CYP	32.0	38.1	70.0	69.3	18.8	90.1	60.6	CYP	MAR	95.0	95.3	85.9	93.6	94.2	88.3	87.7	56.4	10.6	84.8	CZE
CZE	24.4	33.0	97.2	91.8	33.5	55.6	78.3	43.4	CZE	NLD	97.5	98.2	93.5	93.8	98.3	100.0	90.3	35.9	99.5	DNK
DNK	52.0	82.2	98.1	99.7	60.7	72.7	82.2	80.0	95.5	DNK	POL	84.2	33.9	37.0	78.6	46.8	97.9	68.1	84.5	EGY
EGY	57.4	91.8	54.0	71.0	17.9	62.5	32.3	95.0	23.9	68.5	EGY	PRT	69.9	87.2	88.5	98.2	52.8	10.9	84.8	EST
EST	14.0	30.9	83.8	91.3	12.7	35.9	32.6	44.2	67.2	97.4	32.2	EST	SVK	78.8	97.5	99.3	71.0	27.0	97.4	FIN
FIN	30.9	52.4	95.2	98.5	24.9	54.7	56.3	66.1	79.6	99.3	54.9	99.9	FIN	SVN	99.2	97.8	96.6	37.4	99.8	FRA
FRA	65.0	98.2	98.7	99.9	65.0	71.2	86.1	84.2	91.9	99.6	73.6	87.7	97.5	FRA	ESP	99.7	92.4	40.1	99.6	DEU
DEU	64.8	83.9	99.2	100.0	62.5	82.1	94.7	73.5	95.8	99.8	73.6	94.3	98.1	99.2	DEU	SWE	98.0	81.4	94.9	GRC
GRC	93.5	88.7	97.3	98.1	80.9	97.9	95.8	99.4	87.3	97.4	97.6	89.9	93.9	97.6	98.6	GRC	TUN	23.3	79.1	HUN
HUN	52.7	46.6	95.4	90.6	74.5	82.6	87.6	56.2	95.0	93.3	40.9	71.1	76.4	89.4	96.2	96.6	HUN	TUR	99.9	IRL
IRL	45.0	84.7	96.3	99.6	37.3	71.6	79.2	82.2	89.6	99.4	85.5	91.4	97.7	99.5	99.3	98.8	87.6	IRL	GBR	
EU	***58.4***	***80.6***	**93.4**	**93.8**	***56.2***	**71.5**	**82.6**	**73.6**	**88.0**	**95.2**	***69.9***	**86.0**	**92.8**	**81.9**	**77.1**	**93.3**	**87.2**	**95.0**		
Non-EU	**41.7**	**46.5**	***49.4***	***62.8***	**21.8**	***54.1***	***27.5***	***79.9***	***23.8***	***60.4***	**59.0**	***27.0***	***46.2***	***66.9***	***64.1***	***89.5***	***37.8***	***64.3***		

