Firm heterogeneity and lobby participation

Matilde Bombardini*
University of British Columbia

June 2005

Abstract

The structure of protection across sectors has been interpreted as the result of competition among lobbies to influence politicians, but lobbies have been treated as unitary decision makers and little attention has been devoted to the importance of individual firms in this process. This paper builds a model where individual firms rationally decide whether to enter the lobby and determine the amount of resources to allocate to political contributions. Firms of different sizes are shown to have different incentives to participate in the lobby. Therefore the size distribution of firms plays a crucial role in determining the equilibrium level of protection in a sector. The model is tested employing data on protection measures, political contributions, and characteristics of the size distribution of firms. The empirical evidence shows that, accounting for individual firm behavior, the model explains a larger fraction of the variation of protection across sectors. In particular, the model rationalizes a pattern that I uncover in the data according to which industries characterized by higher firm size dispersion obtain a higher level of protection.

JEL Classification codes: D7, H42, F13, L7

I am indebted to Pol Antràs and Andrés Rodríguez-Clare for invaluable guidance and support. The data were kindly provided by Kishore Gawande and Jim Snyder. I would like to thank Daron Acemoglu, David Autor, Haluk Ergin, Ivan Fernandez-Val, Kishore Gawande, Rodney Ludema, Marc Melitz, Veronica Rappoport, Francesco Trebbi, Ivan Werning, and seminar participants at Boston College, Boston University, Chicago GSB, CU Boulder, EIIT conference, Federal Reserve Board, IIES Stockholm, Kellogg, LSE, MIT, SAIS Johns Hopkins, Stanford GSB, University of British Columbia, UT Austin, for helpful comments. I am specially thankful to Elhanan Helpman for encouraging me to pursue this project.

*Department of Economics, matildeb@mit.edu

1 Introduction

Why do some industries receive more protection than others? This question has been the subject of a large body of theoretical and empirical literature. The idea that the structure of trade policy is mainly the result of interest groups lobbying the government to be shielded from foreign competition has gathered large consensus among trade economists, but little attention has been devoted to the role played by individual firms in shaping the structure of protection across sectors. In this paper I uncover a strong empirical link between the level of protection of a given sector and the characteristics of the size distribution of individual firms in the industry. While traditional models of endogenous trade policy cannot account for this pattern, this paper shows that from a theoretical point of view this evidence can be reconciled with the "Protection for Sale" paradigm, first introduced by Grossman and Helpman (1994), if we shift the focus to the behavior of individual firms.

The literature on the political economy of trade policy has illustrated the importance of the interaction between governments and interest groups in the determination of the structure of trade policy in a variety of settings. The common element to these studies is the description of a specific channel through which interest groups influence the policy maker in the choice of trade policy. Most contributions can be summarized by a common scheme: pressure groups attempt to influence the government's choice of trade policy through the promise of votes, monetary donations, and general campaign support; the government grants protection from foreign competition to a sector by comparing the benefits that it receives from the industry's lobby and the social welfare loss brought about by protection measures.

Nevertheless, these studies have failed to investigate the behavior of the individual members (firms) that form interest groups, thus potentially disregarding important aspects of interest groups' aggregate behavior. In particular, none of these studies can account for an important empirical feature that I uncover in this paper. Controlling for all the variables employed in previous empirical studies of the "Protection for Sale" framework, the dispersion of the size distribution of firms within a sector is positively correlated with the level of protection granted to an industry. Why are sector with higher firm size dispersion more protected? According to traditional models of the political economy of trade policy the distribution of firms within a sector should not matter for the determination of trade policy, so the evidence I bring forth calls for a model that incorporates the behavior of individual firms and for further empirical investigation of this new framework.

In this model I embed firm heterogeneity in a menu auction set-up à la Grossman and Helpman (1994). Each firm that enters a lobby presents the government with a contribution schedule that associates a monetary contribution to each potential degree of protection. Therefore in this set-up each firm takes individually rational decisions by maximizing its own private benefit. The government chooses the level of protection by trading off contributions and loss of aggregate welfare. The framework is further enriched by allowing the presence of set-up costs associated with political contributions. Firms therefore decide whether to contribute in the political game and incur the initial expenses necessary to play an active role in the sector lobby. The lump-sum nature of these set-up costs is such that they do not depend on the size of the firm and they determine which firms are active in the lobby in equilibrium. The model predicts that what matters for the strength of the lobby (and therefore for the equilibrium level of protection) is not the size of the sector per se, but the share of the total industry output produced by firms that make positive contributions in the sector's lobby.

The share of industry output produced by firms participating in the lobby, in turn, is determined in equilibrium as a result of the coordination of individual firms and depends crucially on the size distribution of firms within the sector. I show that the efficient equilibrium for the lobby is one in which individual firm's participation depends on the firm's size: larger firms gain more from protection and are able to make larger contributions, therefore in the presence of an initial fixed cost of entering the lobby, it is efficient for larger firms to participate in political activity. This logic is suggested (but not empirically verified) in the work of Masters and Keim (1985) on the motivation behind a corporation's choice to set up Political Action Committees[1]: controlling for other determinants of political participation "the economic size of the firm should also be positively related to the probability of having a PAC [i.e. entering the political game]. This is because the initial fixed costs of organizing for political activity may be spread over a larger asset base".

As a result, the model predicts that under certain conditions (that I find verified in the data), industrial sectors where the distribution of firm size is more dispersed are more likely to have a larger fraction of the sector output produced by firms large enough to incur the fixed cost of contributing and participate in the lobby. Therefore a larger firm dispersion will result in a larger participation share and in a higher level of protection.

The empirical section of this paper tests the predictions of the model employing data on protection measures, the size of sectors, the size distribution of firms within sectors and political contributions by individual firms. It is worth emphasizing that, differently from this paper, previous empirical studies of the "Protection for Sale" model, like Goldberg and Maggi (1999) and

[1] Commonly referred to as PAC's. I will describe in the empirical section what PAC's are and how they work. For now I simply take the choice of setting up a PAC as the decision of entering into the political game.

Gawande and Bandyopadhyay (2000), have made use of sector-level aggregate political contributions data. By matching firm-level contributions data obtained from records of the Federal Election Commission to individual company information available on COMPUSTAT, I am able to test a number of predictions about individual firms' lobbying behavior.

The empirical section starts by providing reduced form evidence that characteristics of the size distribution of firms are important in explaining the pattern of protection across industries. More importantly, in this section I provide evidence that a number of predictions of the model are consistent with the data. First, I verify that, both at the industry level and over all sectors, larger firms are more likely to participate in the political game and make larger contributions. Second, using firm-level data on output and political contributions, I measure the share of total output in a sector produced by firms that lobby and I show that this share is an increasing function of the average firm size and the firm size dispersion within the sector, as predicted by the model. Third, I show that accounting for differences in participation shares across sectors in the way predicted by the model gives sensible parameter estimates. I find strong support that the correct specification for the equation that determines the level of protection should account for participation shares as well as the total size of the sector. Finally I test the model presented in this paper against the "Protection for Sale" benchmark and I show that the Heterogeneity model helps explain a larger fraction of the variation of protection levels across sectors.

This paper builds on the important strand of literature that has explored and established a series of models formalizing the interaction between the government and interest groups. The literature is so vast that I do not attempt at being exhaustive and simply refer to the survey by Rodrik (1995), where the various approaches are analyzed and linked to one another. Rodrik offers a clear perspective on the work in this area going from the Political Support Function introduced by Hillman (1989) to the Tariff-Formation Function approach proposed by Findlay and Wellisz (1982) to the Campaign Contribution approach explored by Magee et al. (1989) and more recently by Grossman and Helpman (1994). While previous approaches had provided a reduced form link between the characteristics of a sector and the benefit to the government of granting protection, the Grossman-Helpman model (henceforth GH) describes a specific channel through which interest groups affect government decisions. In GH lobbies enter a game with the government and bid for protection through campaign contribution offers which the policy maker takes into account when maximizing its own utility (which is a function of aggregate welfare and total campaign funds). This last contribution by Grossman and Helpman is the most carefully micro-founded model up to date and I therefore build on this work to introduce firm heterogeneity and individual firm decisions.

This paper is also related to a more recent but fast-growing area of international trade concerned with the importance of relaxing the assumption of identical firms within sectors. This literature has emphasized, from both a theoretical and an empirical point of view, that allowing for differences in firm productivity and size within a sector helps explain a number of facts that the representative firm approach cannot account for.

One of the first and most influential papers in this literature is Melitz (2003), where a new theoretical framework is introduced and firm heterogeneity plays an important role in the amount of factors reallocation following trade liberalization. As factors of production are reallocated to larger and more productive firms that self-select into the export market, the economy experiences an increase in productivity. The evidence on this self-selection of larger firms into the export market has been widely documented by Bernard and Jensen (1999) and Aw et al. (2000). Self-selection of larger firms is also a feature of the model presented in this paper. By modelling firm heterogeneity Bernard et al. (2003) are able to explain a number of facts about the link between productivity and the exporting status of firms. In Helpman et al. (2004) firm heterogeneity plays a role in determining the prevalence of exports versus foreign direct investment as the channel for domestic firms to access foreign markets. In their model the degree of firm productivity dispersion strongly affects the choice of foreign market access. Antràs and Helpman (2004) introduce firm heterogeneity as a determinant of the choice of integration versus outsourcing by multinational firms. Trade policy is an area where the firm dimension has not been carefully considered and where interest groups have generally been treated as unitary decision makers, characterized by some aggregate dimensions like total sector output or the total number of firms.

An account of where this contribution stands in the literature would be incomplete without mentioning previous work on the provision of public goods and lobby formation. In his seminal contribution Mancur Olson (1965) informally advanced the idea that "in groups of members of unequal 'size'...there is the greatest likelihood that a collective good will be provided". The motivation for this statement relies on the presumption that larger members will find it economically viable to participate in lobbying activities and that groups with a few large members will be more effective than groups with a large number of small members. The notion that large members will ensure a sizeable provision of a collective good is common to the literature on public goods and free-riding. In this paper I intend to revisit Olson's insightful contribution and provide a more rigorous micro-foundation of firms behavior.

Two papers, one by Pecorino (1998) and one by Magee (2002), tackle the issue of free-riding in the interaction between firms in a lobby. While the two papers offer an interesting insight into the issue of how the number of identical firms in a sector affect the likelihood of free-riding, these models do not analyze the decision of the firm to enter the lobby and do not have an immediate application to the case where firms are heterogeneous. Moreover, they adopt a repeated game framework where the participation decision of the firm in the lobby is not explicitly modeled. Gawande (1997) adopts the reduced form model of private provision of public goods first introduced by Bergstrom et al. (1986) and presents empirical evidence that the concentration of firms in a sector increases the level of protection. Their papers share a common failure to carefully micro-found the decision of firm participation into the lobby and still adopt the tariff function approach without having an explicit mechanism of interaction between the government and individual firms. The model I present substantially improves on this reduced form approach by modelling explicitly the firms' decisions about participation and the level of contributions.

