WORKING PAPER

DIPARTIMENTO DI ECONOMIA PUBBLICA

Working Paper n. 100

Marianna Belloc

Protection for Sale in the EU

Roma, Luglio 2007

UNIVERSITA' DEGLI STUDI DI ROMA
"LA SAPIENZA"

Protection for Sale in the EU

Marianna Belloc*

This version: March 6, 2007

Abstract

This paper presents estimation of the Grossman-Helpman (1994) model for the EU. We try to address a number of pitfalls that surround the previous empirical literature. First, we suggest a new identification strategy that enables to single out politically organized sectors with specific regard to trade policy. Second, we utilize two alternative measures of trade protection (ad-valorem equivalent of non-tariff barriers and overall restrictiveness) and adopt industrial and trade data at a high level of disaggregation. Finally, we directly obtain the structural coefficients of the model by estimating a non-linear simultaneous equations system that is tightly linked to theory. We find that the theoretical model is broadly consistent with data and conclusions turn out internally coherent.

Keywords: Trade policy; European Union; Lobbying; Political economy; Simultaneous Equations Models

JEL Classification: F13; F15; D71; D72; C31

*Marianna Belloc: Department of Economics, Sapienza University of Rome (Italy). Email: marianna.belloc@uniroma1.it. I would like to thank Dirk De Bièvre, Giancarlo Gandolfo, Justin Greenwood, Alessandro Nicita, and Massimiliano Tancioni for very useful suggestions. I am also indebted to Andreas Dür and Dirk De Bièvre for providing me with the data on the European DG Trade's Civil Society Dialogue. Usual disclaimers apply.

1 Introduction

The protection-for-sale model by Grossman and Helpman (1994)[1] is one of the main reference points for the empirical literature in the political economy of trade. It aims at microfounding the behavior of organized lobbies and politicians in order to derive a clear-cut expression for the level of endogenous protection as a function of industry characteristics. Very different implications for the structure of trade protection are derived depending on whether a given sector is organized in a lobby or not. In particular, the model predicts that: (*i*) protection is higher in organized sectors than in unorganized ones; and (*ii*) protection is a decreasing (increasing) function of the trade penetration ratio and of the trade elasticity in the organized (unorganized) sectors.

The GH model has been tested by a number of studies that provide evidence in support of its predictions: Gawande and Bandyopadhyay (2000) and Goldberg and Maggi (1999) for the United States; Mitra et al. (2002) for Turkey; McCalman (2004) for Australia. Furthermore, in the more recent literature, it has offered the basis to a number of extensions and adaptations for the analysis of related topics: quotas' trade restrictions (Facchini et al. 2006), foreign versus domestic lobbies (Krishna et al., 2006); lobbying for upstream versus lobbying for downstream industries (Gawande and Krishna, 2005); the effects of preferential trade agreements on multilateral trade liberalization (Karacaovali and Limão, 2005); national versus supranational lobbying in the EU (Belloc and Guerrieri, 2007). The mentioned empirical approaches to the political economy of trade policy are however not free from concerns that shelter on data issues, the empirical strategy adopted (especially to identify organized groups), and the consistency of the estimation results with the theoretical framework. In this paper we estimate the GH model for the European Union ($1999-15$ countries). Our empirical strategy is designed to face some of the weaknesses that have been ascribed to previous empirical tests of this model.

The economic literature has recently started to show an increasing attention on European policy formation and interest groups from both theoretical (Broscheid and Cohen, 2003; Cohen, 1997; Crombez, 2002), and empirical (Karacaovali and Limão, 2005; Belloc and Guerrieri, 2007; Balaoing and Francois, 2006; Belfrage, 2004) points of view. Yet, the absence of data on political contributions has so far made it impossible to appropriately identify lobbies within the EU arena[2]. Our paper is the first to attempt an estimation of

[1] Hereafter GH.

[2] For instance, Karacaovali and Limão (2005) modify the GH model to study the effects of preferential trade agreements upon multilateral trade liberalization. However, in their paper, the authors assume that all the industry are organized. The same assumption is made by Belfrage (2004). Belloc and Guerrieri (2007) aim at evaluating

the GH model using European trade and industrial data and employing an identification strategy designed to distinguish organized from unorganized sectors with a special regard to trade policy.

While the European institutional environment departs from that characterizing the United States' (which has inspired the protection-for-sale framework) in a host of different ways, in this work we have decided to not significantly detach from the original theoretical model. It follows that the test conducted in this paper has to be interpreted as a test for the European Union behaving as a single entity that negotiates with its trade partners relying on a common trade policy. This is in line with the reference model presented by the Treaty Establishing the European Community. Article 133 declares: "The common commercial policy shall be based on uniform principles, particularly in regard to changes in tariff rates, the conclusion of tariff and trade agreements, the achievement of uniformity in measures of liberalization, export policy and measures to protect trade such as those to be taken in the event of dumping or subsidies"[3].

Generally speaking, the measures for lobbying action at the EU level are similar to national ones: political pressure, economic contributions, information provision, and so on. However, the policy making environment that characterizes the EU is more complex as it relies on a very open and decentralized decision process (Mazey and Richardson, 1993). Before the adoption of the Single European Act in 1986, the policy-making was mainly in the hands of the Council of Ministers and, as a consequence, most of the EU lobbying was wielded by national groups through political and administrative channels. After 1986 this situation has changed: the Commission was endowed with the power to initiate EU policies and became a crucial actor in the formulation of policy proposals. This entailed a special need for coalition-building at the EU level at the early stages of the lobbying process. Unlike the Council, the Commission is meant to be a supranational body. As a consequence, commissioners are supposed to be independent of member states in taking their decisions. Furthermore, the Commission is essentially a technical bureaucracy that resorts in large measure to private actors to gather information needed to draft legislation. Thus, lobbies represent an important source of grass-roots information and play a very active role during the legislative process: they bring issues to the policy makers' attention, provide information, and often take part in the committees (Directorates-General) that assist the

the distinct roles of national and supranational special interest groups in the EU. They approximate the political organization dummy resorting to proxies such as the industry scale.

[3]Furthermore, Belloc and Guerrieri (2007) find that the effects of national and European lobbies on protection are very similar.

Commission in preparing proposals (Broscheid and Cohen, 2003). The previous considerations suggest that participation in official meetings on trade issues is very informative to track lobbying for trade policy at the European level. This motivates our identification strategy as it is explained in the next sections.

The remainder of this paper is organized as follows. Section 2 summarizes the main predictions of the GH model and analyses the previous approaches used to test them. Also it reviews some of the critiques moved against these tests and presents how they are intended to be overcome here. Section 3 introduces the empirical strategy and the data used in the present work. Sections 4 and 5 present respectively the core results and the robustness checks. Section 6 draws the conclusions.

2 Previous literature

2.1 The GH model

The intriguing feature of the GH model is the microfoundation of the behavior of organized lobbies and politicians to derive a clear-cut expression for the level of endogenous protection as a function of industry characteristics. While the GH model is too well known to be repeated here, in what follows we recall its most important elements.