ISR	ISR	ALB	60.5	76.3	88.5	10.7	53.6	29.0	98.8	43.6	79.6	97.6	58.1	76.2	91.9	81.8	97.8	50.0	95.2	ISR
ITA	91.4	ITA	DZA	98.3	99.2	86.2	90.9	98.1	91.2	95.4	98.3	83.0	88.3	97.4	98.9	99.2	99.7	96.3	99.5	ITA
LVA	49.9	87.6	LVA	AUT	87.8	17.3	50.5	47.5	51.4	82.2	97.1	32.9	97.8	98.6	85.8	94.9	88.8	87.5	90.3	LVA
LBN	99.4	77.8	35.9	LBN	BEL	45.5	52.6	24.0	99.2	24.3	73.5	99.1	35.2	56.6	90.4	70.8	95.4	41.0	82.5	LBN
LTU	48.5	89.1	99.3	44.0	LTU	BIH	57.7	51.5	57.5	89.4	97.7	32.7	97.2	97.2	87.6	94.8	89.7	91.0	90.2	LTU
LUX	80.5	99.6	88.7	70.6	89.9	LUX	BGR	86.5	73.7	96.7	99.6	78.3	90.5	97.7	99.8	99.8	97.2	91.0	99.5	LUX
MLT	87.9	99.0	30.3	71.5	32.3	91.8	MLT	HRV	62.2	40.2	82.4	80.5	32.5	86.4	93.8	84.7	97.0	48.8	95.6	MLT
MAR	73.3	94.7	46.3	79.9	44.8	95.1	70.8	MAR	CYP	42.9	87.9	96.0	48.7	71.1	98.7	91.2	94.1	51.4	94.6	MAR
NLD	78.5	99.1	90.6	59.4	90.9	100.0	85.6	91.4	NLD	CZE	99.9	66.2	90.9	98.1	99.9	99.9	96.9	90.1	99.8	NLD
POL	55.9	94.7	92.1	38.2	88.3	96.4	45.3	49.3	94.7	POL	DNK	34.5	83.0	92.0	90.5	92.6	90.5	92.7	91.6	POL
PRT	84.4	99.4	85.3	82.8	86.1	99.4	93.7	99.8	98.7	87.1	PRT	EGY	87.9	95.1	99.8	98.2	98.7	85.2	99.6	PRT
SVK	40.6	95.7	74.2	37.2	83.0	93.0	41.2	49.5	89.1	97.3	88.8	SVK	EST	77.1	92.9	94.6	93.4	98.6	85.2	SVK
SVN	62.7	98.4	74.2	33.0	77.4	97.2	67.1	63.6	94.3	86.4	94.9	94.5	SVN	FIN	96.4	97.3	97.0	90.8	92.3	SVN
ESP	91.6	99.9	84.3	71.3	85.3	99.5	97.4	99.6	99.2	89.9	99.9	88.2	96.3	ESP	FRA	98.7	99.5	88.2	99.8	ESP
SWE	76.0	96.8	97.3	44.1	96.2	98.8	76.6	71.1	99.5	94.7	94.7	80.3	84.1	97.3	SWE	DEU	92.9	79.2	98.7	SWE
TUN	83.3	99.0	49.9	92.3	56.6	95.7	97.7	99.2	90.0	61.0	98.4	68.2	82.2	98.2	70.4	TUN	GRC	74.9	89.8	TUN
TUR	53.8	54.4	15.7	86.8	15.6	43.2	19.3	29.7	34.2	20.0	48.9	17.4	14.6	38.1	19.5	43.6	TUR	HUN	30.3	TUR
GBR	87.6	98.3	82.9	81.1	82.5	99.8	96.4	93.1	99.9	86.2	99.4	77.5	87.1	99.1	98.2	89.3	31.0	GBR	IRL	GBR
EU	***81.1***	**84.7**	**85.9**	***70.1***	**86.5**	**96.6**	**84.5**	***86.3***	**92.3**	**84.8**	**93.8**	**87.6**	**91.5**	**87.0**	**92.4**	***87.8***	***36.4***	**82.1**		
Non-EU	**64.8**	***75.3***	***28.1***	**86.5**	***29.0***	***67.9***	***55.5***	**61.2**	***58.9***	***32.6***	***73.5***	***31.2***	***36.1***	***68.0***	***40.8***	**72.1**	**29.3**	***63.5***		

Annex provides the correspondence table between the country names and country codes.

Table 5. Probability of Regional Integration

Country Group	European Med. Countries	Non-European Med. Countries
European Countries	82.6%	60.5%
Non-European Med. Countries	60.5%	46.6%

In contrast, the explanatory power from column (1) is much higher than what column (4) suggests. This provides us an adequate reason to favor the original model over the Liu-based model, because our main objective here is predicting future FTA (and Non-FTA) events and, of course, the prediction power of the model is very important to that end.

In the same way, columns 5 through 11 provide other sensitivity tests by replacing some of the independent variables by their alternatives (columns 5-8), including multilateral variables (columns 9-10) as suggested by Baier et al. (2014), and using an alternative estimator: probit model (column 11). Overall, the estimates from these last columns are similar in sign, magnitude, and significance level to what we obtained based on our initial model in column (1). This makes us more confident of the robustness of our results in columns (1) and (2).

We rely now on the estimated model (10) to calculate the probability of FTA creation for all country pairs in our database. This can be achieved through simulations based on the estimated parameters in Column (1) of Table 3. The calculated probability matrix is given by Table 4. In order to remove the worthless temporal variations, we report the averaged values over the last 5 years, i.e., 2015-2019. Notably, the averaged probabilities are ranged within the full scale [0,1].