Finally, the most thorough analysis of lobby formation in the framework of the Grossman and Helpman model is due to Mitra (1999). In his paper lobby formation is a discrete process: either a sector organizes into a lobby or it is unorganized. In this sense sectors are again treated as black boxes where firms do not play any role: lobby formation realizes on the sole condition that total surplus is greater than the set up cost. This seems a reasonable assumption for lobby formation if we consider, as Mitra does, sectors where firms are all identical and symmetry arguments can justify a coordination outcome. This characterization seems less innocuous if there are large differences among firms within a sector, which is what we observe in the data. Moreover, while Mitra's paper helps explain the presence or the absence of the lobby, the model in this paper describes the "intensity" of lobbying, that is the share of total industry resources that are directed to the political activity of a sector.

The remainder of the paper is divided into two main sections: a theoretical model and empirical methods and results. Section 2 presents the structure of the economy, Section 3 describes the political game and Section 4 presents the structure of the lobby participation decision. Section 5.1 describes the data used in this study and presents preliminary regressions that emphasize the distribution of firms within a sector as a statistically and economically significant determinant of protection levels. Section 5.2 tests several direct predictions of the model and compares the performance of the model presented in this paper with the performance of the "Protection for Sale" model. Section 6 concludes.

2 Structure of the economy

Consider an economy that trades with the rest of the world and faces fixed world prices. There are several goods traded. The numeraire good, x_0, is not taxable, but all the other m goods can potentially bear an import or export tax. I will focus on import tariffs and therefore on the import-competing part of the economy, but the expressions obtained describe export subsidies as well. Denote ad valorem tariff on good x_i by τ_i such that the domestic price for good i is:

$$p_i = (1 + \tau_i)\, p_i^* \tag{1}$$

where p_i^* is the international price of good i. Normalize all international prices to one so that the expression (1) simplifies to $p_i = 1 + \tau_i$.

The population in this economy is of size one and its preferences are represented by the following quasi-linear utility function:

$$U(c_0, c_i) = c_0 + \sum_{i=1}^{m} u_i(c_i)$$

where c_0 is consumption of good x_0 and c_i is consumption of good x_i. The function $u_i(\cdot)$ is differentiable, increasing, and strictly concave.

Quasi-linear preferences allow the demand for each good x_i to depend only on its domestic price p_i (relative to the price of the numeraire good x_0), under the condition that income is high enough to guarantee a positive consumption of good x_0:

$$c_i = d_i(p_i)$$

Each consumer spends the amount $c_i(p_i)$ on good x_i and devotes the rest of his income I to the numeraire good. I assume that consumption of the numeraire good x_0 is always positive.

Under these preferences, the indirect utility function $V(I, \mathbf{p})$ is:

$$\begin{aligned} V(I, \mathbf{p}) &= I + S(\mathbf{p}) = \\ &= I + \sum_{i=1}^{m} u_i(d_i(p_i)) - p_i d(p_i) \end{aligned}$$

where I is income and $S(\mathbf{p})$ is consumer surplus.

The numeraire good is produced one-to-one with labor and does not require any other input. Each of the other goods x_i is produced using labor and a factor specific to the sector and to the firm. Free trade in the numeraire good and the production technology for x_0 assures that the wage is equal to 1, assuming that the production of the numeraire good is always positive. Each sector i

is populated by a set of firms that are endowed with different amounts of the specific factor. Firm j in sector i produces according to the following production function:

$$x_{ij} = f\left(K_{ij}, L_{ij}\right)$$

where $f(\cdot)$ is increasing and concave in both arguments, K_{ij} is the specific factor endowment[2] of firm j in sector i and L_{ij} is the labor employed by this firm.

The return to the firm's specific factor, Π_{ij}, depends on the domestic price for the good produced and the amount of specific factor owned. By Hotelling lemma, as the domestic price p_i increases the rent increases by the amount of output produced:

$$\frac{\partial \Pi_{ij}}{\partial p_i} = x_{ij}$$

The government is not a pure welfare maximizer. In particular, the incumbent government cares about aggregate welfare as well as monetary contributions that can be used for re-election or for other purposes. The government's objective function G depends on aggregate welfare gross of contributions, W , and on the level of contributions that it receives from interest groups, C

$$G = C + aW$$

The composition of C is discussed later in the model. I restrict the set of policy tools available to the government to trade taxes and subsidies and as indicated above I allow these taxes and subsidies to apply to non-numeraire goods only.

Aggregate welfare is the sum of labor income $l = 1$, consumer surplus $S(\mathbf{p})$, tariff revenues $r(\mathbf{p})$ (that are redistributed back to consumers) and rents, Π_i,[3] that accrue to the owners of the specific factors used in the production of non-numeraire goods, as described below.

$$W = 1 + r(\mathbf{p}) + S(\mathbf{p}) + \sum_{i=1}^{m} \Pi_i$$

Tariff revenues from each sector i will be the product of the ad valorem tariff and imports, $m_i(p_i) = d_i(p_i) - X_i$:

$$r(\mathbf{p}) = \sum_{i=1}^{m} \tau_i \left(d_i(p_i) - X_i\right)$$

where $d_i(p_i)$ is demand for good i and X_i is total output in sector i.

[2] An alternative interpretation is that firms are endowed with the same amount of the specific factor, but each firm' specific factor exhibit a different productivity level. In this paper the underlying cause of firm heterogeneity is taken as given. This is a standard assumption, common to the literature on firm heterogeneity (Melitz (2003) and Bernard et al. (2003)).

[3] Sector-level aggregate rents are $\Pi_i = \sum_{j \in S_i} \Pi_{ij}$ where S_i is the set of firms in sector i.

3 Tariff setting game

The structure of the game is similar to the one described by Bernheim and Whinston (1986) and adopted by Grossman and Helpman (1994). Here firms act as a set of principals trying to induce the agent, the government, to implement a policy that might be costly for the government itself, but would benefit the firm in terms of increased specific factor rent. In this framework each firm is an individual player and therefore a lobby will result from the aggregation of contribution offers that are decided on an individual level by all firms in a given sector i. This paper follows Grossman and Helpman (1994) in adapting menu auctions to the tariff setting game, with the fundamental difference that each individual firm is considered as a different principal, which decides independently on its contribution schedule. More precisely the owner of the firm's stock of specific factor is a principal that decides whether to participate in lobbying efforts and how much to contribute to the government. The concepts of firm and owner of the firm's specific factor are used interchangeably.

To this framework the model adds the presence of a fixed cost F, independent of firm size, to participate in lobbying activities. When firms decide to make political contributions the first F dollars do not reach the hands of politicians and are spent to channel resources. This cost might represent the apparatus of legal expenses and resources necessary to set-up a section of the firm devoted to contacting policy-makers or more in general it might represent the amount of money that the firm perceives as a "minimum" to be able to play some role in the policy decision. The presence of this friction substantially modifies the game initially introduced by Bernheim and Whinston (1986).

The players of this game are the individual firms in each sector and the government. The government chooses a vector of prices $\mathbf{p} \in P$ while the strategy space for firm j in sector i consists of a contribution schedule $C_{ij}(\mathbf{p})$ that associates a level of monetary contribution to each price vector. The presence of a fixed cost requires to distinguish between a gross contributions $\widetilde{C}_{ij}(\mathbf{p})$ (gross of the fixed cost), which is the amount of money the firm disburses, and net contributions $C_{ij}(\mathbf{p})$ which is the amount of money that the government receives.

This firm's gross payoff is $W_{ij}(\mathbf{p}) = l_{ij} + \Pi_{ij}(\mathbf{p}) + \alpha_{ij}(r(\mathbf{p}) + S(\mathbf{p}))$ where α_{ij} is the share of population represented by the owner of specific factors in firm ij, l_{ij} is the labor income of the owner of firm ij specific factor. The firm's net payoff is $V_{ij} = W_{ij}(\mathbf{p}) - \widetilde{C}_{ij}(\mathbf{p})$ while the government's payoff is $G = C + aW(\mathbf{p})$ as specified above.

The extensive form of the game is the following:

(i) in the first stage, firms present the government with a contribution schedule $C_{ij}(\mathbf{p}) = \max\left\{0, \widetilde{C}_{ij}(\mathbf{p}) - F\right\}$

(ii) in the second stage, the government chooses a price vector $\mathbf{p}$ and collects $C_{ij}(\mathbf{p})$ from each firm ij

The timing of the game is reported in $Figure$ 1.

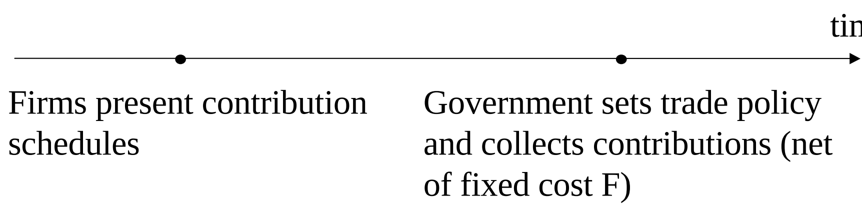


Figure 1

What follows is a set of conditions for an equilibrium in the tariff setting game between the set of firms that make positive contributions and the policy maker. The interpretation of these conditions is given below. Denote by L_i the set of firms in sector i that make positive contributions, which is now taken as given. Denote by S_i the set of all firms in sector i.

A configuration $\left(\left\{C_{ij}^o\right\},\ \mathbf{p}^o\right)$ is a subgame-perfect equilibrium of the game if and only if:

1. C_{ij}^o is feasible

2. $\mathbf{p}^o \in \arg\max \sum\limits_i \sum\limits_{j \in S_i} C_{ij}^o(p) + aW(p)$

3. Given $\widetilde{C}_h^o(\mathbf{p}^o)$, $h \neq j$, $\widetilde{C}_j^o(\mathbf{p}^o) \in \arg\max\ W_j(\mathbf{p}^o) - \widetilde{C}_j(\mathbf{p}^o)$ where $\mathbf{p}^o$ satisfies condition 2.

Condition 1 states that contributions cannot be larger than total income of firm ij and cannot be negative. Condition 2 states that the government chooses $\mathbf{p}$ to maximize its welfare, given the equilibrium contribution schedules presented by each firm. Condition 3 states that each firm chooses its contribution schedule to maximize its net welfare.

Condition 3 can be decomposed into sub-conditions that help to illustrate the characteristics of the equilibrium. :

(a) The joint surplus of the government and firm ij is maximized at $\mathbf{p}^o$ (otherwise the firm could modify its contribution schedule to increase the joint surplus and would retain a fraction of this increased surplus):

$$W_{ij}(\mathbf{p}^o) - \widetilde{C}_{ij}^o(\mathbf{p}^o) + \sum_i \sum_{j \in S_i} C_{ij}^o(\mathbf{p}^o) + aW(\mathbf{p}^o) \geq W_{ij}(\mathbf{p}) - \widetilde{C}_{ij}^o(\mathbf{p}) + \sum_i \sum_{j \in S_i} C_{ij}^o(\mathbf{p}) + aW(\mathbf{p}) \quad \forall \mathbf{p}$$

(b) Due to the timing of the game, firm ij manages to extract all the available surplus from the government (it contributes just enough to maintain the government at the same level of welfare that it would achieve if firm ij were not participating in the political game):

$\exists\ \mathbf{p}^{-ij}\ \forall ij \in L$ such that $\mathbf{p}^{-ij} \in \arg\max \sum_i \sum_{j \in S_i} C_{ij}^o(\mathbf{p}) + aW(\mathbf{p})$ and $C_{ij}^o\left(\mathbf{p}^{-ij}\right) = 0$

This condition will be used to calculate contributions in equilibrium, so details are provided later in the section.