There is a small open economy and the rest-of-the-world. Among the several goods produced, one (the numeraire) is obtained using only labor, while all the others require labor and a sector specific factor. The economy is populated by a number of individuals identical in every relevant respect but factor endowments: all of them own labor but only some also possess a specific factor (at most one). Each individual is welfare maximizer. Her income consists of three elements: labor income, government transfer, and, possibly, rewards from the ownership of the specific factor. Specific factor owners gain from higher protection in their sector. Then, a number of them manage to get organized in order to sheer government's decisions in their favor. To this purpose the lobby group can spend a part of her resources in political contributions that are perceived by the government. Hence, while the lobbies maximize their net-of-contributions welfare, the government aims at maximizing the weighted sum of aggregate general welfare and the contributions from the pressure groups, that is:

$$G = \sum_{i \in L} C_i + aW \tag{1}$$

where W is gross-of-contribution social welfare, C_i is the contribution schedule for group i, L is the number of groups organized in lobbies and a the relative weight that the government attaches to aggregate welfare with respect to special interests. The policy game can be solved in two steps. In the first stage the organized interest groups define the amount of resources to allocate to lobbying action, taking as given the decisions of the other groups. In the second stage, after including the contribution schedules in her welfare function, the government defines the level of trade protection. This is finally given by:

$$\frac{t_i}{1+t_i} = \frac{I_i - \alpha_L}{\alpha_L + a} \cdot \frac{z_i}{e_i} \tag{2}$$

where t_i is a measure of trade protection, I_i is a dummy variable that equals one if sector i is organized and zero otherwise, z_i is the inverse of the trade penetration ratio (production over imports -exports- for import -export- competing industries), e_i the import-demand (export-supply) elasticity. α_L and a are the structural parameters of interest: the former represents the fraction of voting population that owns a specific factor, while the latter is the weight that the government attaches to aggregate social welfare.

Three main predictions can be obtained from equation (2): (*i*) protection is higher in organized sectors than in unorganized ones; and (*ii*) protection is a decreasing (increasing) function of the trade penetration ratio and of the trade elasticity in the organized (unorganized) sectors[4].

2.2 Previous empirical approaches and related issues

The GH model has been verified by a number of works considering different countries and adopting alternative econometric techniques (see section 3). These tests have however risen concerns and critiques that dwell upon data, methodological issues as well as problems of internal coherence for the results obtained (Imai et al. 2006; Gawande and Krishna, 2003; Mitra et al., 2006). They are briefly presented in this subsection where we also explain how we intend to deal with them.

A) Political organization variable

Let us start with concerns regarding the data used in the empirical tests. The first is related to the identification of industries that are politically organized to sheer trade policy. Since a political organization variable is not directly observable, the construction of the dummy I_i used in equation (2) must largely rely on subjective elements. Gawande and Bandyopadhyay (2000) and Goldberg and Maggi (1999) resort to corporate campaign contributions and define as organized industries that allocate resources to political contributions over a certain cutoff level. This choice presents a major problem: we cannot clearly

[4] See Helpman (1995) for a comparative evaluation of these predictions with those obtained by previous political economy approaches to trade policy.

distinguish the amount of money that lobbies address to sway trade policy's decisions from those that are directed to other aims (for instance domestic support not related to trade). To deal with this, Gawande and Bandyopadhyay (2000) run an auxiliary regression to associate contributions to purely trade-related variables. However, in our case contributions data simply do not exist, as they would denote illegal activities in European countries. Then we have to come out with a different strategy. With regard to the EU, Karacaovali and Limão (2005) circumvent the problem assuming that all the industries are organized, while Belloc and Guerrieri (2007) use proxy variables that suffer from the same limitation above illustrated. Outside the EU, another couple of options are suggested by Mitra et al. (2002) and McCalman (2004). The former first map the membership to one of the most important Turkish industrialist organizations (the TUSIAD[5]); then validate the identification made by statistical discriminant and probit estimation techniques (see subsection 3.3). Again the identification cannot capture pure trade policy concerns being the TUSIAD a general (both domestic and international) interest business organization. Finally, McCalman (2000)'s strategy is based on information by Australian trade policy institutions. Considering the Australian practice of reviewing any operation of trade policy by an independent advisory body (the "Tariff Board"), the author classifies an industry as organized if a Tariff Board report was prepared in the given period. While this strategy has the merit to only focus on trade policy, it remains specific to Australia. Our identification methodology is specifically designed to fit the EU institutional environment.

We use data from the EU Directorate-General for Trade[6] to single out organizations registered to take part in the Civil Society Dialog[7]. This holds regular meetings on external trade matters between the European Commissioner for Trade, senior Commission officials and trade negotiators. In this way we attain two goals. On the one hand, we tackle a specific feature of the EU institutional arena, where lobbying is wielded mainly at the early stages of the policy formation by information provision to and negotiations with the European Commission. On the other, we are able to focus on interest groups with pure trade policy concerns. The evident weakness of this strategy is that we can consider only pressure made at the European level and overlook the national channels. Yet, as explained above, this choice is in line with the practice following the ratification of the Single European Act in 1986 (see for instance, Mazey and Richardson, 1993).

[5] Turkish Industrialists and Businessman Association.

[6] Hereafter DG-Trade.

[7] The original dataset that summarizes information on the registered organizations has been compiled by Dür and De Bièvre (2006). More details on this below.

B) Measure of trade protection

Second, the dependent variable in the original GH (1994) model is tariffs. However, while in the theoretical framework the policy maker acts as a noncooperative player, in the real world tariffs are set cooperatively within the WTO arena. This observation has led the protection-for-sale empirical literature to focus on non-tariff barriers (NTBs) as a measure for protection. The choice between tariff and non-tariff barriers as right-hand variable is fraught with important implications (Gawande and Krishna, 2003). Indeed, pure NTB coverage ratios are likely to exhibit large measurement errors that can lead to either over-estimate or under-estimate the extent of protection. On the other hand, adopting tariffs alone would imply, beyond inconsistency with the noncooperative nature of the policy game, large imprecisions given the great use of new trade protectionist measures made by national governments in the recent periods. In this paper we employ ad-valorem equivalents of trade protection constructed by Kee et al. (2006). This measure has two major advantages: first, restrictiveness is estimated at the tariff line level, that is the level of disaggregation at which trade policy is set. Protection measured at the aggregated industry level would indeed be likely to generate severe aggregation bias. Second, Kee et al. (2006) calculate, besides the ad-valorem equivalent of non-tariff barriers, that of overall trade restrictiveness which includes both tariff and non-tariff measures. This is expected to capture the true extent of protection at an acceptable degree of precision even in the presence of combination of old and new protectionist measures.

C) Measure of the trade price elasticities

Third, the estimation of model (2) requires data on import-demand (export-supply) elasticities that in turn must be estimated. Trade price elasticities used in this paper are provided by Kee et al. (2004) and are suitable for our study for a number of reasons. First, they are disaggregated at the tariff line level (more than 4000 products in our case) so to maintain consistency with the level of disaggregation of the protection measure and with the trade policy formation process. This is important since we expect the trade elasticities to be the larger, the higher the level of disaggregation. As a consequence, using more aggregated data for import-demand (export-supply) elasticities than for protection would lead to an under-estimation of the welfare costs of protectionist trade policies. Second, they are obtained by a methodology that is coherent with trade theory and by the same data sources for each country (then for all the 15 countries included in our investigation) and good considered. Finally, the estimates obtained present a good degree of accuracy. Considering the overall study by Kee et al. (2004) (that includes 117 countries), 89% of the elasticities are statistically significant at the 1% level, 91% at the 5% and 93% at the 10%.