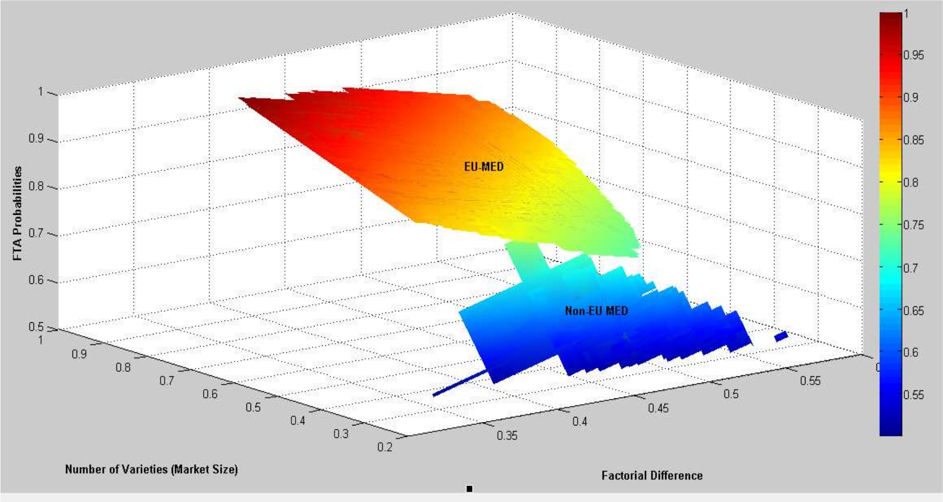
We also calculated the average probability (weighted by GDP) that each country possesses in order to form an FTA with the entire set of European and Non-European Mediterranean countries respectively. The results have been presented by the last two rows of Table 4 from which each number is (vertically) linked to the first country listed in 3-letter codes on the diameter of the matrix. As shown, the results are diverse according to the country and the chosen partnership (EU or Non-EU Med). In addition to the majority of EU countries, Algeria, Egypt, Lebanon, Israel, Morocco, and Tunisia also show high probabilities of forming an FTA with the EU set. In the same way, Belgium, Cyprus, France, Germany, Greece and Ireland have great opportunities for concluding an FTA with the Non-European Mediterranean set.

These calculated probabilities at country level allow us to step in further by calculating a single probability for the EU set to conclude an FTA with the non-EU set. We can achieve this by taking the probabilities of EU countries (only with the non-EU set) and weighting them by their GDP shares. The result shows that this overall probability is equal to 60.5%. A similar calculation for non-EU countries to create an FTA with EU countries provides, of course, the same result (Table 5). In addition, one can calculate the probabilities of regional integration within the European and Non-European Mediterranean sets by income weighting the calculated probabilities of countries within their own set and adding them up to a single number. This gives us 82.6% for the EU set and 46.6% for the non-EU Med set (Table 5). Interestingly, the probability of forming an interregional FTA within the Euromed area is higher than the probability of forming a regional Non-European Mediterranean FTA.

The 3D graph in Figure 1 shows how the probability of creating bilateral FTAs dif-fers between the European and Non-European Mediterranean countries based on the number of varieties (market size) of partners and their differences in factor endow-ments. Factorial difference is measured by the average of normalized values of differ-ences in technology, wages and rental rate of capital as specified in model (10). The number of varieties is also proxied by the normalized value of the number of compa-nies. The advantage of normalization is to get variables that vary between 0 and 1.

In line with our numerical findings in Table 5, the FTA probability surface for the European set is above that for the Non-European Med set. In addition, we can see that a decrease in factor difference and/or an increase in the number of varieties (market size) lead to a rise in the probability of creating FTA in both of the country groups. The surface for the European countries is however deeper, showing more sensitivity of FTA creation to the fundamental factors specified in our model.

Figure 1. Sub-Regional FTA Probability Surface: EU versus Non-EU



4. PERFORMANCE EVALUATION

In this section, we employ ROC (Receiver Operating Characteristic) metrics to evaluate the performance of our model in truly predicting the FTA and no-FTA events. The ROC analysis is usually used in medical sciences for comparing the operating characteristics as the criterion changes. This technic is also increasingly used in Finance to evaluate the performance of PD (Probability of Default) models. The procedure consists of finding how much close is the performance of the real model to the performance of a (presumed) perfect model, or how much far it is from the null performance of a (presumed) random model, in predicting the past events. Such a relative performance is obtained through the comparison of cumulative distribution functions (CDFs) across these models in predicting the FTA and no-FTA events.

To calculate the CDFs for FTA events, the first step consists of sorting the entire database by the estimated probabilities, in a descending order. If our model was *perfect*, the first *n* observations (with the highest probabilities) should correspond to the real FTAs; where *n* is the total number of FTAs in the database (*n*=2798 in 2019). That is because a perfect model would assign, by definition, the highest estimated probabilities to the FTA realizations in an inclusive and exclusive manner. In this case, the CDF for such a perfect model would increment by 1/n from the top of the database down to the n^{th} observation and remains equal to 1 thereafter.

