A Quantile-based Test of Protection for Sale Model

Susumu Imai, Queen's University

Hajime Katayama, University of Sydney

Kala Krishna, Princeton University, Pennsylvania State University, and NBER

October 27, 2007

Abstract

This paper proposes a new test of the Protection for Sale (PFS) model by Grossman and Helpman (1994). Unlike existing methods in the literature, our approach does not require any data on political organizations. We formally show that the PFS model predicts that the quanitle regression of the protection measure on the inverse import penetration ratio divided by the import demand elasticity, should yield a positive coefficient for quantiles close to one. We test this prediction using the data from Gawande and Bandyopadhyay (2000). The results do not provide any evidence favoring the PFS model.

JEL Classification: F13,F14

Keywords: Protection for Sale; Lobbying; Political Economy; Quantile Regression

1 Introduction

Recently there has been much interest in political economy aspects of trade policy. This growing interest is in part triggered by the easy to use theoretical framework in the Grossman and Helpman (1994) "Protection for Sale" model (hereafter the PFS model). Empirical studies such as Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) have shown that as predicted by the PFS framework, protection is positively related to the import penetration ratio for politically unorganized industries, but negatively for politically organized ones.

One of the key explanatory variables in the estimating equation for the PFS model involves a dummy variable for whether the industry is politically organized. Therefore, an important issue in these empirical studies is how to classify industries into politically organized and unorganized ones. When classifying industries, past studies using US data have encountered the following problem: while only politically organized industries are assumed to make campaign contributions in the PFS model, the data indicate that all industries make Political Action Committees' (PAC) contributions. Thus, if one follows the assumptions in the PFS model that organized industries lobby while unorganized ones do not, all industries should be classified as politically organized. But in this case, the PFS model predicts the equilibrium level of protection should be the optimal tariff, which for a small country is zero.

To overcome this problem, past studies have used some simple rules for classification. Goldberg and Maggi (1999) classified an industry as politically organized if its PAC contribution is greater than a pre-specified threshold level. Gawande and Bandyopadhyay (2000) use a regression-based procedure. Their procedure is based on the idea that if industries are politically organized, then industries with higher import penetration ratios are likely to make higher campaign contributions.[1]

[1] More recently, a second generation of empirical studies has taken a different approach to reconciling theory and the data. For example, Ederington and Minier (2006) extend the PFS model by hypothesizing that industries

Several questions naturally arise about these classification rules. First, are their rules consistent with the PFS model? Second, do their rules correctly distinguish between politically organized and unorganized industries? And if there are classification errors, would that lead to bias in the parameter estimates of the PFS model?

In this paper, we argue against the classification rules in Goldberg and Maggi (1999) and in Gawande and Bandyopadhyay (2000). We formally derive the equilibrium relationship between campaign contributions and the inverse import penetration ratio. We then use the theoretical result to provide a simple numerical example of the PFS model where the level of the industry's contribution varies greatly depending on its import penetration. Specifically, politically organized industries may make very small contributions if their import penetration is high, i.e., inverse import penetration is low. This implies that using a particular threshold of campaign contribution as a device to distinguish between politically organized and unorganized industries as is done in Goldberg and Maggi (1999) results in mis-classification and is inconsistent with the PFS model. Furthermore, in our numerical example, import penetration and equilibrium campaign contributions are negatively correlated. This is exactly the opposite of the relationship that is assumed by Gawande and Bandyopadhyay (2000) and most papers using their data, that classify industries as politically organized when the import penetration and the PAC contributions per value added are positively correlated. We argue that if we were to reclassify the political organized industries, then their parameter estimates no longer support the PFS hypothesis.

We also argue that due to classification error, the estimation strategies used in Goldberg

can lobby for both trade and domestic policies. In their model, it is possible that some industries are politically unorganized for trade policies and yet make contributions for domestic policies. Matschke (2006) takes a similar approach. Since the models by Ederington and Minier (2006) and by Matschke (2006) are more comprehensive than the PFS model, the authors impose additional assumptions to make the models tractable for estimation.

and Maggi (1999) cannot provide consistent estimates. Estimation of the PFS model involves regressing a trade protection measure on the inverse import penetration ratio and its interaction term with the political organization dummy. The inverse import penetration ratio should be treated as an endogenous regressor, as has been discussed in the literature (e.g., Trefler, 1993). Potential mis-classification of industries makes it even more challenging to estimate the PFS model, since the political organization dummy would also be econometrically endogenous in the presence of classification error. As Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) were both fully aware of these problems, they used an IV strategy which, at a first glance, appears to provide consistent estimates. This paper shows that if the PFS model is true, the classification error in GM is a function of the inverse import penetration ratio, and consequently, so is the disturbance term in the estimating equation. It is therefore impossible to find an instrument that is correlated with the inverse import penetration ratio and uncorrelated with the disturbance term as needed.

In sum, we argue that if we are to structurally estimate the PFS model on the data used by Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000), we should not use an arbitrary classification scheme along with the campaign contributions to generate political organization dummies. The structural estimation and testing of the PFS model would require treatment of the political organization dummies to be fully consistent with the prediction of the PFS model. To our knowledge, this has not been done in the literature.

Given the shortcomings of the classification rules used in the literature, this paper proposes a new approach to testing the PFS model. Our approach heavily relies on the relationship between observables (i.e., the protection measure, import penetration, and import demand elasticity) implied by the PFS model and thus it is entirely consistent with the PFS framework. Moreover, since our estimating equation involves those observables only, it does not require classification of industries into organized and unorganized ones in any manner; our approach is therefore free from the risk of mis-classification unlike past studies. Furthermore, with this approach, the realm of testing the PFS model can be expanded, as it is applicable for many countries where contribution data are unavailable.

Our approach exploits the following prediction of the PFS model: politically organized industries should have higher protection than unorganized ones given the inverse import penetration ratio and other control variables. This suggests that industries with higher protection are more likely to be politically organized, and thus for those industries, we should expect a positive relationship between the inverse import penetration ratio and the protection measure. We provide a formal proof of this argument within the framework of recent works on quantile regression (Koenker and Bassett, 1978) and instrumental variable quantile regression (Chernozhukov and Hansen, 2004a; 2004b; 2006). We empirically test the prediction using the same data as Gawande and Bandyopadhyay (2000). We find that the estimated relationship is negative instead of positive, and insignificant, casting a serious doubt on the validity of the PFS model. We then discuss several possible explanations for the results.

The remainder of the paper is organized as follows. In Section 2, we review the PFS model and past empirical studies. Section 3 details our approach to testing PFS. Section 4 briefly describes the data used in this study. Section 5 presents the estimation results. In Section 6, we further discuss our results and also examine the validity of an alternative model. Section 7 concludes.

2 The PFS Model and Its Estimation in the Literature

2.1 The PFS Model

The exposition in this section relies heavily on Grossman and Helpman (1994). There is a continuum of individuals, each of infinitesimal size. Each individual has preferences that are linear in the consumption of the numeraire good and are additively separable across all goods. As a result, there are no income effects and no cross price effects in demand which comes from equating marginal utility to own price. On the production side, there is perfect competition in a specific factor setting: each good is produced by a factor specific to the industry, k_i in industry i, and a mobile factor, labor, L. Thus, each specific factor is the residual claimant in its industry. Some industries are organized, and being organized or not is exogenous to the model. Tariff revenue is redistributed to all agents in a lump sum manner. Owners of the specific factors in organized industries can make contributions to the government to try and influence policy if it is worth their while.