D) Estimation, results and their interpretation

The last issue is a general concern about the political economy approach to trade theory in the light of the estimation results obtain by the already mentioned empirical studies. As emphasized by some critical articles (Imai et al., 2006; Gawande and Krishna, 2003; Mitra et al., 2006), but also noted by the original empirical works on the GH model (e.g. Gawande and Bandyopadhyay, 2000), the estimated value of a is very large (for instance for the United States it ranges between 100 and 3000). The larger is a, the more the government behaves as a welfare maximizer, and the less she cares about political pressure by interest groups. Although this results does not invalidate the theoretical framework since no clear expectations are suggested on the magnitude of a, the fact that the government seems not very exposed to political pressure lobbying hints that the political economy approach is not the best choice for trade policy analysis. Actually, a large value of a is not a discouraging result in itself as far as $1/a$ is statistically different from zero (what means that the government distinguishes between special and general interests). But, a large value of a is not realistic if associated to the large estimates of the parameter α_L that come out from the previous empirical studies (between 0.85 and 0.98 for the United States). According to Mitra et al. (2006) the internal contradiction of this conclusion hinges on the unappealing result that a so large fraction of the voting population (between 85% and 95%) finds the resources to get organized in the presence of a government that is not exposed to the political pressure they exert. Mitra et al. (2006) suggest two major reasons for this unrealistic conclusion. On the one hand, the empirical strategy must be more tightly linked to theory, that is, equation (2) should not be contaminated with additional terms not included in the GH model. On the other, all the sectors should be treated as organized as they all enjoy a positive (or non-negative) protection in the various countries object of study. The latter comment is also tied to point A), that relating to the difficulty in singling out the industries that wield their influence for exclusive trade policy concerns.

While in this paper we carefully follow the former suggestion in designing our econometric strategy (see subsection 3.1), we have chosen to distinguish organized from unorganized sectors. Indeed, as already noted, our data from the DG-Trade allow us to identify those organizations that take part in the European policy negotiations with a specific interest for external trade policy. This not only makes our investigation more interesting, but also enables us to simultaneously identify the two parameters a and α_L. Without an identification strategy for lobby groups (i.e. setting the variable I_i always equal to one in equation (2)) we could only estimate the parameter on z_i/e_i (let us say $b = (1 - \alpha_L) / (\alpha_L + a)$) and infer one between α_L and a after setting the other equal to an exogenously chosen constant (see Mitra et al., 2006).

3 Empirical strategy

3.1 Econometric model

Estimation of model (2) presents some issues that we need to tackle carefully. First, the trade price elasticity (e_i), entering model (2) as an explanatory variable, is not observed, but is estimated with possibly severe measurement errors. Two strategies have been suggested by the literature to deal with this: either moving e_i on the left-hand side as in Goldberg and Maggi (1999) and McCalman (2004), or keeping it on the right-hand side and using instrumental variables estimation as in Gawande and Bandyopadhyay (2000) and Mitra et al. (2002). Even if we believe the trade price elasticities provided by Kee et al. (2004) are estimated at an good degree of precision for the reasons exposed in subsection 2.2.C, we cannot exclude the existence of measurement errors. To deal with this, we have decided to adopt the first strategy; this is mainly motivated by the choice illustrated below of implementing simultaneous equations system.

Second, z_i and I_i can turn out to be endogenously determined in the model. Again two approaches are possible. Simultaneous equations system estimation is chosen by Goldberg and Maggi (1999) (FIML[8]) and Gawande and Bandyopadhyay (2000) (Kelejian, 1971,'s 2SLS[9] estimator). Single equation and instrumental variables estimation is employed by Mitra et al. (2002) (both 2SLS and panel-GMM[10]) and McCalman (2004) (2SLS). We have decided to use simultaneous equations estimation by FIML, but, departing from the previous studies adopting this methodology, we directly estimate model (2) in its nonlinear form (for FIML estimation of nonlinear equation systems see Chow, 1973). Despite the additional computational and specification costs, the evident advantage of applying nonlinear econometric techniques consists in immediately obtaining the structural parameters and their standard errors. Furthermore, under correct model specification, FIML estimation is fully efficient.

[8] FIML stands for full information maximum likelihood.

[9] 2SLS stands for two-stage least squares.

[10] GMM stands for generalized method of moments.

We estimate the following model:

$$\frac{t_i e_i}{1+t_i} = c_0 + \frac{1}{\alpha_L + a} \cdot I_i \cdot z_i - \frac{\alpha_L}{\alpha_L + a} \cdot z_i + \varepsilon_i \qquad (3)$$

$$\ln z_i = c_1 + \beta_1^{'} \ln \Delta_{1i} + \zeta_{1i} \qquad (4)$$

$$I_i = c_2 + \beta_2^{'} \ln \Delta_{2i} + \zeta_{2i} \qquad (5)$$

where c_0, c_1 and c_2 are constant terms, ε_i, ζ_{1i} and ζ_{2i} are the additive residual terms, Δ_{1i} and Δ_{2i} are vectors of industrial variables, β_1 and β_2 are vectors of parameters, and finally " $'$ " stands for transposed. Although the intercept, c_0, is not directly derived by the theory in equation (3), it is useful to statistically justify the fact that all sectors (either organized or not) are associated to a positive (or non-negative) level of protection. The system formed by equations (3), (4) and (5) is estimated by FIML.

3.2 Data description

Data on protection are from Kee et al. (2006) and are at the 6-digit of the Harmonized System[11]. We employ two measures of protection: ad-valorem equivalent of core NTBs and ad-valorem equivalent of overall protection (tariff, domestic support and NTBs).

The main problem in measuring trade protection barriers is that trade policy can take a host of different forms: tariffs, quotas, antidumping duties, technical regulations, monopolistic measures, subsidies, and so on. To solve this problem it is necessary to convert all the information contained in the non-tariff impediments to trade into an ad-valorem equivalent that is directly comparable to a tariff. This goal is attained by Kee et al. (2006) mainly following Anderson and Neary (see in particular 1994 and 2004). They rely on two broad classes of NTBs: the so called "core NTBs" (namely price and quantity control measures, technical regulations, monopolistic measures) and agricultural domestic support. For each of them Kee et al. (2006) estimate the effects on imports adopting the comparative advantage approach. The latter has its theoretical foundation in a multi-good multi-factor general equilibrium model with log-linear utilities and log-linear constant returns to scale technologies (see in particular Leamer, 1990). The obtained quantity impact is then transformed into a price equivalent by using the trade elasticities. This gives directly the first measure of protection that we employ in this paper (*NTB-core*). The second is a composite index that combines the ad-valorem equivalent of core NTBs with ad-valorem equivalent of the domestic support and tariff barriers (*overall-protection*).

[11] Harmonized Commodity Description and Coding System (HS) is a universally accepted classification system for tradable goods. The core system identifies basic commodities by a 6-digit number. A complete list of HS codes and product descriptions is available at http://www.foreign-trade.com/reference/hscode.htm

Import-demand (export-supply) elasticities are provided by Kee et al. (2004) and are disaggregated at the 6-digit of the HS.

The variable z_i is defined as the ratio of domestic output over imports or exports depending on whether the sector is import or export-competing. External trade series are obtained from the Eurostat. They consist in *extra*-EU trade statistics and cover the trading of goods between a member state and a non-member country. For "goods" are intended all movable properties. Data are at the 6-digit level of disaggregation according to HS.

To calculate z_i we also need data on production. We faced two possible options: to take production series from the Eurostat - Industry, Trade and Services database that provides quite complete series aggregated at the 4-digit level according to NACE Rev 1.1[12] (as Karacaovali and Limão, 2005), or get it from the Eurostat - PRODCOM database that offers more incomplete series but disaggregated at 8-digit according to the PRODCOM[13] classification system. Using the latter rather than the former leads us to lose several observations and expose us to some problems of reclassification but allows to exploit the more detailed information embedded in the disaggregated data. The PRODCOM classification is directly linked to HS so to maintain the same level of disaggregation as the trade series. Since the number of observations we have at our disposal using the PRODCOM data is very large (1597 products), we decided to opt for this dataset[14]. In this way we can obtain more accurate estimation results. Nonetheless, as a robustness check, we also replicate the estimation using the more aggregated production data (see section 5). Conclusions do not change in any significant way.