Keeping the same order of observations, the CDF for the real model is obtained through incrementing by 1/n (moving from the highest probability to the lowest) wherever the observation corresponds to a real FTA (FTA=1). In this case, the real CDF reaches its maximum value (i.e. one) when all the real FTAs are covered. In contrast, the CDF for the random model increments by 1/k in every observation over the entire sorted database; where k is the total number of non-missing observations, i.e., 310,163 obs. This is because the random model does not distinguish the FTAs from no-FTAs and has a null explanatory power by definition. As a result, this latter achieves its maximum (i.e. one) at the very end of the database.

Figure 2-1. Relative performance of the CDFs in predicting FTAs

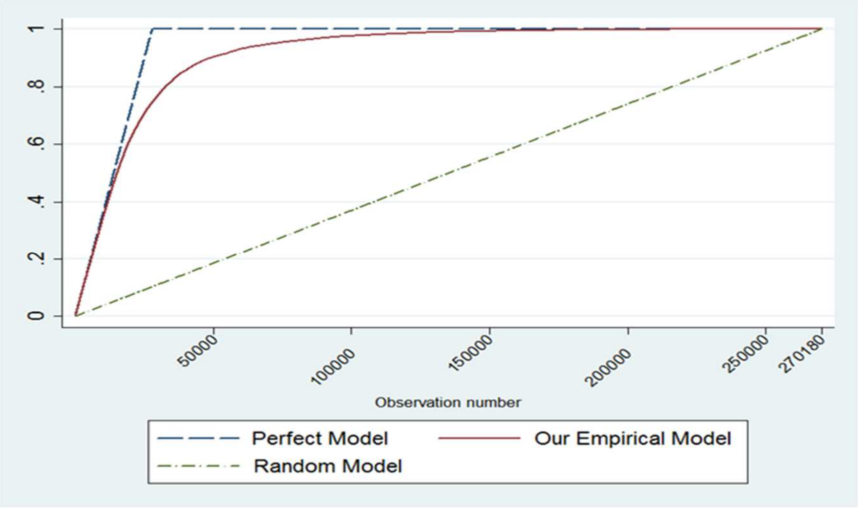
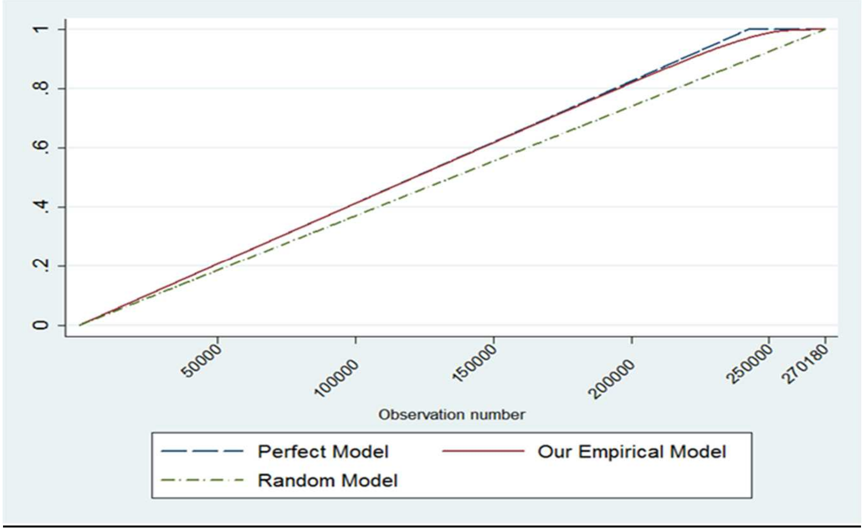


Figure 2-2. Relative performance of the CDFs in predicting No-FTAs



Similarly, we can obtain the same set of three CDFs for No-FTA cases. In addition, note that the database should be sorted in an ascending order in this case and that *n* corresponds now to the number of No-FTA cases. These changes imply, for example, that the real CDF increments by 1/n wherever a No-FTA record is observed in the database (moving from the lowest probability to the highest in the newly sorted database).

Figure 2-1 and Figure 2-2 show the calculated CDFs respectively with regard to FTA and No-FTA events as explained above. As shown, the solid lines (calculated from the real model) are remarkably close to the dashed lines (calculated from a presumed perfect model) and far away from the dash-dotted lines (calculated from a presumed random model). This clearly indicates that our empirical model performs far beyond a random model. Our model is particularly close to the perfect model in the first percentiles of the sorted fitted probabilities. This means that high estimated probabilities perfectly match real FTA events while low estimated probabilities perfectly match real No-FTA events. In other words, our model is extremely precise over these ending percentiles. This exceptional precision accompanies, of course, the generally good performance of the model in the middle percentiles of the estimated probabilities.