I assume that contribution schedules are differentiable at the around the equilibrium price vector $\mathbf{p}^o$ which, also according to Grossman and Helpman (1994), is reasonable if we want to prevent mistakes in the calculations of the individual firm from resulting in large swings in the contributions offered. For the subset of firms that make positive contributions in equilibrium, combining condition 2 and sub-condition (a) one obtains the following condition:

$$\nabla W_{ij}(\mathbf{p}^o) = \nabla C_{ij}^o(\mathbf{p}^o)$$

I indicate the subset of firms that make positive contributions in equilibrium in sector i with L_i. This condition implies that contribution schedules are locally truthful for the subsets L_i's of firms, that is, around the equilibrium price vector $\mathbf{p}^o$ they reflect the willingness to pay of the firm for an increase in the domestic price. In the presence of sunk costs of contributing there are firms that in equilibrium do no contribute any amount and do not pay the initial fixed cost F. For these firms it will not be the case that the contribution schedule is not locally truthful as at $\mathbf{p}^o$ the gross payoff might have a positive slope, but the contribution schedule is flat at zero at that price vector. A contribution schedule that is flat at zero around $\mathbf{p}^o$ is optimal for some firms because any price vector that could be induced by a positive contribution would not compensate for the initial fixed cost of contributing.

Although only differentiability is needed to obtain local truthfulness of the contribution schedule and ultimately the equilibrium price, more restrictive assumptions are needed to obtain the level of contributions of each firm. I therefore restrict the attention to contribution schedules that are truthful over the range of price vectors that entail positive contributions:

$$C_{ij}(\mathbf{p}) = \max\left[0, W_{ij}(\mathbf{p}) - B_{ij} - F\right] \tag{2}$$

where B_{ij} indicates a level of welfare to be determined in equilibrium.The presence of a fixed cost does not allow to extend to this framework all the results derived by Bernheim and Whinston (1986). In the absence of fixed costs Bernheim and Whinston (1986) have shown that truthful contribution schedules and truthful Nash equilibria may be focal. First, firms' best-response sets always include a truthful strategy so firms cannot lose from choosing a truthful contribution schedule. Second, TNE are Pareto optimal, robust to communication among players and are therefore coalition-proof. In this framework, Pareto optimality for the entire sector might not be the most realistic property because it takes into account the welfare of firms that do non make any contributions in equilibrium and therefore "free-ride" on other firms contributions. Therefore I will argue in the following section that we should look at Pareto optimality for the subset of firms L_i for each sector.

Given the simple expression in (2) for firms' contributions schedules the government will choose $\mathbf{p}^o$ solving this program:

$$\mathbf{p}^o = \arg\max \left[\sum_i \sum_{j \in L_i} W_{ij}(\mathbf{p}) + aW(\mathbf{p}) \right]$$

According to this program the objective function of the government is represented by a weighted sum of consumers' and producers' surplus, where the weight on the welfare of firms that make positive contributions in equilibrium is larger than the weight on the welfare of firms that do not make contributions and of the rest of the population. The first order condition for this multivariate maximization problem is the following:

$$\sum_i \sum_{j \in L_i} \nabla W_{ij}(\mathbf{p}^o) + a \nabla W(\mathbf{p}^o) = 0 \tag{3}$$

Consider the impact of the increase in price p_k of good x_k on the welfare of firm ij owner:

$$\frac{\partial W_{ij}}{\partial p_k} = (\delta_{ik}\theta_{ij} - \alpha_{ij}) X_k + \alpha_{ij}(p_k - 1) m'_k$$

where $\delta_{ik} = 1$ if $i = k$ and $\delta_{ik} = 0$ otherwise, $x_{ij} = \theta_{ij} X_i$ and θ_{ij} is the share of total output in sector i produced by firm j:

$$\theta_{ij} = \frac{x_{ij}}{\sum\limits_{j \in S_i} x_{ij}}$$

Therefore aggregating over all firms in sector i that make positive contributions, I obtain the impact of an increase of price p_k on the welfare of the set of firms L_i (the firms that lobby in sector i), call it W_{Li}:

$$\sum_{j \in L_i} \frac{\partial W_{ij}}{\partial p_k} = \frac{\partial W_{Li}}{\partial p_k} = (\delta_{ik}\theta_i - \alpha_{Li}) X_i + \alpha_i (p_k - 1) m'_i \tag{4}$$

where $\alpha_{Li} = \sum\limits_{j \in L_i} \alpha_{ij}$ and $\theta_i = \sum\limits_{j \in L_i} \theta_{ij}$. θ_i is the share of total output in sector i produced by firms that make positive contributions and it might be zero if in sector i no firm makes political contributions (the sector is therefore not politically organized). Now aggregating (4) over all sectors:

$$\sum_i \frac{\partial W_{Li}}{\partial p_k} = (\theta_k - \alpha_L) X_i + \alpha_L (p_k - 1) m'_k \tag{5}$$

where $\alpha_L = \sum_i \alpha_{Li}$ is the share of the population in the economy that owns some specific factor and participates in the political game and $\theta_k = \sum_i \delta_{ik}\theta_i$.

Aggregate welfare is affected by an increase of the price of good x_k according to the following expression:

$$\frac{\partial W}{\partial p_k} = (p_k - 1) m'_k + m_k - d(p_k) + X_k = (p_k - 1) m'_k \tag{6}$$

Notice that in the absence of lobbying the welfare maximizing domestic price is the international price $p_k = 1$. Now substitute expressions (6) and (5) into the first-order condition (3) and rearrange to obtain the following expression for the domestic price of good x_k:

$$p_k^o - 1 = -\frac{\theta_k^o - \alpha_L}{a + \alpha_L} \frac{X_k^o}{m_k'^o}$$

The first-order condition can be rewritten in a fashion similar to the "Protection for sale" equation:

Proposition 1 *If firms' contribution schedules are truthful, for a given set of firms participating in the political game, the equilibrium domestic price of good x_i is given by the following expression:*

$$\frac{\tau_i^o}{1 + \tau_i^o} = \frac{\theta_i^o - \alpha_L}{a + \alpha_L} \left(\frac{z_i^o}{e_i^o} \right) \tag{7}$$

where $z_i^o = \frac{X_i^o}{m_i^o}$ is the inverse import penetration ratio, $e_i^o = -\ m_i'^o p_i^o / m_i^o$ is the price elasticity of imports and θ_i^o is the equilibrium share of total output of sector i represented by firms that make positive contributions:

$$\theta_i^o = \frac{\sum_{j \in L_i} x_{ij}}{\sum_{j \in S_i} x_{ij}}$$

The level of protection τ_i depends on several factors (apart from a which is constant across sectors). First, the larger the level of output of a sector relative to imports the larger the deviation from free trade. This is a result of the relatively smaller distortion imposed on sectors that have low levels of imports. The size of output affects the level of protection because a larger industry will, ceteris paribus, receive a larger benefit from the increase in price p_i and the government can expect to receive larger contributions therefore protection granted will be higher. Second, sectors characterized by lower price elasticity of imports receive larger protection as the distortion created by protection is lower. The third factor is going to be the focus of the remaining part of the theoretical section and of much of the empirical section: the equilibrium share of total output produced by firms lobbying, θ_i^o. For a given level of output the larger this share the larger the marginal contributions the government can expect, the higher the level of protection. The share θ_i^o can be seen as the "intensity" of lobbying and allows sectors to be characterized by different degrees of lobby participation shares. This is conceptually different from GH insofar as it allows the choice of lobbying in a sector to be a smooth function. Remember that in GH the equilibrium tariff has the following comparable expression:

$$\frac{\tau_i^o}{1+\tau_i^o} = \frac{I_i - \alpha_L}{a + \alpha_L}\left(\frac{z_i^o}{e_i^o}\right)$$

where I_i is one if sector i is politically organized and zero otherwise. Therefore in GH and in Mitra (1999) lobbying is a binary choice: either the sector is politically organized or it is not. This is at odds with the observation that participation shares are different across sectors as will be shown in the empirical section. In the following sections of the paper the goal is to show how θ_i^o is determined in equilibrium, what are the factors that affect its size, and how an empirical measure of this share significantly affects the explanation of the variation of protection across sectors.

Equation (7) is the one referred to in the empirical section, while the remainder of the theoretical section makes use of a simplified version of this equation that allows to identify the set of firms L_i for each sector. The following is a set of assumptions used in the remaining section of the model. None of the assumptions I am about to lay out should significantly affect the generality of the results.

Assumption A1: Leontief Production Function

The supply side of this economy is characterized by a very simple production decision. Remember that firm j in sector i is endowed with a certain amount of the specific factor K_{ij}. The production function for good i is Leontief:

$$x_{ij} = \min\{K_{ij}, L_{ij}\}$$

Since there is a perfectly elastic supply of labor at $w = 1$ this amounts to the firm producing:

$$x_{ij} = \begin{cases} K_{ij} & if\ p_i \geq 1 \\ 0 & otherwise \end{cases}$$

Because the lower bound on the domestic price is the international price $p_i^* = 1$ then firms always produce the maximum amount they can and they employ K_{ij} units of labor. The rent earned by the owner of the firm's specific factor depends positively on the domestic price:

$$\Pi_{ij}(\mathbf{p}) = K_{ij}(p_i - 1)$$

Assumption A2: Linear Demand Function

The demand function for good x_i is linear in price p_i:

$$d_i(p_i) = D_i - b_i p_i$$

This assumption simplifies the expression for m_i' which is now a constant, $-b_i$.

Assumption A3: Concentrated specific factor ownership

I assume that the owners of a sector's specific factor represent a negligible fraction of the voting population, that is $\alpha_{ij} = 0$. As a result there will be no "competition" among lobbies representing different sectors. In the absence of this assumption lobbies would make contributions in order to lower the price of all the goods they consume (and increase the domestic price of the good they produce). I therefore assume that each lobby makes contributions with the only goal of raising the price of the good it produces.

Proposition 2 *Under assumptions A1-A3 the equilibrium domestic price p_i^o takes the following expression:*

$$p_i^o = \frac{\theta_i^o K_i}{ab_i} + 1$$

where K_i is the total output of sector i and θ_i^o is the share of total output in sector i produced by firms making positive contributions:

$$\theta_i^o = \frac{\sum_{j \in L_i} K_{ij}}{K_i} \tag{8}$$

Having determined the equilibrium price for a given set of participants in the lobby, I move to discuss how to determine the set of firms that lobby in equilibrium and therefore pay the fixed cost of contributing.

4 Lobby participation

This section focuses on the determination of firm's participation in the lobbying process. From now on I suppress the subscript i since the equilibrium tariff in each sector is determined independently from other sectors. The sector has n firms ordered such that firm 1 is the smallest and firm n is the largest:

$$K_1 < ... < K_n$$

This game admits multiple equilibria, both in terms of the set of firms that make positive contributions in equilibrium and in terms of the contribution levels that support the equilibrium price. For a given set of firms L the equilibrium price in the sector is determined by equation (8), but in general many levels of contributions by individual firms can support this price level (implying different net payoffs for the contributing firms).