Government cares about both social welfare and contributions made to it and puts a relative weight of α on social welfare. The timing of the game is as follows: first, lobbies simultaneously bid contribution functions that specify the contributions made contingent on the trade policy adopted (which determines domestic prices). The government then chooses what to do to maximize its own objective function. In this way, the government is the common agent all principals (organized lobbies) are trying to influence. Such games are known to have a continuum of equilibria.[2] By restricting agents to bids that are "regret free" equilibrium bids have the same curvature as welfare, and a unique equilibrium can be obtained.[3] The equilibrium outcome in this unique equilibrium is as if the government was maximizing a weighted social welfare func-

[2] Given the bids of all other lobbies, each lobby wants a particular outcome to occur, namely, the one where it obtains the greatest benefit less cost. This can be attained by offering the minimal contribution needed for that outcome to be chosen by the government. However, what is offered for other outcomes (which is part of the bid function) is not fully pinned down as given other bids, it is irrelevant. However, bids at other outcomes affect the optimal choices of other lobbies and as their behavior affects yours, multiplicity arises naturally. Uniqueness is obtained by pinning down the bids at all outcomes to yield the same payoff as at the desired one, i.e., the bids are "regret free".

[3] For a detailed discussion of this concept, see Bernheim and Whinston (1986).

tion (where $W(p)$ is social welfare and p is the domestic price and equals the tariff vector plus the world price vector, p^*) with a greater weight on the welfare of organized industries. In other words, equilibrium tariffs can be found by maximizing

$$G(p) = \alpha W(p) + \sum_{j \epsilon J_0} W_j(p),$$

where J_0 is the set of politically organized industries. We provide a new elementary proof of this in Appendix 1 below.

In their model, the welfare of agents in industry j is

$$W_j(p) = \pi_j(p_j) + l_j + \frac{N_j}{N}\left[T(p) + S(p)\right],$$

where $\pi_j(p_j)$ is producer surplus in industry j, l_j is labor income of the owners of the specific factors employed in industry j, wage is unity, $\frac{N_j}{N} = \alpha_j$ is the fraction of agents who own the specific factor j, while $T(p) + S(P)$ is the sum of tariff revenue and consumer surplus in the economy.

Differentiating $W_i(p)$ with respect to p_j gives[4]

$$x_j(p_j)\delta_{ij} + \alpha_i\left[-x_j(p_j) + (p_j - p_j^*)m_j'(p_j)\right],$$

where so $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, α_i is the fraction of the population that owns the specific factor employed in industry i, $m_j'(p_j)$ is the derivative of the demand for imports, and $x_j(p_j) = \pi_j'(p_j)$ denotes supply of industry j. Differentiating $W(p) = \sum_i W_i(p)$ with respect to p_j gives

$$(p_j - p_j^*)m_j'(p_j).$$

Hence, maximizing $G(p)$ with respect to p_j gives

$$\alpha\left[(p_j - p_j^*)m_j'(p_j)\right] + \sum_{i \in J_0}\left[x_j(p_j)\delta_{ij} + \alpha_i\left[-x_j(p_j) + (p_j - p_j^*)m_j'(p_j)\right]\right] = 0.$$

[4] This follows from the derivative of consumer surplus from good j with respect to p_j being equal to $-d_j(p_j)$, where $d_j(p_j)$ is the demand for good j.

Now $\sum_{i \in J_0} \alpha_i = \alpha_L$, the fraction of the population that owns the specific capital of organized industries and $\sum_{i \in J_0} \delta_{ij} = I_j$ is unity if j is organized and zero otherwise. Therefore, this equation can be reduced to

$$x_j(p_j)(I_j - \alpha_L) + (p_j - p_j^*)m_j'(p_j)(\alpha + \alpha_L) = 0. \tag{1}$$

If we further use the fact that $(p_j - p_j^*) = t_j p_j^*$, it can be also expressed as

$$\frac{t_j}{1+t_j} = \left(\frac{I_j - \alpha_L}{\alpha + \alpha_L}\right)\left(\frac{z_j}{e_j}\right),$$

where $z_j = x_j(p_j)/m_j(p_j)$ and $e_j = -m_j'(p_j)p_j/m_j(p_j)$. This is the basis of the key estimating equation. Note that protection is predicted to be positively related to z_j/e_j if the industry is organized, but negatively related to it if the industry is not organized, and that the sum of the coefficients is predicted to be positive.

2.2 A Problem in Estimation — the Classification of Industries

To make the key equation estimable, an error term is added in a linear fashion:

$$\frac{t_j}{1+t_j} = \gamma \frac{z_j}{e_j} + \delta I_j \frac{z_j}{e_j} + \epsilon_j. \tag{2}$$

The error term is interpreted as the composite of variables potentially affecting protection that may have been left out and the measurement error of the dependent variable. To deal with the fact that a significant fraction of industries have zero protection in the data, equation (2) can be modified as follows:

$$\frac{t_j}{1+t_j} = Max\left\{\gamma \frac{z_j}{e_j} + \delta I_j \frac{z_j}{e_j} + \epsilon_j, 0\right\}. \tag{3}$$

The PFS model provides the following well-known predictions on the coefficients on z_i/e_i and $I_i z_i/e_i$: $\gamma < 0$, $\delta > 0$ and $\gamma + \delta > 0$.[5] To test these predictions, equations (2) and (3) (hereafter

[5] Goldberg and Maggi (2000) and others note that $\gamma < 0$, $\delta > 0$ and $\gamma + \delta > 0$ are only necessary conditions for the validity of the PFS specification. However most empirical research in the political economy of trade claim

called the PFS equations) have been estimated in a number of previous studies (e.g., Goldberg and Maggi, 1999; Gawande and Bandyopadhyay, 2000; McCalman, 2001).

Although data on the measure of trade protection, the import penetration ratio, and the import-demand elasticities are often available, it is harder to define whether an industry is politically organized or not. To deal with this problem, Goldberg and Maggi (1999) used data on campaign contributions at the three-digit SIC industry level. An industry is categorized to be politically organized if the campaign contribution exceeds a specified threshold level. Gawande and Bandyopadhyay (2000) used a different procedure for classification. They run a regression where the dependent variable is the log of the corporate Political Action Committee (PAC) spending per contributing firm relative to value added and the regressors include the interaction of the import penetration from five countries into the sub industry and the two-digit SIC dummies. Industries are classified as politically organized if any of the coefficients on its five interaction terms are found to be positive. This procedure is based on the idea that in organized industries, an increase in contributions would likely occur when import penetration increased.

Note that both these procedures are questionable. The arguments against the procedure used in Goldberg and Maggi (1999) (hereafter often referred to as GM) are that it implicitly assumes that all the contributions are directed towards influencing trade policies. Moreover, taking any nonzero cutoff level of contributions as indicating organization seems relatively arbitrary. In addition, the procedure does not control for other variables that potentially influence political clout such as industry size and electoral districts where the industry is concentrated. Gawande and Bandyopadhyay (2000) (hereafter often referred to as GB) also do not provide a

that the right sign of the coefficients of the PFS equation gives strong empirical support of the PFS paradigm. Recently, Imai et al. (2006) criticize them by pointing out that even when estimating the PFS equation on an artificial data simulated from a simple non optimizing model without a PFS element, one obtains parameter estimates consistent with the PFS model. This suggests that to truly test PFS, other implications of the model need to be considered.

rigorous rationale for their approach.