Finally we need data on industrial characteristics employed entering regressions (4) and (5). They are provided by the Eurostat by country and detailed on NACE Rev 1.1 at 4-digit level. The conversion to HS is made possible by international systems concordance tables[15]. Industry data are more aggregated than trade data. We are however not very concerned with this issue since (4) and (5) are not the primary object of our empirical analysis. Yet, since errors in the estimation of these equations are likely to affect the estimation of

[12] NACE Rev. 1.1 is a 4-digit activity classification used by the European Union since 2002. More details are available at: http://ec.europa.eu/eurostat/ramon/relations/index.cfm

[13] PRODCOM is a system for the collection and dissemination of data on the production of manufactured products. PRODCOM headings are derived from HS; this enable comparisons between production and foreign trade data.

[14] We implemented reclassification from PRODCOM to HS by concordance tables available at: http://ec.europa.eu/eurostat/ramon/relations/index.cfm

[15] The classification of economic activities according to NACE Rev. 1.1 corresponds to ISIC Rev.3 at European level. Even though the former is more disaggregated than the latter, it is totally coherent with it and can be considered its European counterpart. We have then first passed from NACE 1.1 to ISIC 3 and then from ISIC 3 to HS. Concordance tables from NACE to ISIC 3 are available at: http://www.fifoost.org/database/nace/nace-en_2002c.php, whereas correspondence tables from ISIC 3 to HS are courtesy of Colin Webb (OECD).

equation (3), we control for the accuracy in the estimation of each equation in the model and, as a robustness check, we also estimate regression (3) alone obtaining unchanged qualitative conclusions. Industrial variables included are: measure of industry scale; share of turnover in production value; investments in long-term rentals, investments in tangible goods, investments in land, investments in existing buildings, investments in constructions; labor share; percentage of agency workers, unpaid persons, apprentices, R&D personnel, and employees in total of persons employed; share of social security costs in total of wages and salaries; value added per person employed; share of personnel costs in production value; wage; share of production value in total manufacturing, share of persons employed in total manufacturing, and percentage of enterprises in total manufacturing.

The construction of the political organization dummy is described in subsection 3.3.

We conclude this subsection with some notes on the sample used and the aggregation procedures. Our study considers the European Union in 1999 consisting in (ordered by accession date): Belgium, France, Italy, Luxemburg, Netherlands, Germany, Denmark, Ireland, United Kingdom, Greece, Portugal, Spain, Austria, Finland, and Sweden. The model is estimated for the EU as a whole. Tariffs and NTB ad-valorem equivalents are already provided from source for the EU as a whole, as well as imports and exports (that is extra-EU trade). Production and total employment are aggregated by simple sum. Finally, ratios (for instance ratios over GDP or total employment) and elasticities are aggregated by weighted sum. The weights used in the aggregation are given by the share of the individual country's constant GDP (millions of euro at 1995 prices and exchange rates) over total GDP for EU15 in 1999[16].

3.3 Political organization dummy

The political organization dummy is not observable. Then we need to come out with a strategy that enables us to identify sectors that are organized to wield pressure on the European government with regard to trade policy decisions. To attain this goal, we have proceeded in two steps. First, we construct some priors on the base of data from the European Commission DG-Trade as explained below, then we validate our identification by means of discriminant function analysis, cluster analysis and probit estimation techniques[17].

1) The Civil Society Dialogue-External Trade (European Commission DG-Trade) holds

[16] Weights used in the aggregation are: Belgium (0.0322), Denmark (0.0208), Germany (0.2781), Greece (0.0137), Spain (0.0716), France (0.1786), Ireland (0.0103), Italy (0.1231), Netherlands (0.0506), Austria (0.0276), Portugal (0.0138), Finland (0.0162), Sweden (0.0290), UK (0.1319), Luxemburg (0.0026).

[17] A similar strategy is also used by Mitra et al. (2002).

regular meetings on trade issues in Brussels with the European Commissioner for Trade, senior Commission officials and trade negotiators. The main goal of these meetings is to foster confident working relationship between all interested stakeholders in trade policy issues and to gather the grass-roots information that the Commission necessitates to draft legislation. The DG-Trade maintains a database of all the civil society organizations registered to take part in the Civil Society Dialogue - External Trade[18]. Dür and De Bièvre (2006) collected data on all the organizations in the database including information on whether or not it is European-wide, i.e. representative of a membership or stakeholders are from all over the EU. Relying on this set of data we have coded the organizations included in the database according to the ISIC Rev. 2 system at the 3-digit level. If, in a given sector, there are at least five European-wide organizations registered in the Civil Society Dialogue External Trade, the dummy I_i is set equal to one; it is zero otherwise[19]. Since the other variables used in the system estimation are classified at the 6-digit of the HS we have implemented a reclassification from ISIC 2 to HS[20]. Here, the fact that the political organization dummy is more aggregated than the trade variables should not create problems as producers, especially at the European level, have great advantages from organizing at the industry level to lobby for protection. This is confirmed by the fact that, in the EU, there is more variation in protection across industries than within them (on this point see also Karacaovali and Limão, 2005).

In the second stage of our identification strategy we validate our priors defined in point 1) on the bases of three techniques.

2.a) The discriminant function analysis (DFA)[21] is a statistical tool to classify cases into the values of a categorical dependent variable or to test a theory by observing whether cases are classified as predicted. The ingredients are the dependent variable (in our case the political organization dummy, I_i, obtained as explained above), and n discriminant variables (in our case the 20 industrial characteristics included in Δ_{1i} and Δ_{2i}, let us say x_1, x_2, $...x_{20}$). The categorical variable, or latent variable, is predicted by the discriminant variables by means of a discriminant, or canonical, function that is as follows:

$$I_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{3i} + ...b_n x_n$$

where b_k $(k = 0, 1, 2, ...n)$ are the discriminant coefficients, that looks like a multiple regression function. The difference between the discriminant function and a multiple regression

[18] More information on the registered organizations is available at http://trade.ec.europa.eu/

[19] We have also experienced with alternative cutoff levels; our final conclusions are not altered in any relevant way.

[20] Concordance tables are available at http://www.macalester.edu/research

[21] For a detailed treatment of the DFA see for instance Huberty (1994).

is that the latter maximizes the regression relationship, whereas the former maximizes the distance between the groups as initially classified in the categorical dependent variable. Then the classification scores are computed, which means the predicted values of the latent variable are obtained by applying the discriminant function to the data for a given group at a time. In symbols:

$$S_{ij} = \widehat{b}_{0j} + \widehat{b}_{1j}x_{1j} + \widehat{b}_{2j}x_{2j} + \widehat{b}_{3j}x_{3j} + ...\widehat{b}_{nj}x_{nj}$$

where S_j is the score of group j, x_{kj} $(k = 0, 1, 2, ...n)$ is variable x_{kj} as observed for group j, and $\widehat{b}_{kj}$ $(k = 0, 1, 2, ...n)$ the estimated discriminant coefficient for group j.