Although we cannot eliminate indeterminacy over the level of contributions, I argue that there is an optimal set of contributing firms in equilibrium and that it is reasonable to expect the selection of such equilibrium if firms are allowed to communicate. Because of the fixed nature of the cost of contributing I show that it is optimal to exclude firms below a certain size from contributing since the surplus created does not justify the payment of a fixed cost.

Consider firm h of size K_h and an arbitrary set of contributing firms L. In considering whether it is optimal for firm h to join the lobby let us consider the joint surplus of firm h and the lobby L. If the joint surplus is higher under firm h participation then it is optimal for firm h to contribute in equilibrium, otherwise it is optimal for firm h to be excluded from the lobby (and therefore save the fixed cost F).

Due to the timing of the game and the absence of competition from other lobbies, firms contribute just enough to compensate the government for the welfare loss relative to free trade. So the necessary contributions to compensate the government for the choice of a positive level of protection p^o are:[4]

$$a\left(W\left(1\right)-W\left(p^{o}\right)\right)=\frac{ab}{2}\left(p^{o}-1\right)^{2}$$

We now have all the elements to calculate the surplus of lobby L and candidate participant firm h under firm h participation and firm h exclusion from the lobby. Denote by p^o the equilibrium domestic price resulting from the interaction of the set of firms L and the government:

$$p^{o}=\frac{\sum\limits_{j\in L}K_{j}}{ab}+1$$

The joint surplus for the lobby is the difference between the gross profits and the necessary contri-

[4]The general expression for the aggregate welfare of the economy is
$W\left(p_{1},...,p_{m}\right)=1+\sum\limits_{i=1}^{m}\left(D_{i}-b_{i}p_{i}\right)\left(\frac{1}{2}p_{i}+\frac{D_{i}}{2b_{i}}-1\right)$
This separable form of aggregate welfare allows to isolate the term that depends on the domestic price of a specific good, say good 1: $W\left(p_{1},...,p_{m}\right)=A+\left(D_{1}-b_{1}p_{1}\right)\left(\frac{1}{2}p_{1}+\frac{D_{1}}{2b_{1}}-1\right)$
where $A=1+\sum\limits_{i=2}^{m}\left(D_{i}-b_{i}p_{i}\right)\left(\frac{1}{2}p_{i}+\frac{D_{i}}{2b_{i}}-1\right)$

butions, including the fixed cost of contributing for each firm:[5]

$$\sum_{j \in L} \left(\Pi_j \left(p^o\right) - F\right) - a \left(W\left(1\right) - W\left(p^o\right)\right) = \sum_{j \in L} \left(K_j \left(\frac{\sum_{j \in L} K_j}{ab}\right) - F\right) - \frac{\left(\sum_{j \in L} K_j\right)^2}{2ab} \tag{9}$$

Now consider the price that would prevail if firm h did not make positive contributions in equilibrium, p^{-h}:

$$p^{-h} = \frac{\sum_{j \in L, j \neq h} K_j}{ab} + 1$$

The joint surplus for the lobby and firm h when firm h does not contribute in equilibrium depends on the gross profits for both the lobby and firm h and the necessary contributions at the lower price p^{-h}. With firm h exclusion there is also a benefit in terms of reduced resources spent on fixed costs of contributing:

$$\begin{aligned} & \sum_{j \in L, j \neq h} \left(\Pi_j \left(p^{-h}\right) - F\right) + \Pi_h - a \left(W\left(1\right) - W\left(p^{-h}\right)\right) & (10) \\ = \; & \sum_{j \in L, j \neq h} \left(K_j \left(\frac{\sum_{j \in L, j \neq h} K_j}{ab}\right) - F\right) + K_h \frac{\sum_{j \in L, j \neq h} K_j}{ab} - \frac{\left(\sum_{j \in L, j \neq h} K_j\right)^2}{2ab} & \end{aligned}$$

It is optimal to include firm h in the lobby if expression (9) is larger than expression (10) and this inequality is satisfied if and only if:

$$\frac{K_h^2}{2ab} \geq F \tag{11}$$

This condition tells us that there is an optimal threshold for the inclusion of firms in the lobby. Therefore if we allow non-binding communication among firms we can expect the firms that consider forming a lobby to select such an equilibrium among the set of Nash equilibria in this game.

Condition (11) can be interpreted considering two opposite effects. As a large firm considers lobbying, it must pay larger contributions because it induces the government to grant larger protection (which brings about a larger welfare loss), but it also experiences a larger gain as it produces a greater quantity of output. Only the participation of larger firms produces a benefit large enough from protection to find it profitable to pay the initial fixed cost and enter the lobby. Notice that in the absence of a fixed cost it is efficient for all firms to make positive contributions in equilibrium.

The derivation above is summarized in the following proposition where I define the lobby selected using the criterion above as the Optimal Lobby.

Proposition 3 *In the Optimal Lobby equilibrium:*

[5] In reporting the firm's benefit I omit the amount of labor income as it is a constant that drops out of all the relevant expressions.

(i) if firm h enters the political game then all firms j, with $j \geq h$, enter the political game

(ii) if firm h^* is the smallest firm participating in the lobby, the following conditions must be satisfied:

$$\begin{aligned} \frac{K_{h^*}^2}{2ab} &\geq F \\ \frac{K_{h^*-1}^2}{2ab} &< F \end{aligned}$$

Proposition 3 allows to determine the equilibrium share of total output produced by firms lobbying for protection. Given the threshold participating firm h^* the equilibrium share θ^o is:

$$\theta^o = \frac{\sum\limits_{j \geq h^*} K_j}{\sum\limits_{j \in S} K_j}$$

4.1 Size distribution and the level of protection

This section investigates whether the model built in the previous sections can shed light on the impact on protection levels of certain characteristics of the size distribution of firms in a sector. In particular this section shows that basic moments of the distribution, like mean and standard deviation of firm size, affect the equilibrium level of protection.

The intuition for this result relies on the likelihood that in a sector where, holding the average constant, the size distribution of firms has a larger standard deviation, we can find a greater number of firms that are large enough to overcome the initial fixed cost of lobbying and find it profitable to participate in the political game.

To simplify the problem and without loss of generality the distribution of firm size is approximated using a continuous Pareto distribution. The model so far has dealt with firms of discrete size: it is fundamental to the structure of the model that the firm perceives its impact on the price level as it decides to lobby. Nevertheless employing a continuous density function in this section does not affect the results about the impact of dispersion on protection and simplifies considerably the relevant expressions. The choice of density function is dictated by documented empirical evidence that the distribution of firm size is well approximated by a Pareto distribution as reported by Axtell (2001) and Helpman et al. (2004).

Take a continuous of firms of size K_j and let K_j be distributed over the support $[K_M, \infty]$ according to the following probability density function:

$$K_j \sim \varepsilon \frac{K_M^{\varepsilon}}{K_j^{\varepsilon+1}}$$

where ε is a parameter[6]. Construct θ^o as the share of output produced by firms that participate in the lobby. Indicate the threshold firm as h^*.

$$\theta^o = \frac{\int_{K_{h^*}}^{\infty} K_j \varepsilon \frac{K_M^{\varepsilon}}{K_j^{\varepsilon+1}} dK_j}{\int_{K_M}^{\infty} K_j \varepsilon \frac{K_M^{\varepsilon}}{K_j^{\varepsilon+1}} dK_j} = \left(\frac{K_{h^*}}{K_M}\right)^{1-\varepsilon}$$

We are interested in the effect on θ^o of a mean-preserving spread in the size distribution of firms. As ε decreases, the right tail of the distribution gets thicker and dispersion increases, but in order to maintain the average size of the firm constant the lower bound K_M has to decrease. Indicate μ as the average size of the firm. In order to keep the average size constant the lower bound K_M is lowered according to the following expression:

$$K_M = \mu \frac{\varepsilon - 1}{\varepsilon}$$

Therefore we can rewrite θ^o as:

$$\theta^o = \left(\frac{K_{h^*}\varepsilon}{\mu(\varepsilon - 1)}\right)^{1-\varepsilon}$$

The impact of a mean-preserving spread in the size distribution of firms on θ^o (and therefore on the domestic price level) will depend on whether the threshold firm is sufficiently above the mean of the distribution as characterized in the following proposition.

Proposition 4 *If the following condition is satisfied:*

$$\ln \frac{K_{h^*}}{\mu} + \ln \frac{\varepsilon}{\varepsilon - 1} > \frac{1}{\varepsilon} \tag{12}$$

then a decrease in ε (an increase in dispersion), ceteris paribus, brings about an increase in the domestic price p.

Condition (12) is always satisfied if $K_{h^*} > \mu$. As described in the following section, I can identify the threshold participating firm and I verify that for all 226 sectors used in this study, with the exception of one, the threshold contributing firm is larger than the average firm.

5 Empirical strategy and data description

The empirical section is organized according to the following road map. First, I provide evidence of the motivating fact of this paper using a reduced form approach: sectors characterized by a higher dispersion in firm size present higher levels of protection.

[6] I impose the condition $\varepsilon > 2$ to guarantee that the distribution has finite variance.

Second, the model predicts that larger firms are more likely to take part in the lobby and are likely to contribute more. Making use of firm-level data I show that this prediction is confirmed. Third, employing the same firm-level data, I build the share of total output in each sector produced by firms that are part of the lobby. The model predicts that these participation shares are increasing in the level of firm size dispersion. I show that this is confirmed by the data. Fourth, I test the prediction that the level of protection depends not simply on the sector's total output, but on the share of output produced by lobbying firms. Fifth, I test the model developed in this paper against the benchmark "Protection for sale" model and show that it explains a significantly larger fraction of the variation in protection levels across sectors.

5.1 Reduced form evidence: the effect of firm size dispersion on protection

The empirical section of this paper makes use of several data sources: the data used in previous empirical studies to test the original GH model is the same as in Gawande and Bandyopadhyay (2000); the data on sector-level firm size distribution is from the 1987 US Census of Manufactures; the data on firms political contributions is taken from Federal Election Commission records for electoral cycle 1986-88 and individually matched to COMPUSTAT firm information about sales, employment and industry classification.

The motivation of this paper stems from an empirical pattern observed in the data: sectors where firm size distribution is more dispersed are more highly protected. This section presents evidence of this pattern. I first need to introduce all the other variables that have been used in the previous empirical tests of the "Protection for sale" framework and the data used to measure firm size dispersion.

5.1.1 The benchmark model and data description

I will use the model presented by Gawande and Bandyopadhyay (2000) (from now on GB) as benchmark of the original "Protection for Sale" model, because the same data set is employed here. Their specification is a system of three equations, of which I will emphasize only one as it is relevant to this study:[7]

$$\frac{t_i}{1+t_i} = \gamma_0 + \gamma_1 I_i \frac{z_i}{e_i} + \gamma_2 \frac{z_i}{e_i} + Z_{1i} + \varepsilon_i \tag{13}$$

[7] This is the specification used to test the main "Protection for sale" equation:
$\frac{t_i}{1+t_i} = \frac{I_i - \alpha_L}{a + \alpha_L} \frac{z_i}{e_i}$
(Grossman and Helpman (1994) p.842)

where t_i is the coverage ratio for industry i, z_i is the inverse of the import penetration ratio, e_i is the price elasticity of imports and I_i is a dummy that describes whether the sector is politically organized, while Z_{1i} includes tariffs on intermediate goods as controls as in GB.