Below we offer a formal argument that claims: (1) both of the approaches are inconsistent with the PFS model and (2) both of the approaches result in the mis-classification of industries, which leads to inconsistent parameter estimates. Given the model and the menu auction equilibrium of the PFS model, the following equation which describes the relationship between the campaign contribution and other variables can be derived. (See Appendix 1 for more details.)

$$\begin{aligned} B_i^*(\mathbf{p}^E) + \left[\alpha W(\mathbf{p}^E) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}^E)\right] &= \underset{\mathbf{p}}{Max}\left[\alpha W(\mathbf{p}) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p})\right] \\ &= \alpha W(\mathbf{p}(\mathbf{i})) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}(\mathbf{i})) \end{aligned}$$

where $B_i^*(\mathbf{p}^E)$ is the campaign contribution of industry i at the equilibrium domestic price vector $\mathbf{p}^E$. That is, the equilibrium campaign contribution schedule should be such that government welfare at equilibrium should equal the maximized value of the government objective function when industry i is not making any contributions at all. Thus, the equilibrium campaign contribution can be expressed as follows.

$$B_i^*(\mathbf{p}^E) = -\left[\alpha W(\mathbf{p}^E) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}^E)\right] + \alpha W(\mathbf{p}(\mathbf{i})) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p}(\mathbf{i})). \tag{4}$$

Note that this says that equilibrium contributions are essentially the difference in the value of the function $H(\mathbf{p}) : R^N \rightarrow R$, (where $H(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p})$) from $\mathbf{p}(\mathbf{i})$ to $\mathbf{p}^E$. Let $\mathbf{p}(t)$ be a path from $\mathbf{p}(\mathbf{i})$ to $\mathbf{p}^E$ as t goes from zero to unity. Since the line integral is path independent, we can choose this path as desired. In particular, we can choose it so that $\mathbf{p}(t) = \mathbf{p}(\mathbf{i}) + t\left[\mathbf{p}^E - \mathbf{p}(\mathbf{i})\right]$ so that $D\mathbf{p}(t) = \left[\mathbf{p}^E - \mathbf{p}(\mathbf{i})\right]$

Hence,

$$\begin{aligned} H(\mathbf{p}(\mathbf{i})) - H(\mathbf{p}(\mathbf{i})) &= \int_0^1 H'(\mathbf{p}(t))dt \qquad (5) \\ &= \int_0^1 DH(\mathbf{p}(t)) \bullet D\mathbf{p}(t)dt \end{aligned}$$

where $DH(\mathbf{p}(t))$ is the vector of partial derivatives of the real valued function $H(.)$ with respect to the vector $\mathbf{p}$ and $D\mathbf{p}(t)$ is vector of the derivatives of $\mathbf{p}$ with respect to t and $\bullet$ denotes their dot product.

Next, we derive the vector $\mathbf{p}(\mathbf{i})$, which maximizes the second term of the right hand side of equation (4). Notice that

$$\sum_{j\epsilon J_0, j\neq i} \alpha_j = \alpha_L - \alpha_i,$$

so that differentiating $\alpha W(\mathbf{p}) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p})$ gives the first order condition with respect to $p_l,\ l \notin J_0 - \{i\}$, (which includes $l = i$) to be

$$-(\alpha_L - \alpha_i)\, x_l\,(p_l(i)) + (\alpha + \alpha_L - \alpha_i)\,(p_l(i) - p_l^*)\, m_i'\,(p_l(i)) = 0$$

or

$$\frac{p_l(i) - p_l^*}{p_l(i)} = -\frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i}\frac{z_l}{e_l}, \quad p_l(i) = \frac{p_l^*}{1 + \frac{\alpha_L - \alpha_i}{\alpha + \alpha_L - \alpha_i}\frac{z_l}{e_l}}.$$

Similarly for any $j \in J_0 - \{i\}$, we get

$$\frac{p_j(i) - p_j^*}{p_j(i)} = \frac{1 - (\alpha_L - \alpha_i)}{\alpha + \alpha_L - \alpha_i}\frac{z_j}{e_j}, \quad p_j(i) = \frac{p_j^*}{1 - \frac{1 - (\alpha_L - \alpha_i)}{\alpha + \alpha_L - \alpha_i}\frac{z_j}{ej}}.$$

Of course, the vector $\mathbf{p}\,(\mathbf{i}) \equiv (p_1(i), ..., p_N(i))$. Note the analogy with equation (1). The above allows us to find $\mathbf{p}\,(\mathbf{i})$ from the data.

Now using the line integral defined in equation (5) and substituting for $DH(\mathbf{p}(t)) = \left[\frac{\partial H(\mathbf{p})}{\partial p_j}\right] = \left[x_j(p_j)(I_j - \alpha_L) + (p_j - p_j^*)m_j'(p_j)(\alpha + \alpha_L)\right]$ and $D\mathbf{p}(t) = \left[\mathbf{p}^E - \mathbf{p}(\mathbf{i})\right]$ and remembering that in the PFS setup, consumption, production and imports depend only on own price we get

$$\begin{aligned} B_i^*(\mathbf{p}^E) &= \int_0^1 \sum_j \{(\alpha + \alpha_L - \alpha_i)\,(p_j(t) - p_j^*)\frac{\partial m_j\,(p_j(t))}{\partial p_j} \\ &\quad + [I\,(j \in L - \{i\}) - (\alpha_L - \alpha_i)]\, x_j\,(p_j(t))\}\{p_j^E - p_j(i)\}dt \\ &= \sum_j \int_0^1 \{(\alpha + \alpha_L - \alpha_i)\,(p_j(t) - p_j^*)(z_j\,(p_j\,(t))\,/e_j)^{-1} \\ &\quad + [I\,(j \in L - \{i\}) - (\alpha_L - \alpha_i)]\}x_j\,(p_j\,(t))\,dt. \end{aligned}$$

Thus, depending on α_i,α, α_L, $x(.)$, and z_i/e_i, $B_i^*(\mathbf{p}^E)$ can be small even for politically organized industries[6]. Hence, classifying political organization based on a uniform threshold, as one in GM (1999) and others leads to classification error. Furthermore, notice that the classification error is

$$\eta_j = I\left(C_j\left(\widehat{\mathbf{p}}\right) > \overline{C}\right) - I_j,$$

where I_j is the true political organization dummy and $\overline{C}$ is the threshold. Then, it is clear that the classification error is a function of z_j/e_j, thus any variable correlated with the inverse import penetration ratio cannot be used as an instrument for political organization, which makes the instrumenting of the term $I_i z_i/e_i$ impossible.

Below, we provide a simple example, where we assume there are 400 industries ($N = 400$), of which 200 are politically organized ($N_p = 200$). We set $p_i^* = 2.0$, $\alpha = 50.0$, $\alpha_L = 0.508$, $\alpha_i = \alpha_L/N$. We then set $z_i = i/1000$ for industries $i = 1, ...N_p$ which are politically organized and $z_{N_p+i} = i/1000$ for industries $N_p + i = N_p + 1, ...N$ which are not politically organized. We also set $x_i = 1000000$. We also set the elasticity of import e_i to be 1 for every industry. In Figure 1, we present the equilibrium campaign contributions. Notice that the campaign contribution increases with z, the inverse import penetration ratio for the political organized industries. That is, for politically organized industries the campaign contributions are negatively correlated with the import penetration, which is the *opposite* of the relationship used by GB to classify political organization.