5. INTERPRETATION OF THE RESULTS

The remaining question is "how to interpret the estimated probabilities?" In other words, what do the likelihoods of 82.6% (EU), 46.6% (Non-EU) and 60.5% (Euromed) mean? Are they high enough to back a regional integration in those areas?

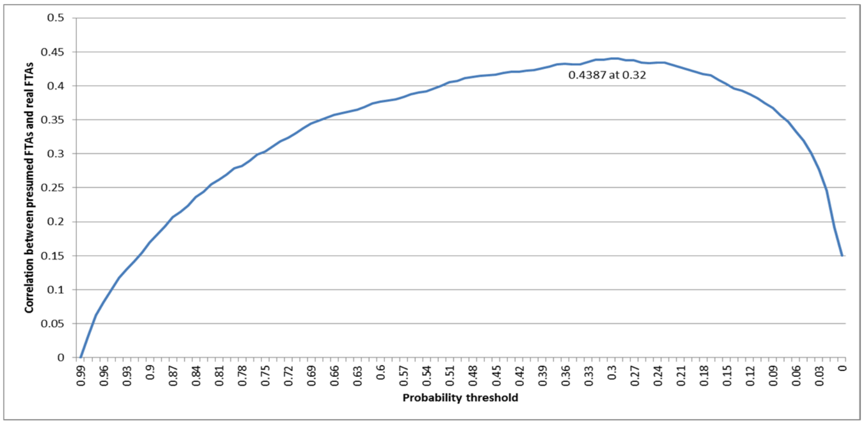
To answer this question, we need to find some probability cutoffs allowing to attribute different qualities to the estimated probabilities. One choice for the probability cutoff is the middle value in the range of [0, 1], i.e., 0.5. In this case, all country pairs with fitted probabilities exceeding 0.5 are qualified for FTA creation, and those for which the probabilities are lower are not. The problem with this measure is that it does not necessarily maximize the predictive power of our model.

The other choice might be the use of unconditional probabilities that are indeed the relative frequencies of FTA versus No-FTA events in the sample, 9.9 versus 81.1 percent respectively. However, as emphasized by Wooldridge (2000), these measures do not effectively separate a good model from a bad one. That is, even a model that fails to predict correctly one FTA can correctly predict almost 81.1 percent of No-FTAs, resulting in a misleading predictive power.

In order to address this problem, Wooldridge (2000) suggests using the proportion of truly predicted FTAs, in the total FTAs, and the proportion of truly rejected FTAs, in the total No-FTAs. The lower the probability cutoff, the higher would be the first proportion and the lower would be the second proportion. The solution then consists of choosing a probability cutoff that maximizes the overall predictive power.

Based on this initiative, we suggest a method that is particularly more efficient in large databases. To do so, we first consider a series of 100 possible probability cutoffs from 0 to 1, by steps of 0.01. By each probability cutoff, a new variable, namely, presumed FTA, is defined that takes the value 1 if the estimated probabilities from the model exceed the corresponding probability cutoff, and 0 otherwise. We then plot the correlation between the presumed and real FTA variables against the corresponding probability cutoffs (Figure 3). The probability cutoff giving the maximum correlation coefficient is chosen as the general probability threshold.

Figure 3. General choice of probability threshold



The correlation curve reaches its maximum at a probability cutoff equal to 0.32 (Figure 3). For the sake of parsimony, we round this number, as well as all following similar numbers, up to the nearest multiple of 0.05. Accordingly, all records with an estimated probability in the range of [0.35, 1] are considered as expected (or feasible) FTAs and those with an estimated probability lower than 0.35 as unexpected FTAs.

However, it does not look very sound to put the estimated probabilities that are slightly higher than 0.35 in the same basket as those that are close to 1. The same argument is true for the estimated probabilities at the left-hand side of the general threshold. As a result, more criteria should be considered in order to more granularly relate the estimated probabilities to FTA and No-FTA realizations.