The equation:

$$\frac{1}{z_i} = \phi \frac{t_i}{1 + t_i} + \xi_i \tag{14}$$

accounts for the simultaneity problem first studied by Trefler (1993): we can expect higher tariffs to reduce import penetration as this equation illustrates. This system accounts for the fact that import penetration and tariff levels are determined simultaneously.

As a measure of protection t_i the literature has widely adopted the use of coverage ratios for non-tariff barriers,[8] which represent the share of products in an industry covered by one or more quantitative or qualitative restrictions to trade. Although the model deals with tariffs, from a practical point of view we generally regard tariffs as the result of multilateral agreements that take place among governments. It is standard to consider NTB's to be a more flexible instrument available to governments since qualitative and quantitative restrictions to trade are not governed by the same strict rules that apply to tariffs and are set by international organizations such as the WTO. Moreover, interest groups are aware that tariffs are set at the international level and that NTB's are easier to set unilaterally, therefore becoming the target of industry lobbying. Furthermore, tariffs are generally very low on all manufacturing products following the rounds of negotiations under the GATT and WTO and variation in tariff levels is not likely to be very wide. Import penetration ratios measure the share of imports to total production in sector i. In the equation above z_i is the inverse of the import penetration. Data on import penetration ratios are the same as in Trefler (1993). As for e_i, the literature has reached a consensus in considering the study by Shiells et al. (1986) as the most accurate estimate of sector-level price elasticity of imports. I_i is a dummy variable that indicates whether the sector is politically organized and represented by a lobby. In this study I use the dummy constructed by GB[9], but I show alternative results under a dummy that I construct using other data sources. The data on political contributions used by GB to construct I_i are originally from the Federal Election Commission.[10]

I employ the instruments used by Trefler to correct for the simultaneity bias intrinsic to the system of equations (13) and (14). The two variables employed here are sector-level capital labor ratios

[8] Data on Non-Tariff Barriers is relative to year 1987 and is collected by UNCTAD. The data was kindly provided by Kishore Gawande.

[9] See Gawande and Bandyopadhyay (2000) for detailed explanation of the derivation of I_i.

[10] The data provided by Kishore Gawande reports sector-level aggregate contributions by Political Activity Commitees (PAC's) for the 1981-82 and 1983-84 election cycles.

interacted with industry dummies and the fractions of managers, scientists and unskilled labor per industry as measures of comparative advantage that determine import penetration independently of the level of protection.

5.1.2 Introducing characteristics of the size distribution of firms

The novel fact that this paper uncovers is that sectors that are characterized by higher firm size dispersion receive higher protection. I present here reduced form evidence, which is not directly a test of the model, but that suggests that relevant variables are omitted in previous empirical studies of the GH model. The basic specification employed in this section is:

$$\frac{t_i}{1+t_i} = \gamma_0 + \gamma_1 I_i \frac{z_i}{e_i} + \gamma_2 \frac{z_i}{e_i} + \gamma_3 \sigma_i + \gamma_4 \mu_i + \gamma_5 I_i + Z_{1i} + \varepsilon_i \tag{15}$$

where μ_i and σ_i are respectively average and standard deviation of firm size within sector i . Notice that GH would imply that γ_3 and γ_4 are both zero as the size distribution of firms should have no impact on the level of protection. The source of data employed to measure these two variables is the 1987 US Census of Manufactures[11]. The Industry Series of the Census of Manufactures include data on 4-digit SIC industry firm size distribution. I treat each 4-digit SIC code as an individual sector and I approximate firm size using total annual sales.[12] The average size and size dispersion are respectively the mean μ_i and the standard deviation σ_i of the firm sales distribution. Publicly available US Census data sets do not report individual firm information, but they report ten size brackets according to employment size. For each bracket total shipments and the number of establishments is reported. As a result of this data limitation the method employed is to calculate the average size per employment bracket and then derive the weighted average and standard deviation across the ten available employment bins.[13]. *Table* 1 reports summary statistics for size distribution variables. The first part of the table includes 226 4-digit SIC sectors.

To correct for the possibility of reverse causality in the relationship between firm size distribution and the level of protection I instrument for the standard deviation of firm size using size dispersion measures for European firms. The data on firm size dispersion for European firms (the mean and the standard deviation of sales), used as an instrument for the characteristics of the size distribution

[11] 1987 Census of Manufactures Volume 1E - Industry Series MC87-I. Industry statistics by employment size of establishment MC87I4

[12] The entry reported in the US Census of Manufactures data is "shipments".

[13] See Appendix for details. A similar calculation of average and standard deviation is made in Helpman et al. (2004).

of US firms is from Helpman et al. (2004)[14]. In principle the instrumenting strategy should also correct for the fact that σ_i and μ_i are estimates of the true moments of the sector-level distribution of firm size.

5.1.3 Results

Column GB in *Table* 2 reports the results for specification (13)[15]. As predicted by the general GH model, for politically unorganized sectors a larger size of the industry output relative to imports and a smaller price elasticity of imports decreases the tariff level ($\gamma_2 < 0$) The point estimate of γ_2 is -1.73 with a robust standard error of 0.70. For politically organized sectors this relationship has the opposite sign ($\gamma_1 > 0$). The point estimate of $\gamma_1$is 1.83 with a standard error of 0.74. Both point estimates are of the same magnitude as in GB. Although the prediction that the sign of γ_1 is positive is confirmed by the data, the positive sign of $\gamma_1 + \gamma_2$ is not statistically significant. It is important to stress that this prediction is confirmed when considering the specification that is closest to the model presented in this paper, as emphasized in the following section.

The results from specification (15) appear in column I of *Table* 2. While the coefficients on $I_i\,(z_i/e_i)$ and z_i/e_i remain of the same sign and magnitude (suggesting that the GH model is robust), the standard deviation of firm size presents a positive and very significant impact on protection levels, controlling for the average size of the firms in the sector. The point estimate of γ_3 is 0.44 (precisely estimated with a robust standard error of 0.063). This is consistent with the prediction of the model that sectors where firms' distribution is more dispersed receive more protection. The coefficient γ_4 is 0.04 on average firm size (not statistically significant with a robust standard error of 0.04).

The centered R^2 for column I is 33 percent larger than the benchmark GB column which suggests that including measures of firm size distribution explains a larger fraction of the variation of protection levels across sectors.

In other columns of *Table* 2 I control for other variables that could be affecting the structure of protection across industries. I control for Total Sales in the sector, accounting for the fact that the data used for import penetration might not be the same as the Sales data used to calculate firm size dispersion. I also control for Total Value Added per sector and for more partial measures of firm distribution as the Herfindahl index and the Concentration4 index[16]. When included in the

[14] Data on European firms is orginally from Amadeus (1997).

[15] The results are consistent with the coefficients reported in the paper by Gawande and Bandyopadhyay (2000) and qualitatively similar to the values obtained by Goldberg and Maggi (1999).

[16] Concentration4 reports the share of total sales accounted for by the top four firms in the sector.

same regression the measure of dispersion presented here remains the only significant distribution variable affecting the structure of protection[17].

5.2 Testing the model

This section employs data on firm-level contributions and firm characteristics to test the last four predictions listed in the road map laid out in the empirical strategy.

One new source of data is a collection of records of the Federal Election Commission of political contributions. I refer to the Appendix for more details about these data. The FEC holds a record of all Political Action Committees formed. PAC's are a channel through which corporations, among other entities, make contributions to politicians (mostly to incumbent politicians).[18] The data set used in this study reports information for various electoral cycles. In order to make the data on contributions compatible with 1987 US Census of Manufactures data I consider the political cycles 1986 and 1988. This data set though, originally lacks a standard identifier for the company sponsor of the PAC, which required to individually match each PAC to a firm. Therefore each PAC was matched to a firm in COMPUSTAT using the PAC name, as explained in detail in the Appendix.

The fourth source of data is COMPUSTAT North America Industrial Annual 1987. This data set provides information on company's employment size, annual net sales and 4-digit SIC. As described in the Appendix, data on PAC contributions and COMPUSTAT data on firm size allow to find the threshold participating firm for each SIC 4 industry. Once the threshold firm is identified, θ_i can be calculated as the share of total output in the industry represented by the firms participating in the political game. In the second and third part of $Table$ 1 firm level summary statistics of PAC contributions, net sales and employment levels are reported. The number of firms for which data are available in 1988 is 3089, of which 478 make positive contributions (15.47% of the available sample). Among the contributing firms the average contribution is $62,241$ dollars. The fact that political contributions are small relative to the size of contributing firms has been documented

[17] Davis (1990) suggests an alternative way to compute statistics of the distribution of firm size. In particular Davis proposes the following "size-weighted" average firm size, define it $w\mu_i$:

$$w\mu_i = \sum_{j \in S_i} x_{ij} \frac{x_{ij} n_{ij}}{X_i}$$

where n_{ij} is the number of firms with same size index j. The relationship between $w\mu_i$ and μ_i and σ_i is the following:

$$w\mu_i = \mu_i + \frac{\sigma_i^2}{\mu_i}$$

Empirically $w\mu_i$ affects positively the level of protection as both its components μ_i and σ_i have been shown to have a positive impact on protection.

[18] I thank Jim Snyder for kindly sharing these data.

by Ansolobehere et al. (2003). The small size of political contributions should not be surprising when what should matter is the leverage that a given amount of money obtains. Furthermore it is plausible that political contributions are not the only channel through which lobbying takes place. In particular many corporations decide to hire lobbyists whose task is to promote legislation that is favorable to a specific industry. Although this paper does not show evidence of this channel, the logic of the model should apply as long as this form of lobbying also involves an initial fixed cost independent of firm size.

I will refer in many instances to the Data Appendix for a thorough description of both the data and the methodology used to link the different data sets.

5.2.1 Firm size and the likelihood of participating in the political game

A prediction of the model is that larger firms are more likely to participate in the political game and make larger contributions than smaller firms. It would be naive to expect that all large firms in an industry contribute. This is because contributions are not the only way in which firms might try to influence policy choices. Nevertheless, in this section evidence is reported to show that larger firms are indeed more likely to make positive contributions and make larger contributions.

The model presented in this paper distinguishes theoretically between the participation and the contribution decision of the firm. Therefore it would be optimal to estimate the participation and the contribution decision simultaneously.[19] This would require a measure of the fixed cost in order to identify the participation decision separately from the contribution decision. Since such measure is not available, evidence of the two decisions is reported separately.

According to the first panel in *Table* 3 the amount of contributions by each firm increases as a function of size (measured as the logarithm of sales). The model also predicts that contributions should be zero up to the industry participation threshold and then increase with the size of the firm. This convex shape of contributions as a function of size is confirmed by *Graph* 1 and by the pooled regression in *Table* 3 where square of the logarithm of sales has the predicted positive coefficient (0.001 with SIC 4 clustered standard error of 0.0002). The average intercept across sectors is -0.013 and the average slope is 0.004 (see column III, *Table* 3).

The second panel in *Table* 3 also shows that the probability that a firm participates in the lobby is increasing in the size of the firm both across sectors and within each sector. Moreover, the third panel in *Table* 3 shows that the ratio of participating firms to non-participating firms is increasing as we consider larger firm size categories (with a coefficient of 0.058), both at the aggregate and

[19] This is the standard Heckman (1979) correction for sample selection.

sector level.