The positive relationship between campaign contributions and $\frac{z}{e}$ in the simulated model is in line with PFS. PFS predicts that for politically organized industries, protection is positively related to $\frac{z}{e}$. Hence, campaign contributions and $\frac{z}{e}$ are likely to be positively related as long as greater campaign contributions result in higher protection.

This has important implications for the interpretation of the results obtained by GB. That

[6] Below we present a simple example of such a case.

is, their parameter estimates are: $\gamma_{GB} < 0$, $\delta_{GB} > 0$, and $\gamma_{GB} + \delta_{GB} > 0$. But if the correct political organized industries should be the ones that they classified as politically unorganized and vice versa, then the correct classification of political organization should be $1 - I_{GB}$ where I_{GB} is the politically organization dummy by GB. Hence, the protection equation should be

$$\frac{t_j}{1+t_j} = (\gamma_{GB} + \delta_{GB}) \frac{z_j}{e_j} - \delta_{GB} (1 - I_{GB,j}) \frac{z_j}{e_j} + \varepsilon_j.$$

Then, $\widehat{\gamma} = \gamma_{GB} + \delta_{GB} > 0$, $\widehat{\delta} = -\delta_{GB} < 0$ and $\widehat{\gamma} + \widehat{\delta} = \gamma_{GB} < 0$ is *inconsistent* with the PFS framework. Furthermore, we also observe that campaign contributions vary from 3 to 18, depending on the value of z, highlighting the possibility that the GM classification based on the threshold of campaign contribution may mis-classify industries with low campaign contribution and low z as politically unorganized.

We have argued above both in the numerical example and the brief discussion that the relationship between z/e and the campaign contributions predicted by the theory is likely to be positive, not negative. We now check the relationship in the data. Figure 2 shows the scatterplot where the x-axis is $log(z/e)$ and y-axis is the log of per value added campaign contributions. This data is that used in GB (1999). Figure 3 depicts the scatterplot where the x-axis is $log(z/e)$ and the y-axis the log of campaign contributions, using the data by Facchini .et. al. (2007) who reconstructed the Goldberg and Maggi (2000) dataset.[7] We used logs to minimize the effect of outliers. In both figures, we can see that the relationship between the two is negative, not positive as predicted by the PFS framework. It is needless to say that these results by no means statistically reject the PFS framework. A more rigorous estimation and testing exercise of the PFS model using campaign contribution data is left for future research.

[7] This had to be done as the data of GM has never been made available to other researchers.

3 A Proposed Approach

3.1 Quantile Regression

In this section, we detail our approach to testing the PFS model. The advantage of our approach is that it allows us to test the PFS model without an arbitrary classification of the political organization. The approach relies heavily on the relationship between observables implied by the PFS model.

Equation (3) and the restrictions on the coefficients have at least two implications. First, as has been discussed in the literature, z/e has a negative effect on the level of protection for politically unorganized industries while it has a positive effect for politically organized ones. Second, given z/e, politically organized industries have higher protection. These implications lead to the following claim: given z/e, high-protection industries are more likely to be politically organized and thus the effect of an the increase in z/e on protection tends to be that of politically organized industries.

The logic of this argument is illustrated in Figure 4 where the distribution of $t/(1+t)$ is plotted for given z/e. The variation of $t/(1+t)$ given z/e occurs for two reasons. First, because some industries are organized while others are not and these two behave differently, and second, because of the error term. This results the distribution of $t/(1+t)$ comes from a mixture of two distributions, namely those for the politically organized industries and the unorganized ones. These two distributions for some given values of z/e are plotted in Figure 4. The two dashed lines give the conditional expectations of $t/(1+t)$ for the organized and unorganized industries as a function of z/e. In line with the PFS model, the two lines start at the same vertical intercept point and the line for the organized industries is increasing while the other is decreasing in z/e. For each z/e, if we look at the industries with high $t/(1+t)$, they tend to be the politically organized ones. Thus, at high quantiles, the relationship between $t/(1+t)$ and z/e should be that for organized industries, i.e., should be increasing as depicted by the solid line labelled the 90th quantile in Figure 4.

The relevant proposition (Proposition 1) and proof can be found in Appendix 2. The proposition essentially states that in the quantile regression of $t/(1+t)$ on z/e, the coefficient on z/e should be close to $\gamma+\delta>0$ at the quantiles close to $\tau=1$. To empirically examine this, we use quantile regression (Koenker and Bassett, 1978) and estimate the following equation:

$$Q_T(\tau|Z)=\alpha(\tau)+\beta(\tau)Z/10000, \tag{6}$$

where τ denotes quantile, $T=t/(1+t)$, $Z=z/e$, and $Q_T(\tau|Z)$ is the conditional τ-th quantile function of T. If the PFS model is correct, it is expected that $\beta(\tau)$ converges to $(\gamma+\delta)>0$ as τ approaches its highest level of unity from below.

In the quantile regression, Z is assumed to be an exogenous variable. However, Z is likely to be endogenous as discussed in the literature and hence the parameter estimates of the quantile regression are likely to be inconsistent. It is therefore important to allow for the potential endogeneity of Z. We formally show that even in the presence of this endogeneity, the main prediction of the PFS model in terms of our quantile approach does not change. The relevant proposition (proposition 2), an analogue of proposition 1, is presented in Appendix 2. To test the prediction in the presence of possible endogeneity of Z, we estimate the following equation by using IV quantile regression (Chernozhukov and Hansen, 2004a; 2004b; 2006):

$$P(T\leq\alpha(\tau)+\beta(\tau)Z/10000|W)=\tau, \tag{7}$$

where W is a set of instrumental variables.

Importantly, nowhere in equations (6) and (7) is the political organization dummy present; these equations involve only variables that are readily available. This way our approach does not require classification of industries in any manner and as a result, we can avoid biased results due to mis-classification.

An issue that we need to deal with is the endogeneity of the political organization. We do so first by controlling for capital-labor ratio. This is essentially equivalent to allowing capital-labor ratio to be a determining factor for the probability of political organization. This specification is motivated by Mitra (1999) who provides a theory of endogenous lobby formation. The model predicts that among others, industries with higher levels of capital stock are likely to be politically organized.

But even after controlling for capital-labor ratio, there still could remain a correlation between the error term of the equation determining the political organization and the error term of the protection equation (3). In this paper, we are less concerned about this correlation for the following reasons. First, our method is not subject to classification error, one of the main sources of correlation between the error terms in the two equations in GM and other studies. In these studies, the classification error is captured by both the disturbance term of the equation determining the political organization and the disturbance term of the protection equation. Thus, classification error would result in positive correlation between the disturbance terms. In our method, however, this source of correlation is not present since the method does not involve classification and hence classification error. Second, as long as the error term of the equation determining political organization and that of the protection equation are positively correlated, or as long as the negative correlation is not too strong, then our quantile IV procedure will still be consistent. This is because the political organization dummies do not enter in the RHS of the estimating equation, and after controlling for Z we still would see most of the industries in high quantiles (i.e. industries whose error term of the protection equation are high) to be politically organized. The only case where the IV quantile regression results for high quantiles gives a biased estimate of $\gamma + \delta$ is when, given Z_j, the politically organized industries have equal or less protection than the unorganized ones, which we believe to be an unlikely scenario.