We are now interested in finding out if the model is discriminating, that is if the two groups (organized and unorganized sectors) as they have been classified according to our priors are sufficiently different. This information is provided by the Wilks' lambda given by the proportion of the total variance in the discriminant scores not explained by differences among the groups. It is distributed as a chi-square with the number of degrees of freedom equal to the number of discriminant variables. A statistically significant lambda induces to reject the hypothesis that the two groups have the same mean discrimination score. Furthermore, to test the validity of our preliminary classification we compare pairwise the predicted categorical variable (I_i^*) with our political organization dummy (I_i). Prediction is produced as follows: if the discriminant score of the function for a certain object (in our case a product-line) is less or equal to the cutoff, the case is classified as zero (it goes into the unorganized group), as one otherwise (it is included among the organized sectors). We obtain the following results:

Table 1. Discriminant function analysis

Wilks' lambda	0.3914	
Chi-sq (p-val)	3800***	(0.0000)
Correctly predicted	90.82%	

***Significance at 1%

As one can notice, the two groups are statistically different. Furthermore our priors are validated by the model in the 90.82% percent of cases.

2.*b*) The second validation method is according to cluster analysis[22]. The way of proceeding is similar to the previous but does not require an *a priori* structure of the groups. Briefly, the cluster analysis is an exploratory data analysis tool with the aim of sorting different objects into groups maximizing (minimizing) the degree of association between

[22] For a detailed treatment we refer the reader to Everitt et al. (2001).

objects that belong to the same (different) group. In our case we suggest a number of industrial characteristics that may influence the status of a sector of being organized in lobby or not, without assuming an *a priori* categorization. Then the objects (our product-lines) are divided into two groups, let us say a meaningful structure, on the bases of similarities (dissimilarities) in such characteristics. Once obtained the new categorization we compare it with our prior definition of I_i and evaluate the percentage of pairwise matching. We verify that our measure of the political organization variable as suggested by the DG-Trade data is validated in the 61.46% of cases.

2.c) Finally we deal with probit estimation and predictions. We estimate a probit regression of I_i (as constructed in point 1) on the already mentioned industrial characteristics included in Δ_{1i} and Δ_{2i}. Then we generate the prediction using a cutoff probability of 0.6 and compare pairwise the prediction with our measure of political organization. We obtain that prediction is accurate in the 92.95% of cases.

4 Estimation results

In this section we report estimation results for model (3)-(5) obtained by FIML techniques. Estimation is repeated using as dependent variable either the ad-valorem equivalent of NTB-core or the ad-valorem equivalent of the overall protection measure.

We start by commenting the model for NTB-core; results are in table 2. The estimated value of α_L turns out to be equal to 0.4653 whereas the relative weight that the government attaches to aggregated social welfare is 148, both being statistically significant at any conventional level. This conclusion hints that the GH model fits well the EU environment. Our results suggest that the percentage of organized groups is around 47% of the voting population. This value seems realistic and coherent with the value of a. Indeed even if European industrial groups have a large incentive to look for alliances and coalitions given the complex and decentralized institutional environment they have to face (Mazey and Richardson, 1993), the lobby formation at the EU level is a quite recent phenomenon. Hence it is reasonable that the share of organized groups turns out much smaller than that estimated for the US by previous studies (that falls in the range $85\% - 95\%$)[23]. With regard to the estimated value of a, we observe that, coherently with that found by the previous studies, it is "large" (in the sense explained in subsection 2.2.D). Hence, we test the hypothesis that the government is a pure welfare maximizer, i.e. that $1/a$ equals zero.

[23] With regards to the EU, Belloc and Guerrieri (2007) obtain that α_L equals 0.75. We remark that, however, in their model α_L represents the percentage of national *plus* European lobbies in total population, then it is not comparable with that estimated in this paper.

Table 2. FIML estimation: $\mathbf{t}_i$ *= NTB-Core*

Dependent variable:	$\mathbf{t}_i\mathbf{e}_i/(\mathbf{1}+\mathbf{t}_i)$		$\ln \mathbf{z}_i$		$\mathbf{I}_i$	
	coeff.	**s.e.**	**coeff.**	**s.e.**	**coeff.**	**s.e.**
Constant	0.0805***	0.0134	1.1430***	0.1463	-0.0018	0.2452
α_L	0.4653***	0.0658				
a	147.9427***	17.5316				
Scale			-0.2902***	0.0919	0.1006***	0.0368
Turnover in prod value			0.4170*	0.2249	-0.1958**	0.0837
Long-term rentals			-0.9394***	0.2942	0.2309**	0.0951
Tangible goods			0.2433**	0.1168	0.0446	0.0518
Land			1.8222***	0.4291	-0.4587***	0.1674
Existing buildings			0.2465	0.3070	-0.1517	0.1201
Constructions			-0.3166**	0.1526	0.1066	0.0775
Labor share			-0.7129*	0.3644	1.4688***	0.1085
Agency workers			0.4951***	0.0809	-0.4193***	0.0378
Apprentices			0.1017	0.1316	-0.3068***	0.0528
Unpaid persons			-0.3891***	0.0860	0.0800**	0.0361
R&D personnel			-0.4736***	0.0535	-0.1269	0.0233
Employees			0.0775	0.3609	-0.1424	0.1500
Social security			-1.2198***	0.4494	-0.2209***	0.1618
Apparent productivity			0.3484	0.3674	0.5435***	0.1329
Personnel costs			1.0646***	0.1938	-0.7994***	0.0773
Wage			0.3197	0.5481	-0.5767***	0.2004
Share prod value			2.0534***	0.2865	-1.0606***	0.1273
Share persons employed			-2.3800***	0.3635	1.1312***	0.1545
Share enterprises			29.7059***	5.0943	-5.1392***	1.7190

Observations = 1597

Log likelihood = -3346.514

Akaike information criterion = 4.4363

Schwarz information criterion = 4.5962

***Significance at 1%; **Significance at 5%; *Significance at 10%

Output is in table 4 showing that the null hypothesis is rejected at any confidence level. Also we test the hypothesis that the fraction of voting population organized in lobbies is zero, either taken individually or jointly with the previous one. In both cases, we are led to reject H_0. Finally we verify if the sum of the coefficients over z_i/e_i and $I \cdot z_i/e_i$ is null; we conclude that $(1-\alpha_L)/(a+\alpha_L)$ is not statistically equal to zero.

To confirm validity of the whole model, we further control if the estimation results for equation (5) are susceptible of a sensible interpretation. In particular we explore consistency with predictions suggested by three main theories for endogenous lobby formation (for a survey see for instance Gawande and Krishna, 2003; and Gawande, 1998): pressure groups and collective action (Olson, 1965), adding machine (Caves, 1976), and social justice (Ball, 1967; Constantopoulos, 1974; Fieleke, 1976). *Pressure groups*: industries differ in their ability to coordinate and get organized in lobbies due to free-riding problems. Accordingly we expect the probability of a sector to get organized to be the larger, the smaller the number of enterprises in that sector. As one can notice in table 2, the number of enterprises has indeed a negative effect on the dependent variable and is statistically significant at any conventional level. Second, we expect the probability for a sector to be organized to be the larger, the larger the industry scale (measured by value added per enterprise). The bigger industry size is associated with larger stakes involved in cooperation among producers to ask for protection. The larger the industry asking for protection, the greater is the incentive to take part in the tariff-setting process. Again this prediction is verified by data as the coefficient associated to industry scale is positive and statistically significant at the 1% level. *Adding machine*: a larger number of employees is associated to a larger number of votes in the elections, then lobbying is likely to be more effective. Consistently we find that the value of the coefficient associated to the share of employment in total manufacturing is positive and statistically significant at any confidence level. *Social justice*: altruistic government officials have a special regard for lowest income groups. Then the ability for a sector to sheer political favor is likely to be the larger, the more numerous low-income workers and the larger the labor share in that sector. Consistently we find that the probability of lobbying is linked positively with the number of apprentices, the number of agency workers, and with the labor share, while it is negatively associated with the average wage. The other variables included in the equation are mostly controls as there are not unambiguous theoretical predictions on the effects they exert on the probability of lobbying.