5.2.2 The relationship between θ_i and size distribution parameters

The model predicts that the share of total output represented by firms that participate in the lobby increases as the average size and standard deviation of the size distribution of firms increase.

$$
\begin{aligned}
\frac{\partial \theta_i}{\partial \sigma_i} &\geq 0 \\
\frac{\partial \theta_i}{\partial \mu_i} &\geq 0
\end{aligned}
$$

Using firm-level data on size and contributions one can identify the threshold contributing firm for each sector i, that is the smallest firm that makes positive contributions (in the model firm h^* for each sector i) and one can measure the share of total output of sector i represented by firms participating in the political game, that is what I call "true" θ_i[20]

Identifying the threshold participating firm allows to check that the condition, under which a mean-preserving spread of the firm size distribution induces an increase in θ_i, is satisfied. In a the sample of 226 sectors, the condition that the threshold participating firm is larger than the average firm is satisfied for 225 of these industries and is not satisfied for 1.

The share of total output of sector i represented by the firms contributing is positively correlated to both the mean and the standard deviation of the firm size distribution. *Table* 4 reports a correlation of 0.47 between θ_i and σ_i and a correlation of 0.39 between θ_i and μ_i. Both correlations are positive and significant at the 1 percent confidence level, consistently with what is predicted by the model.

5.2.3 The Heterogeneity specification

The model predicts that the correct specification that describes the equilibrium level of protection of an industry should account for different participation shares θ_i's. Once firm-level data allows to calculate the share of total output produced by firms lobbying, the model predicts exactly how it should enter the protection equation:[21]

[20] I define this θ_i "true" because it is calculated using actual data on contributions. I distinguish this from the "constructed" θ_i which is the function described below.

[21] The specification still includes I_i to account for the fact that some sectors cannot be considered politically organized even though they make some small political contributions. This choice is made in order to maintain the set of organized sectors the same as in GB. The results are therefore comparable and the differences are not due to a different set of organized sectors.

$$\frac{t_i}{1+t_i} = \gamma_0 + \gamma_1 I_i \frac{z_i}{e_i} \theta_i(\sigma_i, \mu_i) + \gamma_2 \frac{z_i}{e_i} + Z_{1i} + \varepsilon_i \tag{16}$$

where $\theta_i(\sigma_i, \mu_i)$ is an increasing function of both σ_i and μ_i (as shown in the previous section).

I show in the following section the results for this specification using:

- the participation shares, that are calculated using firm-level contribution and size data (call them "true" participation shares)
- the "constructed" shares, that is shares that are function of parameters that, according to the model, determine θ_i's: $\theta_i(\sigma_i, \mu_i)$

The model predicts that $\gamma_1 > 0$: in sectors where firms make political contributions the level of protection is higher the higher the output, the lower the imports, the lower the price elasticity of imports and the higher participation share θ_i. The model also predicts $\gamma_2 < 0$: in sectors that are not politically organized the level of protection is lower the higher the output, the lower the imports, the lower the price elasticity of imports.

The third prediction is that $\gamma_1\theta_i + \gamma_2 > 0$. In principle the sign of $\gamma_1\theta_i + \gamma_2$ should depend on whether $\theta_i - \alpha_L$ is smaller or greater than zero, but if we maintain the assumption that ownership is very concentrated, α_L will be relatively small and the sign of $\gamma_1\theta_i + \gamma_2$ should be positive. In other words it seems reasonable to expect the political effort to lower a product's price on the part of consumers of the good to be weaker than the political effort to increase it by its producers.

5.2.4 The heterogeneity specification with "true" participation shares

I here test the main prediction of the model, the equation:

$$\frac{t_i}{1+t_i} = \gamma_0 + \gamma_1 I_i \frac{z_i}{e_i} \theta_i + \gamma_2 \frac{z_i}{e_i} + Z_{1i} + \varepsilon_i \tag{17}$$

using the share θ_i's calculated by finding the threshold participating firm for each sector.

As *Table* 5 reports, both coefficients γ_1 and γ_2 have the predicted sign: $\gamma_1 > 0$ and $\gamma_2 < 0$. The point estimate for γ_1 is 13.78 with a standard error of 3.24 and the coefficient γ_2 is -0.28 with a standard error of 0.11. The implication of the results of this specifications is that an increase of 1 percent of the share of sales represented by contributing firms induces an increase of 13.78 percentage points in the coverage ratio. Moreover this specification supports the hypothesis that $\gamma_1\theta_i + \gamma_2 > 0$: calculated at the average $\overline{\theta}$, $\gamma_1\overline{\theta} + \gamma_2 = 3.34$ and the 95 percent confidence interval is [2.88, 3.80]. Notice that the equivalent prediction in the GH model is not supported by the results obtained by GB.

Estimating γ_1 and γ_2 allows to find the weight that the government places on aggregate welfare relative to contributions. We can rewrite the government's utility as $G = C + a(W^N + C)$ where W^N is the aggregate welfare net of contributions. Therefore the weight on net aggregate welfare is a while the weight on contributions is $(1 + a)$. The estimates of γ_1 and γ_2 imply that $a = 725$ and that the government places an equal weight on net welfare and contributions.

Another important implication of this regression is that the value for α_L is around 0.02,[22] which implies that a low fraction of the population is represented by interest groups. This value for α_L is much lower than the values obtained by Gawande and Bandyopadhyay (2000), who find $\alpha_L \approx 1$, and Goldberg and Maggi (1999), who find values of α_L between 0.84 and 0.88. These values of α_L have been recognized as unrealistically high for the US and are commonly indicated as a failure of the empirical tests of the "Protection for sale" model. The more reasonable estimate of α_L obtained under the Heterogeneity specification offers further support to the need to account for different participation shares (θ_i's) across sectors.

5.2.5 The heterogeneity specification with "constructed" participation shares

I here test the equation (17) using "constructed" thetas. θ_i's are shown to depend positively on both the mean and the standard deviation of the size distribution in the sector. I choose the simplest functional form for $\theta_i\left(\sigma_i, \mu_i\right)$:

$$\widetilde{\theta}_i = \rho \mu_{iF} + (1 - \rho)\, \sigma_{iF} \tag{18}$$

where $\mu_{iF} = F_\mu\left(\mu_i\right)$[23] and $\sigma_{iF} = F_\sigma(\sigma_i)$[24] are normalized values of σ_i, μ_i.[25]

Results in *Table* 6 show that different weights (ρ's) do not substantially affect the coefficients γ_1 and γ_2 in specification (17). The coefficients have the predicted sign with γ_1 positive and significant and γ_2 negative and significant. Moreover the prediction that $\gamma_1 \theta_i + \gamma_2 > 0$ is supported by the data. These last findings might also be a consequence of the re-scaling of the variable $I_i\left(z_i/e_i\right)$ by θ_i which is a quantity between 0 and 1, but using θ_i's seems to bring the data closer to the predictions of the model.

5.2.6 Comparing the Heterogeneity model and the "Protection for sale" model

It is difficult to draw any sharp conclusions from comparing results from specifications (13) and (16) because the two models are such that model (16) is not nested into model (13).

[22] The model implies that $\alpha_L = -\frac{\gamma_2}{\gamma_1}$.

[23] Empirical cumulative distribution function of μ_i

[24] Empirical cumulative distribution function of σ_i

[25] $\rho \in (0, 1)$ is a weight to which I assign several values. I report results for different values of ρ in *Table* 6.

I adopt a methodology introduced by Davidson and MacKinnon (1981) and employed by Eicher and Osang (2002) to compare the two models' power in explaining the pattern of protection across sectors. The goal is to test whether the model proposed in this paper explains significantly more of the variation in observed NTB's than the original GH model does.

The procedure introduced by (Davidson and MacKinnon (1981)) consists in non-nested J-tests where two types of tests are performed.

Table 7 reports results for these two tests under "true" θ_i's . In *Test I* the null hypothesis is that the GH model is the correct one and the alternative hypothesis is that the model proposed in this paper, which we call "Heterogeneity" model, does not add any explanatory power. I reject the null hypothesis that the GH model is the correct one as I find that the Heterogeneity model adds explanatory power.

Test II considers as null hypothesis that the correct model is the Heterogeneity model and as alternative hypothesis the GH model. This test finds that one cannot reject the null hypothesis that the Heterogeneity model is the correct one at the 1 percent confidence level (one can reject that the Heterogeneity model is the most informative at the 5 percent confidence level, which suggests that there might be some information in the GH model that is not encompassed by the Heterogeneity model).

Table 8 reports results for the same two tests under "constructed" θ_i's. In this case both tests strongly support the Heterogeneity model as having stronger explanatory power than the original GH model. In particular, *Test II* here cannot reject the null hypothesis that the Heterogeneity model is the correct one at all confidence levels.

I conclude that according to these tests the model proposed in this paper seems to have substantially greater explanatory power over the original GH model and therefore that a correct model of the political economy of trade policy should include firm-level decisions and account for the fact that sectors present different size distributions of firms and difference lobby participation shares.

6 Conclusions

This paper provides a micro-foundation of individual firms' lobbying behavior and develops a model that helps explain a number of empirical features shown in the data. In particular the model explains why larger firms are more likely to lobby and to contribute more and offers a channel through which the size distribution of firms affects lobby participation shares and therefore the level of protection in a sector. This paper shows that accounting for individual firm behavior and differences in participation shares across sectors helps explain a larger fraction of the variation of protection across sectors and therefore improves on existing theoretical and empirical studies of endogenous protection that employ the "Protection for sale" framework.

Nevertheless this paper constitutes a first attempt to study individual firm lobbying behavior and focuses on a specific game structure and equilibrium selection criterion. It might be interesting to explore different game structures that might explain participation decisions through different mechanisms than the method involved here.

A promising line of research recently undertaken by Gawande and Li (2004) explores the consequences of considering the government not as a unitary entity, but as a more complex and uncertain counterpart to interest groups. It would be interesting to explore from a theoretical and empirical point of view the possible competition among politicians to obtain contributions from interest groups. The conjecture is that in the presence of fixed costs for the politicians in contacting firms to obtain political support, the size distribution of firms within a sector should affect the level of protection granted to the industry. In future work I intend to develop a theoretical framework that can account for this mechanism.

Moreover it would be interesting to further explore from an empirical point of view how the decision to lobby depends on the industry and other characteristics of the firm: this might shed some light on the nature of the fixed cost necessary to start political activity and other determinants that are not considered in this paper.

7 Appendix

7.1 Federal Election Commission contributions data and COMPUSTAT individual company information

The data set used to identify the firms that participate in the political game was provided by Jim Snyder and is taken from the Federal Election Commission (FEC). The FEC collects information about all Political Action Committees formed: it provides the PAC's name and the sponsor's name, along with data on contributions for all electoral cycles from 1978 to 1998. The FEC identifier does not correspond to any standard company classification so it is necessary to use the name of the PAC sponsor to individually match each PAC to a company listed in COMPUSTAT. In this process I made use of a publication by Congressional Quarterly that describes the sponsors of most corporate PAC's. I was not able to match all the PAC's to an individual firm using COMPUSTAT company information, but a reasonable effort was made to look for links between PAC's and companies through company affiliations and subsidiaries (this information was taken from the each company's website). The PAC data set included 3700 entries of which 2040 were matched to individual companies available on COMPUSTAT. Disregarding the banking, insurance, utilities and health sectors, that are not relevant to this study, I can assess the number of unmatched firms to below 500. This number includes several PAC's that I could not classify in any other sector and I therefore reported as potentially relevant to my study (i.e. manufacturing sectors). It is plausible to have introduced some selection bias using this matching procedure: COMPUSTAT covers publicly traded companies, which are plausibly the largest in the industry. In identifying the industry participation threshold with the smallest contributing firm that matched I am potentially overestimating the participation threshold and underestimating the share of total industry output represented by contributing firms. Nevertheless companies that contribute in manufacturing industries are predominantly publicly traded and, among the PAC's I was not able to match, a large number is private and large (according to CQ PAC's Directory).