4 A Brief Description of the Data

We use part of the data used in Gawande and Bandyopadhyay (2000).[8] The data consist of 242 four-digit SIC industries in the United States. In the dataset, the extent of protection t is measured by the nontariff barrier (NTB) coverage ratio. This is a standard exercise in the literature (e.g., Goldberg and Maggi, 1999; Mitra et al., 2002). z is measured as the inverse of the ratio of consumption to total imports scaled by $10,000$. e is derived from Shiels et al. (1986) and corrected for measurement error by GB. A brief description of the variables used in the current study is provided in Table 1. See GB for more details along with the sample statistics of the variables. Of particular note about the data is that 114 of 242 industries (47%) have zero protection. This suggests the potential importance of dealing with the corner solution outcome of T.

5 Estimation Results

5.1 Quantile Regression Results

Column (1) of Table 2 presents the estimation results of equation (6).[9] The results do not appear to provide any supporting evidence for the PFS model; the null hypothesis that $\beta(\tau) = 0$ cannot be rejected at high quantiles (in fact, at all quantiles) in favor of the one-sided alternative that $\beta(\tau) > 0$. Moreover, the point estimates indicate that contrary to the PFS prediction, the $\beta(\tau)$ are all negative at high quantiles and decrease as τ goes from 0.4 to 0.9.

Note that α and β are estimated to be zero at the 0.1-0.4 quantiles. This suggests that the

[8] We are grateful to Kishore Gawande for kindly providing us with the data.

[9] All the estimation in this study is done by using a MATLAB code written by Christian Hansen (available at http://faculty.chicagogsb.edu/christian.hansen/reserach).

corner solution ($T = 0$) greatly affects the estimates at the lower quantiles. From this evidence, it is conjectured that the existence of corners also affects the estimates at the mean. Thus, findings based on the linear model (i.e., equation (2)) in GB (2000), Bombardini (2005), and others are likely to be subject to bias due to the corner solution problem. In contrast, our method does not suffer from the problem, since the focus is mainly on the higher quantiles where the effect of corner solution is minimal. In fact, our method has a distinct advantage over the other estimation strategy in the literature. To address the corner solution problem, several studies (e.g., Goldberg and Maggi, 1999; Facchini et al., 2006) estimate a system of equations: equation (3) as well as an import penetration equation and an equation for political organization. While dealing with the existence of corners, this strategy requires the joint normality assumption on the error terms which potentially affects the estimation results. In contrast, our results are not driven by the parametric assumption on the error term; it is not required by the quantile regression.

One might wish to control for various factors as well. Following Gawande and Bandyopadhyay (2000), we control for tariff of intermediate goods ($INTERMTAR$) and NTB coverage of intermediate goods ($INTERMNTB$). As column (2) of Table 2 shows, our main findings do not change; $\beta(\tau)$ decreases (for the most part) from zero to a negative value with the increase in τ, contrary to what the PFS model predicts. α and β are found to be zero at the 0.1 and 0.2 quantiles, again suggesting the importance of corner solution.

5.2 IV Quantile Regression Results

Table 3 presents the estimation results of equation (7). Our choice of instruments is guided by GB (2000) where they used 34 distinct instruments, their quadratic terms, and some of the two-term cross products. We use a subset of their instruments (17 instruments) indicated in

Table 1. These are also used in Bombardini (2005) as the basic instruments.[10] First, we use two sets of instruments. Instrument set 1 consists of the 17 instruments, their squared terms and the squares of $INTERMTAR$ and $INTERMNTB$. Instrument set 2 includes instrument set 1 and the interaction terms of the 17 instruments. The IV quantile results for the instrument set 1 are reported in columns (1) of Table 3. As in the quantile regression, we cannot reject the null hypothesis that $\beta(0.9) = 0$ in favor of the one-sided alternative. The point estimates are not favorable for the PFS model, either; even after correcting for the endogeneity of Z, the estimate of β at the highest quantile is not positive as required by the PFS model. The results remain virtually the same when we use the instrument set 2 as IV's, as columns (2) indicates.

The estimation results where the capital labor ratio is controlled for are presented in columns (3) and (4) of Table 3. $\beta(\tau)'s$ are again estimated to be zero at $\tau = 0.1$ regardless of whether we use instrument set 1 or 2. At high quantiles, $\beta(\tau)'s$ are estimated to be negative for most of τ's . Although the point estimate of $\beta(0.9)$ is positive when we use instrument set 2, the null hypothesis cannot be rejected in favor of the one-sided alternative.

Although we use a subset of GB's instruments, our results may be driven by too many instruments. Thus, we further estimate equation (7) using only one of the following instruments at a time: $SCIENTISTS$, $MANAGERS$, and $CROSSELI$ and using all of them (see Table 1 for their definitions). These instruments are found to be strongly correlated with Z in GB (2000). The results are presented in columns (5) - (12) of Table 3. The results suggest that having many instruments affects the estimates of $\beta(\tau)$. Specifically, the absolute magnitude of the coefficients now become far larger than that obtained with the larger number of instruments. Nonetheless, our main findings appear to be robust; regardless of which instrument we use and whether we control for capital-labor ratio, the null hypothesis at the highest quantile cannot be rejected. Moreover, the point estimates of $\beta(\tau)$ are negative at high quantiles, in fact, they are

[10] We are grateful to Matilde Bombardini for providing us with the program for her PFS estimation.

zero at low quantiles and negative at any other quantiles, which is inconsistent with the PFS's prediction.

6 Discussion

There are several possibilities that may explain our results. The first possibility is heteroskedasticity. If the error term has higher variance when the industry is politically unorganized, i.e.,

$$\varepsilon_j = w_j + (1 - I_j)\,\zeta_j, \tag{8}$$

then politically unorganized industries would have error terms with much higher variance.[11] As a result, they would be the ones that dominate in high quantiles as well as in low quantiles, whereas the politically organized industries would be found mostly around the median. Hence, at high quantiles, the negative quantile regression coefficients correspond to γ, which is negative, and not $\gamma + \delta > 0$. This may explain the presence of negative slope coefficients in the higher quantiles. The possibility cannot be completely ruled out. However, given that almost all industries have positive campaign contributions and both GM and GB report that more than half of the industries are politically organized, it is reasonable to think that a significant fraction

[11] If equation (8) is indeed the error structure, then the PFS equation is modified to be:

$$\frac{t_j}{1+t_j} = \gamma\frac{z_j}{e_j} + \delta\frac{z_j}{e_j}I_j + \varsigma_j\,(1 - I_j) + w_j.$$

Importantly, the modified equation has an additional term $1 - I_j$ with a random coefficient ς_j. That is, we are needing an error term with richer stochastic structure to make the model consistent to the data. However, the more we rely on the complexity of the stochastic structure of the error term instead of the model part to fit the data, the less attractive becomes the treatment of the error term as an "add on" to the structural model. And if we decide not to arbitrarily add an error term to the reduced form of the deterministic model, the original lobbying model needs to be substantially modified to explicitly include stochastic shocks so that the reduced form of the stochastic model results to the modified equation above. Then, it would be unclear whether findings in past studies (i.e., $\gamma < 0$, $\delta > 0$, and $\gamma + \delta > 0$) can be interpreted as being in support of the PFS paradigm.

of the industries are likely to be politically organized. In that case, it is surprising to find that the slope coefficients of the quantile regressions are negative at almost all quantiles except for the zeros at low quantiles, which comes from the corners.

Second, the small sample may make it difficult for our approach to provide evidence favoring the PFS model. This problem can be overcome by using more disaggregated data, although such an exercise is beyond the scope of the current paper.