In this regard, we define four FTA and two No-FTA realizations according to the precision that we expect from each one in truly predicting FTA and No-FTA events respectively. Table 6 provides our precision requirements. The general threshold 0.35 has of course, been used to separate the FTA side from the No-FTA side. Accordingly, the values A, B and C in Table 6 should be determined in the way that if we refer the estimated probabilities falling in [35, A), [A, B), [B, C) or [C, 100] to some FTA realizations, we would be right in 50%, 75%, 90% and more than 90% of the cases respectively. This is consistent with the fact that the higher the estimated probabilities are, the more accurately our model predicts FTA realizations. In the same way, D (on the No-FTA side) should be determined so that the probability ranges [D, 35) and [0, D) respectively provide us 75% and more than 90% of precision in predicting No-FTA events.

Table 6. Precision reference of FTA evaluations

Quality attributions	*No-FTA realizations*		*FTA realizations*			
	Impossible	Unexpected	Possible	Reasonably possible	Probable	Highly Probable
Probability range	[0, D)	[D, 35)	[35, A)	[A, B)	[B, C)	[C, 100]
Truly predicted (%)	>90	75	50	75	90	>90
Incorrectly predicted (%)	<10	25	50	25	10	<10

We employ a *Reverse Backtesting*[4] procedure to determine the values A, B, C and D in the way to get our desired precisions in Table 6. Table 7 shows, among others, the probability ranges allowing us to get our precision objectives when predicting the existing FTA and No-FTA events. For example, if we qualify all country pairs with an estimated FTA probability between 35% and 65% as having an FTA, we would be right only in 52.8% of the cases. This relatively low level of precision does not allow us to attribute an FTA realization stronger than "*Possible"* to those probabilities, in line with our defined decision rules in Table 6. In the same way, if we recognize all estimated probabilities in the range [95%, 100%] sufficiently high for representing an FTA, we would be right in 96.5% of the cases. This high level of precision allows us to attribute an FTA realization as strong as "*Highly Probable*" to those estimated probabilities, again according to our precision rules in Table 6. In the same way, one can interpret other probability ranges in Table 7.

---

[4] Backtesting is the process of measuring the performance of a decision rule through judging it on the past data. The procedure is commonly used in different areas in Finance, for example, in order to determine the prediction power of a PD model. Given our precision objectives in Table 7, we apply a reverse backtesting procedure to determine the decision rules which satisfy them.

Based on these intermediate thresholds, it is now possible to tell more about the overall figures of 82.6%, 60.5% and 46.6% that we found respectively for a European, Non-European Med and Euromed FTA creations. Such FTAs are respectively perceived as probable, possible and possible. In other terms, European integration seems to be highly supported by our model. The interregional Euromed integration even appears more possible (though in the same category) than an intra integration in the Non-European Mediterranean area. Of course, these overall figures should be seen with all heterogeneities that we found at bilateral or country level as provided through Table 4.

Table 7. The empirical thresholds

	In favour of No-FTA (< 0.35)			
FTA evaluation	Impossible		Unexpected	
Probability range (%)	[0, 20)		[20, 35)	
	No.	%	No.	%
Truly predicted FTAs	3828	1.7	2802	24.1
Truly predicted No-FTAs	225538	98.3	8829	75.9
Total number	229366	100	11631	100

	In favour of FTA (≥ 0.35)							
FTA evaluation	Possible		Reasonably Possible		Probable		Highly Probable	
Probability range (%)	[35, 65)		[65, 75)		[75, 95)		[95, 100]	
	No.	%	No.	%	No.	%	No.	%
Truly predicted FTAs	6107	52.8	2456	74.9	8034.0	90.3	5240	96.5
Truly predicted No-FTAs	5465	47.2	824	25.1	867.0	9.7	190	3.5
Total number	11572	100	3280	100	8901	100	5430	100

6. CONCLUSION AND POLICY IMPLICATIONS

European integration and more generally regional integration in the Euromed area have been one of the deepest economic integrations in the world but also one of the most questioned especially after the recent events like Brexit.

According to the Median-Voter approach, two countries would proceed with free trade if and only if the FTA improves the relative utility of their median voters. Assuming that the current FTAs in the world are reflecting the revealed preferences of their involved countries towards free trade, we developed a conditional fixed-effect logit model which relates this revealed choice to some observable variables such as differences of countries in factor endowments and technology, their market sizes and trade costs. The estimation of the model then allows to identify the driving forces behind the binary choice of free trade across the world over 1981-2019.

Our model indicates that the more similar countries are in terms of factor rewards and technology, the more likely they are to conclude a free trade agreement. The economic size also stimulates free trade: the bigger countries are in terms of GDP, the more ready they are to develop their free trade. In contrast, we find that the likelihood of FTA is inversely related to the physical and cultural remoteness of countries from each other. The results are robust to a number of alternative estimators and specifications. Our analysis in terms of the Relative Operating Characteristic (ROC) strongly supports the performance of our model in truly predicting the (past) FTA and No-FTA events. This is a way to evaluate the model on its predictions for similar future events.