Table 3. FIML estimation: $\mathbf{t}_i$ *= Overall-protection*

Dependent variable:	$\mathbf{t}_i\mathbf{e}_i/(\mathbf{1}+\mathbf{t}_i)$		$\ln \mathbf{z}_i$		$\mathbf{I}_i$	
	coeff.	**s.e.**	**coeff.**	**s.e.**	**coeff.**	**s.e.**
Constant	0.1072***	0.0177	1.1205***	0.14774	-0.0126	0.2618
α_L	0.3446**	0.1414				
a	112.6851***	23.7256				
Scale			-0.2740***	0.0927	0.1080***	0.0379
Turnover in prod value			0.5876***	0.2193	-0.1281	0.0972
Long-term rentals			-1.0527***	0.2977	0.1649*	0.1000
Tangible goods			0.2678**	0.1179	0.0620	0.0503
Land			1.9435***	0.4309	-0.3936**	0.1673
Existing buildings			0.2616	0.3108	-0.1462	0.1273
Constructions			-0.3559**	0.1542	0.0823	0.0760
Labor share			-0.6100*	0.3684	1.5564***	0.1083
Agency workers			0.4287***	0.0818	-0.4515***	0.0381
Apprentices			0.0426	0.1334	-0.3472***	0.0520
Unpaid persons			-0.3970***	0.0857	0.0785**	0.0374
R&D personnel			-0.5070***	0.0537	-0.1455***	0.0232
Employees			-0.1376	0.3334	-0.2364	0.1623
Social security			-1.3927***	0.4537	-0.3288**	0.1771
Apparent productivity			0.4462	0.3680	0.5935***	0.1349
Personnel costs			1.1244***	0.1915	-0.7809***	0.0866
Wage			0.2933	0.5524	-0.5800***	0.2114
Share prod value			1.8981***	0.2928	-1.1231***	0.1360
Share persons employed			-2.1040***	0.3664	1.2048***	0.1615
Share enterprises			29.4166***	5.0861	-5.0413***	1.7234

Observations = 1597

Log likelihood = -4047.975

Akaike information criterion = 5.3538

Schwarz information criterion = 5.5107

***Significance at 1%; **Significance at 5%; *Significance at 10%

Furthermore, we re-run the system estimation adopting as measure of trade protection the ad-valorem equivalent of the overall trade restrictiveness index, which includes both NTB barriers and tariffs. Results are shown in table 3 and, as one can notice, are very similar to those obtained for the very NTB trade protection measure. The values estimated for respectively α_L and a suggest that around the 34% of the voting population succeeds in organizing in lobbies, and the political weight that the government attaches to aggregated social welfare is about 113. Also in this case we have conducted hypothesis tests that are reported in table 4. Again we are led to the same conclusion as above, i.e. to reject the hypothesis that α_L and $1/a$ are equal to zero, either individually or jointly considered.

Table 4. Hp.s testing

	NTB-core		**Overall protection**	
$\mathbf{H_0}$	**Chi-sq**	**p-val**	**Chi-sq**	**p-val**
$1/a = 0$	71.2108***	0.0000	22.5574***	0.0000
$\alpha_L = 0$	50.0629***	0.0000	5.9384**	0.0148
$1/a = 0,\ \alpha_L = 0$	71.3364***	0.0000	59.7329***	0.0000
$(1 - \alpha_L) / (a + \alpha_L) = 0$	183.4740***	0.0000	188.8440***	0.0000

***Reject H_0 at 1%; **Reject H_0 at 5%.

A possible criticism to the employed estimation strategy is that there may be other variables beyond the trade penetration ratio and the trade price elasticities that directly affect protection. To address this point, we also experiment with alternative specifications of model (3)-(5) including additional regressors in equation (3) and namely: measure of industry scale, labor share, share of employees in total of persons employed, apparent productivity, growth rate of imports and total number of persons employed. The exercise is repeated using alternatively either NTB-core or overall-protection as measure for protection. We verify that, on the one side, previous qualitative conclusions on the structural coefficients remain valid (NTB-core: $a = 134$ with $s.e. = 14.8322$, and $\alpha_L = 0.4700$ with $s.e. = 0.0637$; overall-prot: $a = 107$ with $s.e. = 23.8945$, and $\alpha_L = 0.3510$ with $s.e. = 0.1495$), on the other, none of the additional variables is statistically significant either individually or jointly considered (chi-squared statistics for the Wald test of joint significance and the correspondent p-value equal respectively: NTB-core: $\chi^2_5 = 7.5828$ and $p{-}val = 0.2703$; overall-prot: $\chi^2_5 = 9.38$ and $p{-}val = 0.1530$). Furthermore, both the Akaike Information Criterion (NTB-core: $AIC = 4.4620$; overall-prot: $AIC = 5.3741$) and the Schwarz Information Criterion (NTB-core: $SIC = 4.6455$; overall-prot: $SIC = 5.5576$) fail to reject the original formulations of the model.

Finally, to test for heteroskedasticity we implement a parsimonious version of the generalized White test (see Kelejian, 1982) by regressing the squared residuals from the estimated model (3)-(5) on a constant, the fitted values of the dependent variables and the squared of the latter (note that these are functions of the regressors in the original model). The test is parsimonious in the sense that, given the large number of regressors in our specification, we do not include the cross-product terms in order to conserve on the degrees of freedom. The statistics NR^2 (where the R^2 is that obtained from the auxiliary regression, and N is the number of included observations) is distributed as a chi-squared with two degrees of freedom. In the case of NTB-core the NR^2 statistics turns out to be 1.3761 ($p\text{-}val = 0.5026$), while when we use the overall restrictiveness index it is equal to 0.4465 ($p\text{-}val = 0.7999$). Hence, we are led not to reject the null hypothesis of no heteroskedasticity.

5 Robustness checks

In this section we challenge the robustness of our core estimation results in a host of different ways.

First we test if the estimation results are sensitive to the choice of the variables used in regressions (4) and (5) whose specification, even if suggested by the theory, is not directly derived by a structural theoretical model. We conduct a model specification search by dropping one variable at a time in the two equations and replicating the simultaneous system estimation (in other words we replicate the estimation 20 times for 19 different combinations of the initial 20 regressors). In the case NTB-core is chosen as dependent variable, we obtain that the estimated value for a falls in the range $136 - 181$ (always significant at least at the 1% level), whereas the corresponding value for α_L is between 0.5860 and 0.3391 (significant at least at the 5% level). In the case the overall protection measure is adopted in place of t_i, our econometric output suggests that a is within the interval $105 - 135$ (always significant at least at the 5% level) and α_L within the interval $0.4566 - 0.2226$ (always significant at least at the 5% level but in three cases out of 20). Sign and significance for the other regressors are unchanged with respect to those listed in tables 2 and 3. We have also repeated this exercise by dropping two or more variables obtaining very similar results. Conclusions remain unchanged even estimating regression (3) as a single equation treating z_i and I_i as exogenous variables. Yet the choice of the specification used for our core results (including all the 20 regressors) is motivated by the Akaike and Schwarz Information Criteria that suggest it to be preferred with respect to (or substantially equivalent to) the alternative specifications.