Third, one can argue that there is a possibility that if the political organization were correctly assigned, then our results are not inconsistent with those of GB. Recall that in our simple example where we computed the relationship between the equilibrium campaign contribution and z/e for politically organized industries, it was positive instead of negative. If the positive relationship holds in reality, we argued that the industries that were originally classified as politically organized should be classified as unorganized and vice versa, and then the true results of the GB protection equation estimation should be $\widehat{\gamma} > 0$, $\widehat{\delta} < 0$ and $\widehat{\gamma} + \widehat{\delta} < 0$, which is indeed consistent with our quantile regression and quantile IV results.

7 Conclusion

In this paper, we proposed a new test of the PFS model that does not require data on political organizations. The test is based on a certain prediction of the PFS model which has not been explored in the literature: given that industries with higher protection measures are more likely to be politically organized, the effect of the inverse import penetration ratio on protection at higher quantiles tends to reflect that of politically organized ones. We tested this prediction using the quantile regression and IV quantile regression techniques. The findings are not supportive of the PFS model, unlike those in past studies in the literature. Clearly, more evidence is needed to conclude the empirical validity of the PFS model. One fruitful research avenue is to analyze different countries than United States. Such an exercise can be done relatively easily, as our method does not require data on political organization. Another research avenue is to use more disaggregated data so that our approach can provide statistically more clear-cut evidence.

References

Bombardini, Matilde (2005) "Firm Heterogeneity and Lobby Participation," mimeo, University of British Columbia.

Chernozhukov, Victor and Christian Hansen (2004*a*) "Instrumental Variable Quantile Regression," MIT Department of Economics Working Paper No. 06-19.

Chernozhukov, Victor and Christian Hansen (2004*b*) "The Effects of 401(k) Participation on the Wealth Distribution: An Instrumental Quantile Regression Analysis," *Review of Economics and Statistics*, 86(3), 735-751.

Chernozhukov, Victor and Christian Hansen (2006) "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics*, 132(2), 491-525.

Ederington, Josh and Jenny Minier (2005) "Reconsidering the Empirical Evidence on the Grossman-Helpman Model of Endogenous Protection," University of Kentucky Working Paper.

Facchini, Giovanni, Johannes V. Biesebroeck, and Gerald Willmann (2006) "Protection for Sale with Imperfect Rent Capturing," *Canadian Journal of Economics*, 39(3), 845-873.

Gawande, Kishore and Usree Bandyopadhyay (2000) "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection," *Review of Economics and Statistics*, 82(1), 139-152.

Goldberg, Pinelopi K. and Giovanni Maggi (1999) "Protection for Sale: An Empirical Investigation," *American Economic Review*, 89(5), 1135-1155.

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

Imai, Susumu, Hajime Katayama, and Kala Krishna (2006) "Protection for Sale or Surge Protection?" NBER Working Paper No. 12258.

Koenker, Roger W., and Gilbert W. Bassett (1978) "Regression Quantiles," *Econometrica*, 46(1), 22-50.

Matschke, Xenia (2006) "Costly Revenue-Raising and the Case for Favoring Import-Competing Industries," University of Connecticut Working Paper No. 2005-28R.

McCalman, Phillip (2004) "Protection for Sale and Trade Liberalization: An Empirical Investigation," *Review of International Economics*, 12(1), 81-94.

Mitra, Devashish (1999) "Endogenous Lobby Formation and Endogenous Protection: A Long-run Model of Trade Policy Determination," *American Economic Review*, 89(5), 1116-1134.

Mitra, Devashish, Dimitrios D. Thomakos, and Mehmet A. Ulubasoglu (2002) "'Protection for Sale' in a Developing Country: Democracy vs. Dictatorship," *Review of Economics and Statistics*, 84(3), 497-508.

Shiells, Clinton R., Robert M. Stern, and Alan V. Deardorff (1985) "Estimates of the Elasticities of Substitution between Imports and Home Goods for the United States," *Weltwirtschaftliches Archiv,* 122(3), 497-519.

Trefler, Daniel (1993) "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*, 101(1), 138-160.

Table 1: Definitions of Variables

Variable	IV	Definition
t		All NTB coverage ratios at the 4-digit SITC level
z		(Consumption in 1983)/(Total imports)
$INTERMTAR$		Average tariff on intermediate good use
$INTERMNTB$		Average NTB coverage of intermediate good use
e	1	Absolute import elasticity after correcting for measurement errors
$ln(e)$	2	Log of absolute import elasticity after correcting for measurement errors
$ln(HERF)$	3	Log of Hefindahl index of firm concentration
$ln(DOWNSTREAMSHR)$	4	Log of percentage of an industry's shipment used as intermediate goods in others
$ln(DOWNSTREAMHERF)$	5	LOg of intermediate-goods-output-buyer concentration
$SCIENTISTS$	6	Fraction of employees classified as scientists and engineers, 1982
$MANAGERS$	7	Fraction of employees classified as managerial, 1982
$UNSKILLED$	8	Fraction of employees classified as unskilled, 1982
$CONC4$	9	Four-firm concentration ratio, 1982
$FIRMSCALE$	10	Measure of industry scale: Value added per firm, 1982
TAR	11	US post-Tokyo round ad valorem tariffs (Ratio)
$RERMELAST$	12	Real exchange rate elasticity of imports
$CROSSELI$	13	Cross price elasticity of imports
$(K/L)_1$	14	Capital-labor ratio, food processing
$(K/L)_2$	15	Capital-labor ratio, resourse intensive
$(K/L)_3$	16	Capital-labor ratio, general manufacturing
$(K/L)_4$	17	Capital-labor ratio, capital intensive

Table 2: Quantile Regression Results

	(1)		(2)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.004)	0.000 (0.056)	0.000 (0.013)	0.000 (0.060)
0.2	0.000 (0.005)	0.000 (0.079)	0.000 (0.017)	0.000 (0.080)
0.3	0.000 (0.006)	0.000 (0.091)	−0.026 (0.014)	−0.099 (0.153)
0.4	0.000 (0.006)	0.000 (0.097)	−0.029 (0.014)	−0.020 (0.092)
0.5	0.002 (0.006)	−0.003 (0.099)	−0.026 (0.014)	−0.032 (0.094)
0.6	0.028 (0.006)	−0.046 (0.098)	−0.053 (0.024)	−0.082 (0.093)
0.7	0.077 (0.010)	−0.126 (0.095)	−0.044 (0.017)	−0.125 (0.090)
0.8	0.157 (0.026)	−0.258 (0.094)	−0.046 (0.018)	−0.145 (0.086)
0.9	0.308 (0.040)	−0.505 (0.089)	−0.001 (0.021)	−0.225 (0.075)
GB Controls	No		Yes	

Note: This table provides the estimation results of equation (3). Standard errors are in parentheses. GB Controls indicate whether $INTERMTAR$ and $INTERMNTB$ are controlled for. For the definition of these vriables, see Table 1.