7.2 Construction of the characteristics of the size distribution of firms

This section describes the construction of sector-level size dispersion measure. The 1987 US Census of Manufacturing (henceforth USCM) reports the value of total shipments by SIC 4 and provides a breakdown of the total shipments by employment size of the establishment according to ten brackets reported below. The variable *Emplsize* indicates the employment bracket and associated to it is a range that describes the number of employees per establishment in that size category. Below I will describe the implications for my study of the choice of establishment as the unit of observation and I will report the adjustments that I was able to make.

US Census of Manufacturing 1987

Employment brackets

Emplsize	Number of Employees*
1	$0-4$
2	$5-9$
3	$10-19$
4	$20-49$
5	$50-99$
6	$100-249$
7	$250-499$
8	$500-999$
9	$1000-2499$
10	>2500

*per establishment

The USCM reports the total value of shipments S_i and the total number of establishments n_i for each employment size category $i,\ i=1,...,10$. The average and the standard deviation of shipments are calculated as follows:

$$\mu = \frac{\sum_{i=1}^{10} S_i}{\sum_{i=1}^{10} n_i}$$

$$\sigma = \left(\frac{\sum_{i=1}^{10} n_i \left(\frac{S_i}{n_i} - \mu\right)^2}{\sum_{i=1}^{10} n_i} \right)^{\frac{1}{2}}$$

7.3 Determination of the participation threshold

The FEC reports data on individual firm contributions and therefore cannot be matched to the Census of Manufacturing data directly as firms are generally composed by several establishments. An establishment is defined by the USCM as the "A single physical location where business is conducted or where services or industrial operations are performed", whereas a firm is defined as "A firm is a business organization consisting of one or more domestic establishments in the same state and industry that were specified under common ownership or control". We therefore need to know how many establishments belong on average to each firm of a given size in a given sector. The method used to impute a firm to one of the ten employment categories in the 1987 Census of Manufacturing requires the use of the 1992 Statistics of US Businesses (henceforth SUSB) data on industry (SIC 4) employment. The 1992 SUSB classifies enterprises[26] in each industry according to the number of employees and reports for each of the six employment categories (which are different from the USCM employment breakdown) the total number of firms and the total number of establishments. It is possible to derive the average number of establishments per firm for a company of a certain size (in terms of employees). After assigning each individual company to an industry and employment bracket (using COMPUSTAT data on employment and SIC category), I divide the number of employees of each firm by the corresponding number of average establishment per firm. I finally used this average number of employees per establishment to assign the firm to the USCM employment category. Identifying the smallest firm contributing in each sector then allows to find the participation threshold, which is the USCM employment bin where the threshold firm is classified.

References

Ansolobehere, Stephen, John M de Figueiredo, and James M Snyder, "Why Is There So Little Money in Politics?," 2003.

Antràs, Pol and Elhanan Helpman, "Global Sourcing," *Journal of Political Economy*, 2004, *112* (3), 552–580.

Aw, Bee Yan, Sukkyun Chung, and Mark J Roberts, "Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)," *World Bank Economic Review*, 2000, *14* (1), 65–90.

Axtell, Robert L, "Zipf Distribution of U.S. Firm Sizes," *Science*, 2001.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian, "On the Private Provision of Public Goods," *Journal of Public Economics*, 1986, *29* (1), 25–49.

[26] The SUSB distinguishes between firms and enterprise using the following criterion. An enterprise is a company that operates in more than one industry and therefore controls several firms across different industries (SIC 4). For my scope I do not make a distinction between the two entities.

Bernard, Andrew B and J Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, 1999, *47* (1), 1–25.

_ **, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum**, "Plants and Productivity in International Trade," *American Economic Review*, 2003, *93* (4), 1268–1290.

Bernheim, B Douglas and Michael D Whinston, "Menu Auctions, Resource Allocation, and Economic Influence," *Quarterly Journal of Economics*, 1986, *101* (1), 1–31.

Davidson, Russell and James G MacKinnon, "Several Tests for Model Specification in the Presence of Alternative Hypotheses," *Econometrica*, 1981, *49* (3), 781–93.

Eicher, Theo and Thomas Osang, "Protection for Sale: An Empirical Investigation: Comment," *American Economic Review*, 2002, *92* (5), 1702–10.

Findlay, Ronald and Stanislaw Wellisz, "Endogenous Tariffs, the Political Economy of Trade Restrictions, and Welfare," in J.N. Bhagwati, ed., *Import competition and response*, Chicago: University of Chicago Press, 1982, pp. 223–234.

Gawande, Kishore, "US Non-tariff Barriers as Privately Provided Public Goods," *Journal of Public Economics*, 1997, *64* (1), 61–81.

_ **and Hui Li**, "The Case of the Missing Contributions," 2004.

_ **and Usree Bandyopadhyay**, "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection," *Review of Economics and Statistics*, 2000, *82* (1), 139–52.

Goldberg, Pinelopi Koujianou and Giovanni Maggi, "Protection for Sale: An Empirical Investigation," *American Economic Review*, 1999, *89* (5), 1135–55.

Grossman, Gene M and Elhanan Helpman, "Protection for Sale," *American Economic Review*, 1994, *84* (4), 833–50.

Helpman, Elhanan, Marc J Melitz, and Stephen R Yeaple, "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 2004, *94* (1), 300–316.

Hillman, Arye L, *The political economy of protection*, Fundamentals of Pure and Applied Economics series, vol. 32 International Trade section. Chur, Switzerland; London; New York and Camberwell, Australia: Harwood Academic, 1989.

Magee, Christopher, "Endogenous Trade Policy and Lobby Formation: An Application to the Free-Rider Problem," *Journal of International Economics*, 2002, *57* (2), 449–71.

Magee, Stephen P, William A Brock, and Leslie Young, *Black hole tariffs and endogenous policy theory: Political economy in general equilibrium*, Cambridge; New York and Melbourne: Cambridge University Press, 1989.

Masters, Marick F. and Gerald D. Keim, "Determinants of PAC Participation Among Large Corporations," *The Journal of Politics*, 1985, *47* (4), 1158–1173.

Melitz, Marc J, "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Mitra, Devashish, "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination," *American Economic Review*, 1999, *89* (5), 1116–34.

Olson, Mancur, *The logic of collective action; public goods and the theory of groups* Harvard economic studies, v. 124, Cambridge, Mass.,: Harvard University Press, 1965.

Pecorino, Paul, "Is There a Free-Rider Problem in Lobbying? Endogenous Tariffs, Trigger Strategies, and the Number of Firms," *American Economic Review*, 1998, *88* (3), 652–60.

Quarterly, Congressional, *Congressional Quarterly's federal PACs directory*, Washington, D.C.: Congressional Quarterly Inc., 1998.

Rodrik, Dani, "Political Economy of Trade Policy," in Gene M. Grossman and Kenneth eds Rogoff, eds., *Handbook of international economics. Volume 3*, 1995.

Shiells, Clinton R, Robert M Stern, and Alan V Deardorff, "Estimates of the Elasticities of Substitution between Imports and Home Goods for the United States," *Weltwirtschaftliches Archiv*, 1986, *122* (3), 497–519.

Trefler, Daniel, "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*, 1993, *101* (1), 138–60.

Table 1 Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
		SIC 4 sample - US Census of Manufactures 1987			
Coverage Ratio	226	0.1	0.2	0.0	1.0
Total Shipments	226	6247.4	13639.8	92.9	133345.8
Average Shipments	226	23.3	59.8	0.5	639.2
Shipments St. Dev.	226	43.0	88.8	0.8	797.3
		Firm full sample - COMPUSTAT			
PAC Contributions	3089	8668.1	40517.9	0	526956
Total sales	3089	1524.5	8110.8	0	121816.6
Employees	2893	9389.4	46777.3	1	765700
		Contributing firm sample - COMPUSTAT			
PAC Contributions	478	62241.12	88590	100	526956
Total sales	374	7171.8	18988.1	3.496	121816.6
Employees	368	43507.86	110162.2	60	765700

Notes: Coverage Ratios from UNCTAD 1983. Shipments in millions USD. Compustat sample refers to North America Industrial Annual for year 1988. Net annual sales (Compustat series DATA12) in millions USD. PAC contributions in USD. Full sample refers to Compustat firms with available total sales data. Contributing firm sample includes only firms contributing to respective PACs.

Table 2 - Size distribution characteristics- Reduced form

Dependent Variable: NTB_i	Regression using Gawande Organization Dummy						Organization Dummy FEC data	
	GB*	I	II	III	IV	V	VI	VII
$I_i(z_i/e_i)$	1.83 (0.74)	1.97 (0.87)	1.47 (0.75)	1.56 (1.04)	1.53 (0.77)	1.58 (0.81)	2.3 (1.16)	1.55 (0.6)
z_i/e_i	-1.73 (0.70)	-1.82 (0.85)	-1.38 (0.73)	-1.46 (1.022)	-1.43 (0.75)	-1.44 (0.79)	-2.21 (1.15)	-1.62 (0.58)
σ_i(/1000)		0.44 (0.063)	0.37 (0.12)	0.04 (0.018)	0.39 (0.08)	0.48 (0.06)	0.42 (0.06)	
μ_i(/1000)		0.04 (0.04)	0.04 (0.05)	0.01 (0.08)	0.05 (0.05)	0.037 (0.047)	0.033 (0.049)	
I_i			0.01 (0.02)	0 (0.00)	0.01 (0.02)	0.008 (0.019)	0.021 (0.022)	0.024 (0.006)
Total Sales (/10M)			5.6 (7.61)	0.015 (0.007)				
Total Value Added (/1M)					1.5 (1.69)			
Concentration4						-0.015 (0.37)		
Herfindhal						-0.01 (0.009)		
$I_i\sigma_i$(/1000)								0.42 (0.02)
$I_i\mu_i$(/1000)								0.031 (0.008)
$N_i\sigma_i$(/1000)								-1.93 (0.92)
$N_i\mu_i$(/1000)								6.77 (2.1)
F-test joint σ_i μ_i**		0.00	0.00	0.04	0.00	0.00	0.00	0.00
F-test model**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
J-test overidentification**	0.33	0.2	0.26	0.47	0.23	0.19	0.35	0.37
Centered R^2	0.24	0.32	0.33	0.3	0.32	0.33	0.32	0.33
No. of Observations	226	226	226	226	226	226	226	226
Estimator	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS

Notes: *Gawande and Bandyopadhyay (2000) benchmark. **p-value reported. Robust s.e. in parentheses. All specifications include a constant and controls for intermediate goods tariffs and intermediate goods ntb's, not reported. I_i is a dummy variable taking value 1 if the sector is politically organized (from Gawande and Bandyopadhyay (2000)); z_i is the inverse import penetration ratio divided by 10000 (the import penetration ratio is the ratio of imports to domestic production); e_i is the price elasticity of imports; μ_i is the average of per firm shipments in sector i; σ_i is the standard deviation of per firm shipments in sector i; N_i is defined as $(1-I_i)$. In column III the average and the standard deviation of log(shipments) are reported. Instrument set defined in Appendix Table A1.