Using the estimated logit model, we provided the probabilities of joining FTA for all Euromed country pairs (i.e. 1260 country-pairs) as well as at country and regional level in the Euromed area. The estimates indicate a very heterogeneous pattern of economic integration across this area. We then used a Reverse Backtesting procedure to interpret the estimated probabilities according to the precision that each probability range offers in truly predicting the existing FTA and No-FTA events. Subsequently, each probability receives an FTA realization quality from being "impossible" to "highly probable".

With an FTA likelihood equal to 82.6%, the intra-European integration is perceived as *probable* and so receives a great support from our study. In addition to the majority of EU countries, Algeria, Egypt, Lebanon, Israel, Morocco, and Tunisia are the Non-European (Mediterranean) countries which show a high capacity of sharing an FTA with the EU set.

On the other shore of the Mediterranean Sea, the likelihood of creating a regional FTA is obtained equal to 46.6%. This is far from the feasibility of a regional integration amongst the European countries but still indicates a *possible* FTA according to our Reverse Backtesting procedure. In addition, Belgium, Cyprus, France, Germany, Greece and Ireland are those of the European countries that demonstrate some great opportunities for concluding an FTA with the Non-European Mediterranean set. Finally, the likelihood of forming an FTA over the entire Euromed area is estimated to be equal to 60.5%, indicating again a *possible* FTA. Besides these numerical figures, our structural logit model emphasizes on the role of technological upgrading, factor reward convergence and the creation of bigger markets in supporting a naturally successful FTA in the Euromed area as in any other part of the world.

As a limit and possible extensions of the model, some new debates are not yet considered. For example, the debate about different social or environmental standards when concluding a North-South FTA is becoming more and more important. It should be interesting to integrate this information in a model by taking into account the cost of such differences for partner countries.

Another debate, which has strongly emerged since the Covid-19 crisis, is related to the relevance of national and strategic industries (pharmaceutical, medical, etc.). In this regard, many countries question negotiations of future FTAs and rather look back to their own industries and employment. Whatever these debates, the new international trade theory still highlight the significant gains of multilateral and regional integration, in terms of efficiency, price reduction, scale economies, availabilities of varieties, dynamic gains, etc.

ANNEX. Country names

Country Name	Code	Country Name	Code	Country Name	Code
Albania	ALB	Finland	FIN	Malta	MLT
Algeria	DZA	France	FRA	Morocco	MAR
Austria	AUT	Germany	DEU	Netherlands	NLD
Belgium	BEL	Greece	GRC	Poland	POL
Bosnia and Herzegovina	BIH	Hungary	HUN	Portugal	PRT
Bulgaria	BGR	Ireland	IRL	Slovakia	SVK
Croatia	HRV	Israel	ISR	Slovenia	SVN
Cyprus	CYP	Italy	ITA	Spain	ESP
Czech Republic	CZE	Latvia	LVA	Sweden	SWE
Denmark	DNK	Lebanon	LBN	Tunisia	TUN
Egypt	EGY	Lithuania	LTU	Turkey	TUR
Estonia	EST	Luxembourg	LUX	United Kingdom	GBR

REFERENCES

Abedini, J. and N. Péridy (2018). Is there an economic rationale for leaving or joining the EU? Evidence from a political economy framework, *Economics Bulletin* 38(2), pp. 1-12.

Archibugi D., Coco A. (2004). A New Indicator of Technological Capabilities for Developed and Developing Countries (ArCo). *World Development*, 32 (4): 629–654.

Baier S., Bergstrand J. (2004). Economic determinants of free trade agreements. *Journal of International Economics*, 64(1): 29-63.

Baier S., Bergstrand J., Mariutto R. (2014). Economic Determinants of Free Trade Agreements revisited: Distinguishing Sources of Interdependence. *Review of International Economics*, 22(1): 31-58.

Baldwin R., Venables A. (1995). *Regional Economic Integration*. Handbook of International Economics, III, New York: North Holland.

ECORYS (2013a) Trade Sustainability Impact Assessment in support of negotiations of a DCFTA between the EU and Morocco, http://www.trade-sia.com/morocco/wp-content/uploads/sites/2/2014/02/FR-Morocco-annexes1.pdf

ECORYS (2013b) « Trade Sustainability Impact Assessment in support of negotiations of a DCFTA between the EU and Tunisia", http://www.trade-sia.com/tunisia/2014/01/08/final-report-now-published-online/

Facchini G., Silva P., Willmann G. (2016). The Political Economy of Preferential Trade Arrangements: An Empirical Investigation. The University of Nottingham, Research Paper 2015/16.