Table 3: IV Quantile Regression Results

	(1)		(2)		(3)		(4)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.003)	0.000 (0.407)	0.000 (0.002)	0.000 (0.270)	0.000 (0.013)	0.000 (0.592)	0.000 (0.013)	0.000 (0.582)
0.2	0.000 (0.012)	0.000 (0.402)	0.000 (0.011)	0.000 (0.369)	−0.037 (0.018)	−0.050 (0.496)	−0.036 (0.017)	−0.110 (0.944)
0.3	−0.025 (0.011)	−0.370 (0.357)	−0.026 (0.011)	−0.370 (0.287)	−0.037 (0.012)	−0.050 (0.391)	−0.036 (0.010)	−0.190 (0.338)
0.4	−0.028 (0.009)	−0.200 (0.621)	−0.029 (0.009)	−0.200 (0.421)	−0.060 (0.017)	−0.020 (0.060)	−0.043 (0.017)	−0.140 (0.359)
0.5	−0.031 (0.023)	−0.270 (1.395)	−0.026 (0.023)	−0.270 (1.091)	−0.043 (0.030)	−0.250 (1.241)	−0.060 (0.033)	−0.250 (0.828)
0.6	−0.053 (0.023)	−0.080 (2.153)	−0.053 (0.024)	−0.080 (1.184)	−0.100 (0.033)	−0.540 (2.085)	−0.103 (0.035)	−0.270 (1.602)
0.7	−0.044 (0.015)	−0.130 (2.403)	−0.044 (0.014)	−0.130 (1.611)	−0.080 (0.028)	−0.120 (2.388)	−0.080 (0.031)	−0.120 (2.031)
0.8	−0.046 (0.016)	0.020 (2.722)	−0.046 (0.014)	0.020 (1.826)	−0.059 (0.022)	−0.160 (2.794)	−0.059 (0.024)	−0.160 (2.387)
0.9	−0.002 (0.044)	−0.230 (3.572)	−0.001 (0.042)	−0.230 (3.383)	−0.037 (0.054)	−0.250 (3.077)	−0.070 (0.064)	4.190 (3.444)
GB Controls	Yes		Yes		Yes		Yes	
K/L	No		No		Yes		Yes	
Instruments	Set 1		Set 2		Set 1		Set 2	

Note: This table provides the estimation results of equation (4). Standard errors are in parenthese. They are calculated by 200 bootstrap resampling. GB Controls and K/L indicate whether $INTERMTAR$ and $INTERMNTB$ are controled for and whether $(K/L)_i$, (i =1,2,3,4) are controled for, respectively. Instruments indicates which variables are used as instrumental variables. Set 1 include IV1-17, their quadratic terms, and the quadratic terms of GB controls. Set 2 include Set 1 plus the interaction terms involving IV1. For the definition of these variables, see Table 1.

Table 3: IV Quantile Regression Results (Continued)

	(5)		(6)		(7)		(8)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.013)	0.000 (0.566)	0.000 (0.017)	0.000 (0.667)	0.000 (0.014)	0.000 (0.878)	0.000 (0.018)	0.000 (0.744)
0.2	0.000 (0.017)	0.000 (0.755)	−0.040 (0.037)	0.690 (5.422)	0.000 (0.018)	0.000 (1.171)	−0.038 (0.020)	0.000 (0.958)
0.3	−0.018 (0.047)	−4.270 (15.28)	−0.042 (0.061)	3.340 (14.78)	−0.030 (0.025)	0.870 (4.928)	−0.043 (0.033)	0.440 (3.955)
0.4	−0.024 (0.033)	2.290 (10.25)	−0.034 (0.033)	−0.270 (4.210)	−0.032 (0.024)	−0.600 (4.548)	−0.043 (0.024)	−0.010 (1.162)
0.5	−0.027 (0.018)	−0.140 (1.205)	−0.042 (0.060)	−2.910 (13.43)	−0.037 (0.026)	−1.300 (6.201)	−0.055 (0.035)	−0.840 (5.087)
0.6	−0.032 (0.034)	−4.740 (10.23)	−0.070 (0.076)	−6.210 (18.40)	−0.037 (0.033)	−4.370 (10.02)	−0.078 (0.047)	−3.290 (10.04)
0.7	−0.043 (0.027)	−3.890 (7.039)	−0.060 (0.043)	−3.400 (8.382)	−0.040 (0.033)	−5.350 (9.584)	−0.047 (0.072)	−6.680 (16.39)
0.8	−0.040 (0.022)	−2.910 (4.023)	−0.057 (0.064)	−7.380 (15.69)	−0.002 (0.045)	−9.450 (11.36)	0.053 (0.094)	−12.89 (18.19)
0.9	0.111 (0.047)	−9.590 (8.541)	0.089 (0.114)	−10.53 (21.70)	0.098 (0.046)	−8.69 (8.081)	0.095 (0.069)	−12.14 (10.84)
GB Controls	Yes		Yes		Yes		Yes	
K/L	No		No		No		Yes	
Instruments	*SCIENTISTS*		*SCIENTISTS*		*MANAGERS*		*MANAGERS*	

Note: This table provides the estimation results of equation (4). Standard errors are in parenthese. GB Controls and K/L indicate whether $INTERMTAR$ and $INTERMNTB$ are controled for and whether $(K/L)_i$, (i =1,2,3,4) are controled for, respectively. Instruments indicates which variables are used as instrumental variables. For the definition of these variables, see Table 1.

Table 3: IV Quantile Regression Results (Continued)

	(9)		(10)		(11)		(12)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.013)	0.000 (0.687)	0.000 (0.017)	0.000 (0.713)	0.000 (0.004)	0.000 (0.000)	0.000 (0.015)	0.000 (0.000)
0.2	0.000 (0.017)	0.000 (0.916)	−0.029 (0.087)	−7.490 (23.21)	0.000 (0.009)	0.000 (0.000)	−0.027 (0.019)	11.90 (14.83)
0.3	−0.023 (0.043)	−4.070 (12.70)	−0.034 (0.046)	−2.270 (8.778)	−0.018 (0.018)	−17.81 (23.37)	−0.032 (0.019)	−4.810 (9.368)
0.4	−0.027 (0.032)	−2.570 (8.367)	−0.041 (0.035)	−0.780 (4.799)	−0.021 (0.019)	−7.160 (12.05)	−0.036 (0.026)	−4.740 (9.130)
0.5	−0.038 (0.024)	−1.260 (5.271)	−0.057 (0.036)	−1.570 (5.868)	−0.019 (0.026)	−7.180 (14.49)	−0.040 (0.043)	−8.320 (18.37)
0.6	−0.051 (0.024)	−0.510 (7.201)	−0.077 (0.041)	−4.040 (8.927)	−0.033 (0.026)	−8.500 (13.08)	−0.075 (0.038)	−4.150 (8.596)
0.7	−0.041 (0.038)	−7.720 (11.80)	−0.060 (0.043)	−3.550 (10.07)	−0.043 (0.027)	−3.890 (6.624)	−0.056 (0.037)	−3.610 (6.416)
0.8	−0.031 (0.031)	−5.280 (8.244)	−0.054 (0.048)	−6.450 (10.61)	−0.029 (0.049)	−4.970 (7.401)	−0.057 (0.077)	−7.360 (11.50)
0.9	0.098 (0.053)	−8.690 (11.19)	0.087 (0.092)	−7.870 (15.12)	0.079 (0.070)	−5.780 (7.446)	0.090 (0.152)	−10.90 (13.81)
GB Controls	Yes		Yes		Yes		Yes	
K/L	No		No		No		Yes	
Instruments	*CROSSELI*		*CROSSELI*		ALL 3 IV's		ALL 3 IV's	

Note: This table provides the estimation results of equation (4). Standard errors are in parenthese. When we use all three instruments (i.e.,*SCIENTISTS*, *MANAGERS*, and *CROSSELI*), standard errors are calculated by 200 bootstrap resampling. GB Controls and K/L indicate whether *INTERMTAR* and *INTERMNTB* are controled for and whether $(K/L)_i$, (i =1,2,3,4) are controled for, respectively. Instruments indicates which variables are used as instrumental variables. For the definition of these variables, see Table 1.