Second we have re-estimated model (3)-(5) using t_i (either standing by NTB-core or overall trade restrictiveness measure) as dependent variable in place of $t_i/(1+t_i)$. The estimated parameters are quite consistent with our previous conclusions being a equal to, respectively for NTB-core and overall protection, 117 ($s.e.$ = 15.6480) and 84 ($s.e.$ = 46.4872), and α_L equal to, again respectively for NTB-core and overall protection, 0.5345 ($s.e. = 0.0669$) and 0.3094 ($s.e. = 0.3832$). Notice, however, that in the latter case α_L turns out to be not statistically significant.

Third, we verify if our results are affected by any possible measurement error associated with the trade price elasticities. Following Goldberg and Maggi (1999) we simply replicate the estimation omitting e_i from the specification. Here we are only interested in testing if the sign and the statistical significance of the parameters change. Indeed, we expect the magnitude of the estimated coefficients to be different and, in particular, larger than that obtained including the trade elasticities on the left-hand side. We find that when the dependent variable is NTB-core, a turns out to be 638, while α_L is 0.8968, both statistically significant at any confidence level. When we use the overall protection measure, we obtain 520 for a and 0.9351 for α_L, again significant at any confidence level. This test confirms that the elasticities used in this paper do not present relevant measurement errors able to bias our qualitative conclusions.

Forth, we try a different treatment of the political organization dummy. Accordingly, instead of using I_i as defined in subsection 3.3, we use I_i^* predicted by the discriminant analysis, and treat it as exogenous. Then the model we estimate is only composed of two equation regressions ((3) and (4)). Again our previous results are strongly confirmed (for NTB-core: $a = 118$ with $s.e. = 8.7072$, and $\alpha_L = 0.4691$ with $s.e. = 0.0414$; for overall-protection: $a = 93$ with $s.e. = 11.8823$, and $\alpha_L = 0.3762$ with $s.e. = 0.0821$).

Finally, we estimate the model using production data from the Eurostat - Industry, Trade and Services. As anticipated in subsection 3.2, these data are more aggregated than those used in our previous estimation but the number of observations is remarkably larger (3857 rather than 1597). On the one side, this exercise enables us to verify if any systematic error has been introduced in the reclassification of production data from the PRODCOM system to HS. On the other, the production series from Eurostat - Industry, Trade and Services is almost complete. Therefore, by using this series we check if our previous results are affected by systematic missing values in the PRODCOM dataset. To implement this robustness check, z_i is constructed by dividing imports (exports) for the import-competing (export-competing) sector (6-digit) by the average production in the associated industry (4-

digit level). In this way, however, the aggregate production value associated to trade data for each product line tends to be systematically larger than the disaggregated one. Indeed, the mean for z_i constructed using disaggregated data is 4.38, while the mean for the same variable obtained using aggregated data is 14.82. As a consequence we expect the estimated a to be larger than before by a multiplicative factor of about 10. As usual we replicate this exercise using either NTB-core or overall-protection as dependent variable. We find that in the former case α_L equals 0.5206 ($s.e. = 0.0352$) and a equals 1023 ($s.e. = 62.4706$), in the latter α_L turns out to be 0.4282 ($s.e. = 0.0403$) while a is 761 ($s.e. = 52.8986$). Previous qualitative conclusions are thus again corroborated.

6 Conclusions

The GH (1994) model has received a great deal of attention by the international economics literature, both from theoretical and empirical standpoints. It has been empirically tested by several studies for the US, Turkey and Australia, obtaining support by the data; and has offered the basis to a number of extensions. However, the absence of data on political contributions has so far made it difficult to find for the EU an appropriate discrimination strategy that enables to distinguish between organized and unorganized sectors. On the other hand, the increasing interest showed for the European politics has produced a growing theoretical literature on policy-making in the EU. In this paper we try to fill this gap between theoretical aspirations and lack of reliable empirical evidence.

We suggest an original strategy to discriminate between organized and unorganized sectors in the European Union that was absent in the literature. In doing so we treat the EU as a single entity interpreting the spirit of the common trade policy that characterizes (and will characterize even more in the future) the politics of the various member states.

Our results suggest that the model fits well the data. Consistently with the previous empirical literature, we find that the weight that the government attaches upon general welfare is several times larger than that associated to political contributions. This conclusion seems sensible in our case given that the decision making process in the EU is mainly in the hands of the Commission that is meant to be a technical and supranational body. Hence, we expect such an institution to have a particular regard for aggregate social welfare. Nonetheless, the hypothesis that the political officials are pure welfare maximizers is rejected at a high confidence level, what hints consistency of data with the protection-for-sale framework. On the other side, the percentage of organized groups at the European level turns out smaller than that previously found for single country estimations. Such a result is not surprising, but rather in line with the experience of the European Union where lobby formation at the supranational level represents a recent phenomenon. A smaller percentage of the population organized in lobbies is also coherent with the large value found for the weight attached to social welfare.

References

Anderson, James and Peter Neary, 1996. A New Approach to Evaluating Trade Policy. *Review of Economic Studies* 63, 107-125.

Anderson, James and Peter Neary, 2003. The Mercantilist Index of Trade Policy. *International Economic Review* 44, 627-649.

Balaoing, Annette and Joseph Francois, 2006. The Political Economy of Protection in a Customs Union: What Drives the Tariff Structure of the EU?. *Tinbergen Institute, Erasmus University Rotterdam, Working Paper*

Ball, David S., 1967. United States Effective Tariffs and Labor's Share. *Journal of Political Economy* 75, 183-187.

Belfrage, Carl-Johan, 2004. Special Interest Politics and Trade Policy - An Empirical Challenge. *Department of Economics, Lund University, Working Papers* 2005_31

Belloc Marianna and Paolo Guerrieri, 2007. Special Interest Groups and Trade Policy in the EU. *Open Economies Review*, forthcoming

Broscheid Andreas and David Cohen, 2003. Insider and Outsider Lobbying of the European Commission An Informational Model of Forum Politics. *European Union Politics* 4, 1465-1165

Caves, Richard E., 1976. Economic Models of Political Choice: Canada's Tariff Structure. *Canadian Journal of Economics* 9, 278-300.

Chow, Gregory C., 1973. On the Computation of Full-Information Maximum Likelihood Estimates for Nonlinear Equation Systems. *Review of Economics and Statistics* 55, 104-109.

Cohen, David, 1997. The European Business Lobby", *Business Strategy Review* 8, 17-25.

Constantopoulos, Maria, 1974. Labor Protection in Western Europe. *European Economic Review* 5, 313-318.

Crombez, Christopher, 2002. Information, Lobbying and the Legislative Process in the European Union. *European Union Politics* 3, 7-32.

Dür, Andreas and De Bièvre, Dirk, 2006. *Database on the Sectoral and Issue-specific Composition of DG Trade's Civil Society Dialogue.* Mannheim Centre for European Social Research, University of Mannheim.

Eurostat, 2006a. *Industry, Trade and Services - Horizontal View - Structural Business Statistics* http://epp.eurostat.ec.europa.eu.

Eurostat, 2006b. *External Trade - External Trade Detailed Data,* http://epp.eurostat.ec.europa.eu/.

Everitt, Brian S., Landau, Sabine, and Leese Morven, 2001. *Cluster Analysis*, 4nd edition. Arnold Publication, London.

Facchini, Giovanni, Biesebroeck Johannes Van, and Gerald Willmann, 2006. Protection for Sale with Imperfect Rent Capturing. *Canadian Journal of Economics* 39, 845-873.

Fieleke, Norman S., 1976. The Tariff Structure for Manufacturing Industries in the United States: A Test of Some Traditional Explanations. *Columbia Journal of World Business* 11, 98-104.