Table 3 - Likelihood of participating in the political game as a function of size

Dependent Variable		All Sectors		Distribution of coefficients across sectors				
				Mean	St. Dev.	Min	Max	No. of sectors
		I	II	III	IV	V	VI	VII
Contribution	Intercept	-0.014 (0.003)	-0.003 (0.001)	-0.013	0.033	-0.186	0.124	186
	log(Sales)	0.005 (0.001)	-0.005 (0.001)	0.004	0.008	-0.002	0.049	186
	log(Sales) squared		0.001 (0.0002)					
	No. of Firms	3027	3027					
	No. of Sectors	216	216					
	Estimator	OLS	OLS	OLS				
Probability of participating	log(Sales)	0.032 (0.003)		0.056	0.075	-0.090	0.447	189
	No. of Firms	3032						
	No. of Sectors	189						
	Estimator	Probit		Linear Probability				
Industry-level ratio of participating/non-participating firms	Intercept	-0.248 (0.031)		-0.267	0.423	-1.841	0.108	124
	Employment bin	0.058 (0.007)		0.058	0.079	-0.001	0.350	124
	No. of Sectors	210						
	Estimator	OLS		OLS				

Notes: COMPUSTAT sample. s.e. are clustered by industry in the full sample estimations (columns 1 and 2). Sales in million USD. The all sectors Probit model (column 1) reports the marginal effect of log(Sales) computed at the mean. Probit estimation may not be feasible by sector: in several sectors all firms above a certain threshold participate in the political game. The linear probability model allows to estimate the slope of log(Sales) in such cases as well. The number of SIC 4 sectors for which coefficients are estimated is limited by the number of observations per sector. The minimum number allowed in this Table is 4 (qualitatively similar results obtain with higher minima).

Table 4 - Correlation between θ_i, μ_i and σ_i

	μ_i	σ_i	θ_i
μ_i	-		
σ_i	0.6849 (0.00)	-	
θ_i	0.386 (0.00)	0.4738 (0.00)	-

Notes: p-values for pairwise correlation reported in parentheses. Number of obs.: 226. Variable θ_i is defined as the share of total output in sector i produced by firms making positive contributions; remanining variables defined in Notes of Table 2.

Table 5 - True Participation Shares

Dependent Variable NTB_i	
$\theta_i I_i(z_i/e_i)$	13.78
	(3.24)
z_i/e_i	-0.28
	(0.11)
Implied a/(1+a)	1.00
Implied α_L	0.02
Estimator	GMM
F-test joint significance $\theta_i I_i(z_i/e_i)$ (z_i/e_i) p-value	0.00
F-test model p-value	0.00
J-test overidentification p-value	0.18
Shea** Partl R^2/Partl. R^2	.92/.91
Shea** Partl R^2/Partl. R^2	.86/.86
Centered R^2	0.25

Notes: *γ_1 is the coefficient on $\theta_i I_i(z_i/e_i)$ and γ_2 is the coefficient on (z_i/e_i) **First stage Goodness of fit stats for $\theta_i I_i(z_i/e_i)$ and z_i/e_i. Two-step efficient GMM standard errors in parentheses below coefficients. Intercept included, not reported. All variables are defined in Notes of Table 2, except from θ_i, defined in Notes of Table 2 and Table 4. Instrument set defined in Appendix Table A1. All regressions include a constant and controls for intermediate goods tariffs and intermediate goods ntb's, not reported.

Table 6 - Constructed Participation Shares

Dependent Variable NTB_i	ρ =1/2	ρ =1/3	ρ =2/3	ρ =1/2	ρ =1/2
$\theta_i I_i(z_i/e_i)$	7.52	7.52	7.51	6.6	8.82
	(0.73)	(0.76)	(0.71)	(0.94)	(0.71)
z_i/e_i	-0.98	-0.88	-1.07	-0.82	-1.61
	(0.11)	(0.11)	(0.12)	(0.15)	(0.79)
Implied a/(1+a)	1.00	1.00	1.00	1.00	1.00
Implied α_L	0.13	0.12	0.14	0.12	0.18
Estimator	GMM	GMM	GMM	2SLS	Censored 2SLS
F-test joint $\theta_i I_i(z_i/e_i)$ (z_i/e_i)	0	0	0	0	0.02
F-test model	0	0	0	0	0
J-Test overidentification*	0.26	0.3	0.24	0.26	0.26
Shea** Partl R^2/Partl. R^2	.94/.89	.94/.89	.94/.88	.94/.88	.94/.88
Shea** Partl R^2/Partl. R^2	.91/.86	.91/.86	.92/.86	.92/.86	.92/.86
Centered R^2	0.26	0.26	0.26	0.27	0.38***

Notes: All specifications include an intercept. Two-step efficient GMM standard errors in parentheses below coefficients. * Hansen J-Test p-value reported. For instruments see GB and Melitz et al. **Goodness of fit stats for $\theta_i I_i(z_i/e_i)$ and z_i/e_i. ***Pseudo R^2 reported. Variables defined in notes to Table 2 and Table 4. Instrument set defined in Appendix Table A1. All regressions include a constant and controls for intermediate goods tariffs and intermediate goods ntb's, not reported.

Table 7 - Non-Nested Hypothesis Testing for True Theta

Null Hp.	Alternative Hp.	J-Test p-value	Interpretation	Test
GH	Heterogeneity	0.001	Reject null	I
Heterogeneity	GH	0.041	Reject null	II

Notes: Davidson and MacKinnon (1981) specification test for non nested models.
The null hypothesis tested is that the model associated to the null is the "correct model" and that the model under the alternative is uninformative.
Test I supports the Heterogeneity model. Also see Eicher and Osang (2002)

Table 8 - Non-Nested Hypothesis Testing for Contructed Theta

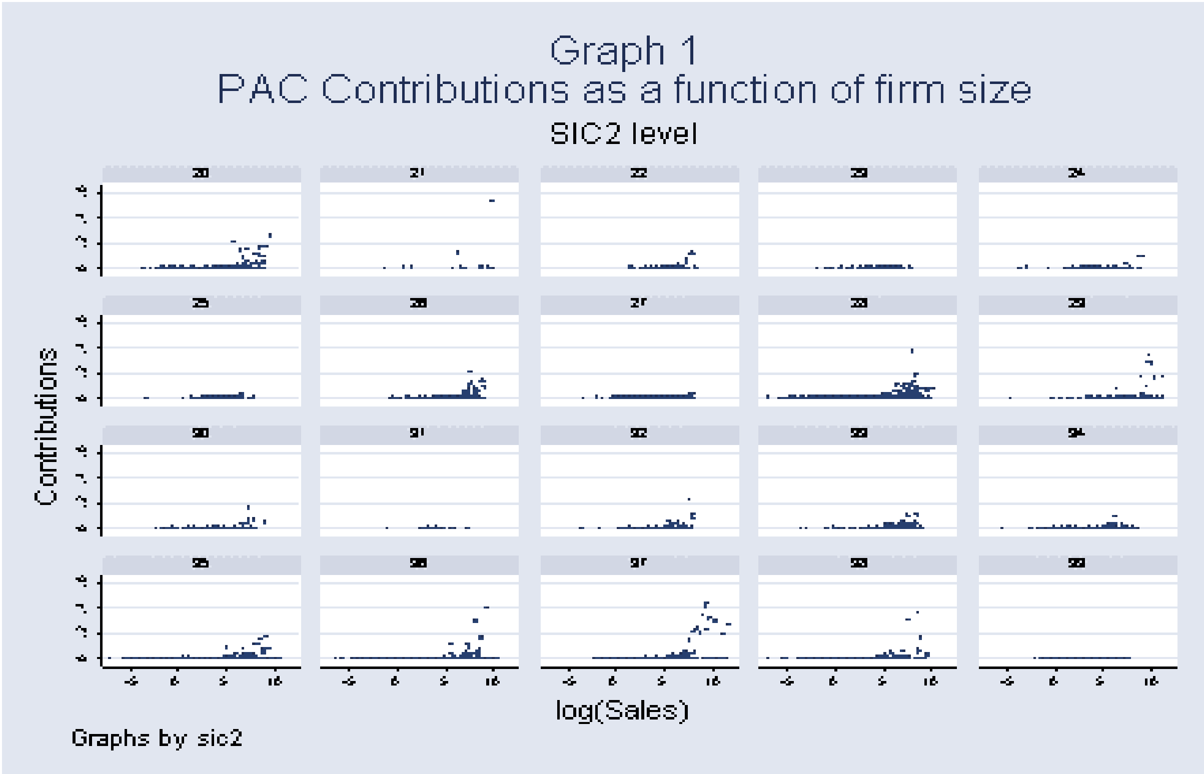
Null Hp.	Alternative Hp.	J-Test p-value	Interpretation	Test
GH	Heterogeneity	0	Reject null	I
Heterogeneity	GH	0.5758	Cannot reject null	II

Notes: Davidson and MacKinnon (1981) specification test for non nested models.
The null hypothesis tested is that the model associated to the null is the "correct model"
and that the model under the alternative is uninformative.
Both test I and II support the Heterogeneity model. Also see Eicher and Osang (2002)

Table A1 - Instruments list

Instruments	Description
1	Average tariff on intermediate goods used in an industry
2	Average coverage ratio on intermediate goods used in an industry
3	Logarithm of the price elasticity of imports (1986)
4	Log percentage of an industry's output used as intermediate good in other sectors
5	Logarithm of the intermediate goods buyer concentration
6	Herfindahl index of the industry
7	Measure of the scale of firms in an industry (value added per firm) (1982)
8	Concentration 4 (share of output in a sector produced by the four largest producers)
9	Share of industry employees defined as Unskilled (1982)
10	Share of industry employees defined as Scientists and Engineers (1982)
11	Share of industry employees defined as Managerial (1982)
12	Real exchange rate leasticity of imports and exports
13	Cross price elasticity between home production and imports (Shiells et al.)
14	Ad valorem tariff
15	Price elasticity of imports (1986)
16	Capital-labor ratio of the industry x Dummy for Food Processing Industry
17	Capital-labor ratio of the industry x Dummy for Resource-intensive Industry
18	Capital-labor ratio of the industry x Dummy for General Manufacturing Industry
19	Capital-labor ratio of the industry x Dummy for Capital Intensive Industry
20	Average log(sales) by industry from European data (Amadeus Dataset)*
21	Average log(sales) by industry from French data*
	Interactions between instrument 1-9 and 1-21

Notes: Instruments 1-19 are obtained from Gawande and Bandyopadhyay (2000). The set of instruments interactions was selected to optimize the fit of the first stage. *As reported in Helpman, Melitz, Yeaple (2003)


Graph 1
PAC Contributions as a function of firm size
SIC2 level
Contributions
log(Sales)
Graphs by sic2