Gasiorek M. (2019) "Analyzing the impact of a EU-Tunisia DCFTA", FEMISE report, FEM43-16.

Heckscher E. (1919). The Effect of Foreign Trade on the Distribution of Income. *Ekonomisk Tidskrift*, 21: 497–512.

Kahouli B. (2016). Regional integration agreements, trade flows and economic crisis: A static and dynamic gravity model. *International Economic Journal*, 30(4): 450-475.

Larch M., Wanner J., Totov Y.V., Zylkin T. (2019). Currency unions and trade: A PPML re-assessment with high-dimensional fixed effects. *Oxford Bulletin of Economics and Statistics*, 81(3): 487-510.

Levy P.I. (1997). A political-economic analysis of free-trade agreements. *American Economic Review*, 87(4): 506-519.

Linder S.B. (1961). *An essay on trade and transformation*. New York: John Wiley & Sons.

Liu X. (2008). The Political Economy of Free Trade Agreements: An Empirical Investigation. *Journal of Economic Integration*, 23(2): 237-271.

Maddala G.S. (1987). Limited Dependent Variable Models Using Panel Data. *Journal of Human Resources*, 22(3): 307-338.

Mayer W. (1984). Endogenous tariff formation. *American Economic Review*, 74(5): 970-985.

Ohlin B. (1933). *Interregional and International Trade*. Cambridge, MA: Harvard University Press.

Parra M.D., Martinez Zarzoso I. and Suárez-Burguet C. (2016). The impact of FTAs on MENA trade in agricultural and industrial products, *Applied Economics*, 48(25): 2341-2353.

Pehlivan G.G. (2014). The economic growth effects of the Euromed FTA, *Journal of Business Economics and Finance*, 3(4): 425-443.

Péridy N. (2015) Outcome of Fifty Years of Euro-Mediterranean Trade Partnership and DCFTA Perspectives, IEMed Mediterranean Yearbook 2015, European Institute of the Mediterranean, 266-270.

Péridy N., Roux N. (2012). Why are the Trade Gains from the Euro-Mediterranean Partnership so Small? *Journal of World Trade*, 46(3): 571-596.

Ricardo D. (1817). *On the Principles of Political Economy and Taxation*. Batoche Books.

Rodrik D. (1986). Tariffs, subsidies, and welfare with endogenous policy. *Journal of International Economics*, 21(3/4): 285-299.

Verbeek M. (2004). *A Guide to Modern Econometrics*. John Wiley & Sons, Ltd.

Viner J. (1950). *The Customs Union Issue*. New York: Carnegie Endowment for International Peace.

Wooldridge J.M. (2000). *Introductory Econometrics*, Cincinnati, OH: Southwestern.

La pertinence de l'intégration régionale dans la zone Euromed : Résultats à partir d'un modèle logistique d'électeur médian

Résumé – Le modèle de l'électeur médian permet de développer une méthodologie qui évalue la pertinence d'une intégration régionale à partir de la comparaison des utilités relatives avant et après l'intégration. Cet article propose un modèle logit conditionnel à effets fixes qui relie ces utilités relatives à des facteurs explicatifs, tels que les différences de dotations factorielles et de technologies, afin d'estimer la probabilité de l'intégration régionale dans l'espace euro-méditerranéen (Euromed) qui fait l'objet de nombreux débats. Le modèle inclut tous les pays ayant conclu une intégration régionale. La performance du modèle est ensuite testée avec la méthode ROC (Receiver Operating Characteristic). Cette méthode confirme la bonne performance du modèle pour prédire la probabilité de créer une intégration régionale. Une procédure de test à rebours (Reverse Backtesting procedure) est ensuite utilisée pour caractériser la probabilité d'intégration régionale de « impossible » à « hautement probable ». Les résultats montrent qu'avec une probabilité de 82,6%, l'UE est une zone d'intégration régionale « très probable » tandis que l'intégration Euromed est jugée seulement possible (60,5%). De manière générale, toutes les probabilités par paire de pays sont calculées afin de qualifier la pertinence de leur intégration régionale.

Mots-Clés
Intégration régionale
Union européenne
Zone euroméditerranéenne
Modèle logistique de l'électeur médian
Modèle logit conditionnel à effets fixes