Gawande, Kishore, 1998. Comparing Theories of Endogenous Protection: Bayesian Comparison of Tobit Models using Gibbs Sampling Output. *Review of Economics and Statistics* 80, 128-140.

Gawande, Kishore, and Usree Bandyopadhyay, 2000. Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection. *Review of Economics and Statistics* 82, 139-152.

Gawande, Kishore, and Pravin Krishna, 2003. The Political Economy of Trade Policy: Empirical Approaches, in E. Kwan Choi and James Harrigan (ed.s), *Handbook of International Trade*, Vol. I. Basil Blackwell, Oxford, pp. 213-250.

Gawande, Kishore, and Pravin Krishna, 2005. Lobbying Competition over US Trade Policy. *NBER Working Paper* 11371.

Gawande, Kishore, Pravin Krishna, and Michael Robbins, 2006. Foreign Lobbies and US Trade Policy. *Review of Economics and Statistics* 88, 563-571.

Goldberg, Penelopi, and Giovanni Maggi, 1999. Protection for Sale: An Empirical Investigation. *American Economic Review* 89, 1135-1155.

Grossman Gene M. and Elhanan Helpman, 1994. Protection for Sale. *American Economic Review* 84, 833-50.

Helpman, Elhanan, 1995. Politics and Trade Policy. *NBER Working Paper* 5309.

Huberty, Carl J., 1994. *Discriminant Analysis. Quantitative Applications in the Social Science Series* 19. Sage Publications, Thousand Oaks, CA.

Imai, Susumu, Katayama, Hajime, and Kala Krishna, 2006. Protection for Sale or Surge Protection?. *NBER Working Paper* 12258.

Leamer, Edward, 1990. Latin America as a Target of Trade Barriers Erected by the Major Developed Countries in 1983. *Journal of Development Economics* 32, 337-68.

Limão, Nuno, and Baybars Karacaovali, 2005. The Clash of Liberalizations: Preferential vs. Multilateral Trade Liberalization in the European Union, *CEPR Discussion Paper* 4793

Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga, 2004. Estimating Import Demand and Export Supply Elasticities. *World Bank Policy Research Working Paper* No. 3452.

Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga, 2006. Estimating Trade Restrictiveness Indices. *World Bank Policy Research Working Paper* 3840.

Kelejian, Harry H., 1971. Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables. *Journal of the American Statistical Association* 66, 373-74.

Kelejian, Harry H., 1982. An Extension of a Standard Test for Heteroskedasticity to a Systems Framework. *Journal of Econometrics* 20, 325-333.

Mazey, Sonia, and Jeremy Richardson, 1993. Transference of Power, Decision Rules, and the Rules of the Game, in Sonia Mazey and Jeremy Richardson (ed.s) *Lobbying in the European Community.* Oxford University Press, Oxford, pp. 3-26.

McCalman, Phillip, 2004. Protection for Sale and Trade Liberalization: an Empirical Investigation. *Review of International Economics* 12, 81-94.

Mitra, Devashish, Thomakos, Dimitrios D., and Mehmet A. Ulubaşoğlu, 2002. Protection for Sale in a Developing Country: Democracy vs. Dictatorship. *Review of Economics and Statistics* 84, 497-508.

Mitra, Devashish, Thomakos, Dimitrios D., and Mehmet A. Ulubaşoğlu, 2006. Can We Obtain Realistic Parameter Estimates for the 'Protection for Sale' Model?. *Canadian Journal of Economics* 39, 187-210.

Olson, Mancur, 1965. *The Logic of Collective Action.* Harvard University Press, Cambridge, MA.

A Detailed data appendix

NTB-core: Ad-valorem equivalent of non-tariff barriers. Source: Kee et al. (2006).

Overall-protection: Composite index that combines the ad-valorem equivalent of core NTBs with ad-valorem equivalent of the domestic support and tariff barriers. Source: Kee et al. (2006).

y: Value of production. Millions of euros. Source: Eurostat (2006b).

Imp (**exp**): External imports (exports) that is trading of goods between a Member State and a non-member country. Millions of euros. Source: Eurostat (2006b).

z: Inverse of the trade penetration ratio (ratio of production value (y) over imports (imp) if imports $\geq$ exports, and over exports (exp) otherwise).

e: Import-demand (export-supply) price elasticity. Source: Kee et al. (2004).

Scale: Index of industry scale given by production per enterprise. Authors' calculations. Source: Eurostat (2006a).

Turnover: Ratio of turnover over production value. Authors' calculations. Source: Eurostat (2006a).

Long-term rentals: Ratio of gross investment in long-term rentals, operational and financial leasing of goods over total purchases of goods and services. Authors' calculations. Source: Eurostat (2006a).

Tangible goods: Ratio of gross investment in tangible goods over total purchases of goods and services. Authors' calculations. Source: Eurostat (2006a).

Land: Ratio of gross investment in land over total purchases of goods and services. Authors' calculations. Source: Eurostat (2006a).

Existing buildings: Ratio of gross investment in existing buildings over total purchases of goods and services. Authors' calculations. Source: Eurostat (2006a).

Constructions: Ratio of gross investment in constructions over total purchases of goods and services. Authors' calculations. Source: Eurostat (2006a).

Labor share: Share of wages and salaries in value added. Authors' calculations. Source: Eurostat (2006a).

Agency workers: Percentage of agency workers in total of persons employed. Authors' calculations. Source: Eurostat (2006a).

Apprentices: Percentage of apprentices in total of persons employed. Authors' calculations. Source: Eurostat (2006a).

Unpaid persons: Percentage of unpaid persons in total of persons employed. Authors' calculations. Source: Eurostat (2006a).

R&D personnel: Percentage of R&D personnel in total of persons employed. Authors' calculations. Source: Eurostat (2006a).

Employees: Percentage of employees in total of persons employed. Authors' calculations. Source: Eurostat (2006a).

Social security: Ratio of social security costs over total of wages and salaries. Authors' calculations. Source: Eurostat (2006a).

Apparent productivity: Ratio of value added over number of persons employed. Authors' calculations. Source: Eurostat (2006a).

Personnel costs: Ratio of personnel costs over total of wages and salaries. Authors' calculations. Source: Eurostat (2006a).

Wage: Average wage per person. Millions of euros. Authors' calculations. Source: Eurostat (2006a).

Share prod value: Ratio of production value of a given sector over production value in total manufacturing. Authors' calculations. Source: Eurostat (2006a).

Share persons employed: Ratio of number of persons employed in a given sector over number of persons employed in total manufacturing. Authors' calculations. Source: Eurostat (2006a).

Share enterprises: Ratio of number of enterprises in a given sector over number of enterprises in total manufacturing. Authors' calculations. Source: Eurostat (2006a).

Pubblicato in proprio
Dipartimento di Economia Pubblica
Facoltà di Economia
Università degli Studi di Roma “La Sapienza”
Via del Castro Laurenziano 9 – 00161 Roma

Working Paper del Dipartimento di Economia Pubblica
Università degli studi di Roma "La Sapienza"
Via del Castro Laurenziano 9 – 00161 Roma

I Working Paper vengono pubblicati per favorire la tempestiva divulgazione, in forma provvisoria o definitiva, dei risultati delle ricerche sulla teoria e la politica economica.
La pubblicazione dei lavori è soggetta all'approvazione del Comitato Scientifico, sentito il parere di un referee.

I Working Paper del Dipartimento di Economia Pubblica ottemperano agli obblighi previsti dall'art. 1 del D.L.: 31.8.45 n. 660.