Table 4: Quantile Regression Results of the SP Model

	SP		Modified SP	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
0.2	0.000 (0.000)	0.000 (0.091)	0.000 (0.000)	0.000 (0.030)
0.3	0.000 (0.006)	0.000 (2.658)	0.000 (0.002)	−0.045 (0.696)
0.4	0.020 (0.066)	−6.672 (5.798)	0.006 (0.006)	−1.973 (1.985)
0.5	0.042 (0.007)	−11.33 (2.641)	0.020 (0.006)	−5.125 (1.759)
0.6	0.044 (0.001)	−9.615 (1.618)	0.033 (0.006)	−6.686 (1.721)
0.7	0.046 (0.001)	−7.841 (1.479)	0.049 (0.006)	−7.854 (2.022)
0.8	0.046 (0.000)	−6.076 (1.388)	0.072 (0.008)	−8.666 (2.469)
0.9	0.047 (0.000)	−4.276 (1.186)	0.111 (0.013)	−9.214 (3.103)

Note: This table provides the estimation results of equation (3), using the data simulated from the Surge Protection model. Standard errors are in parenthese.

Table 5: IV Quantile Regression Results of the SP Model

	SP		Modified SP	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
0.2	0.004 (0.008)	−2.000 (4.755)	0.000 (0.031)	−0.093 (0.009)
0.3	0.031 (0.015)	−15.33 (7.286)	0.000 (0.002)	−2.931 (0.292)
0.4	0.042 (0.004)	−17.00 (2.153)	0.006 (0.670)	−7.210 (0.718)
0.5	0.043 (0.001)	−14.24 (1.613)	0.018 (1.921)	−8.934 (0.891)
0.6	0.044 (0.000)	−11.54 (1.349)	0.038 (1.836)	−9.866 (0.985)
0.7	0.044 (0.000)	−9.450 (1.141)	0.053 (2.162)	−10.83 (1.082)
0.8	0.045 (0.000)	−7.430 (1.138)	0.073 (2.618)	−11.08 (1.109)
0.9	0.045 (0.003)	−5.390 (1.074)	0.110 (3.342)	−10.53 (1.078)

Note: This table provides the estimation results of equation (4), using the data simulated from the Surge Protection model. Standard errors are in parenthese.

Figure 1: Equilibrium Campaign Contribution of Politically Organized Industries

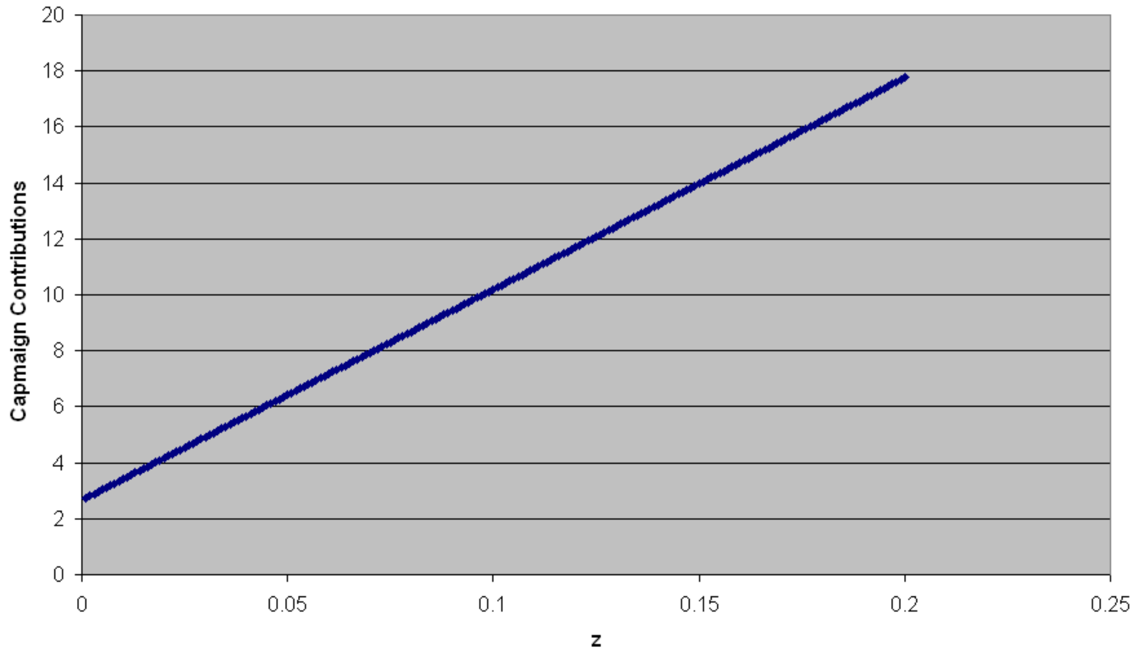


Figure 2: Campaign Contributions and z/e

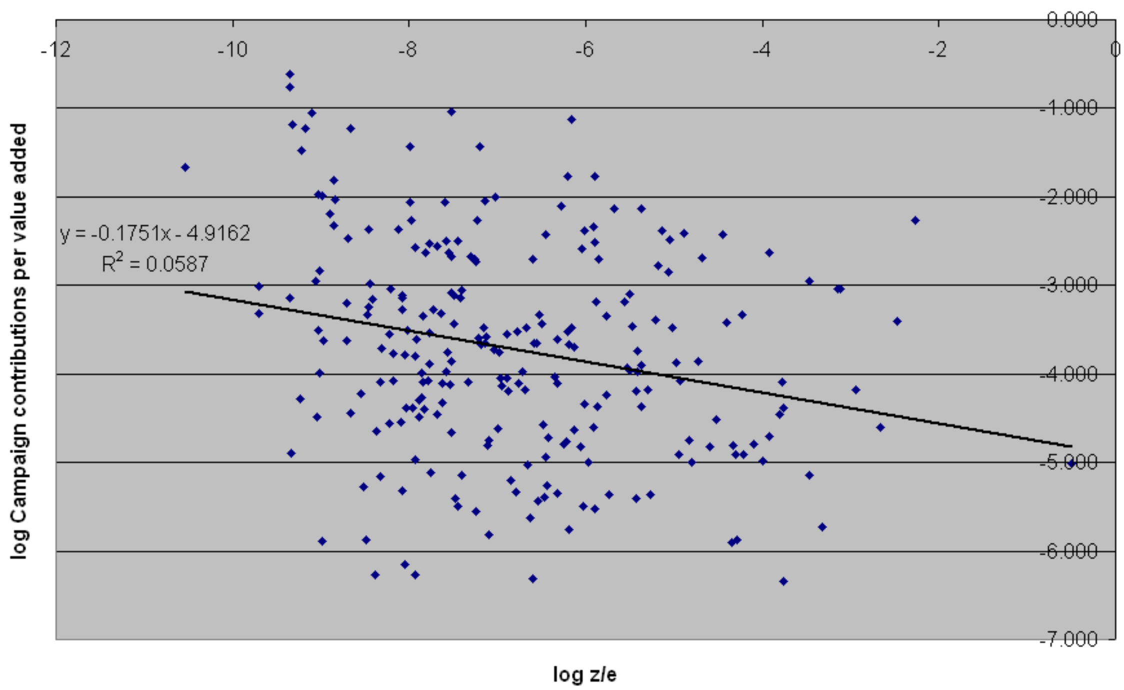


Figure 3: log z/e and log campaign contributions

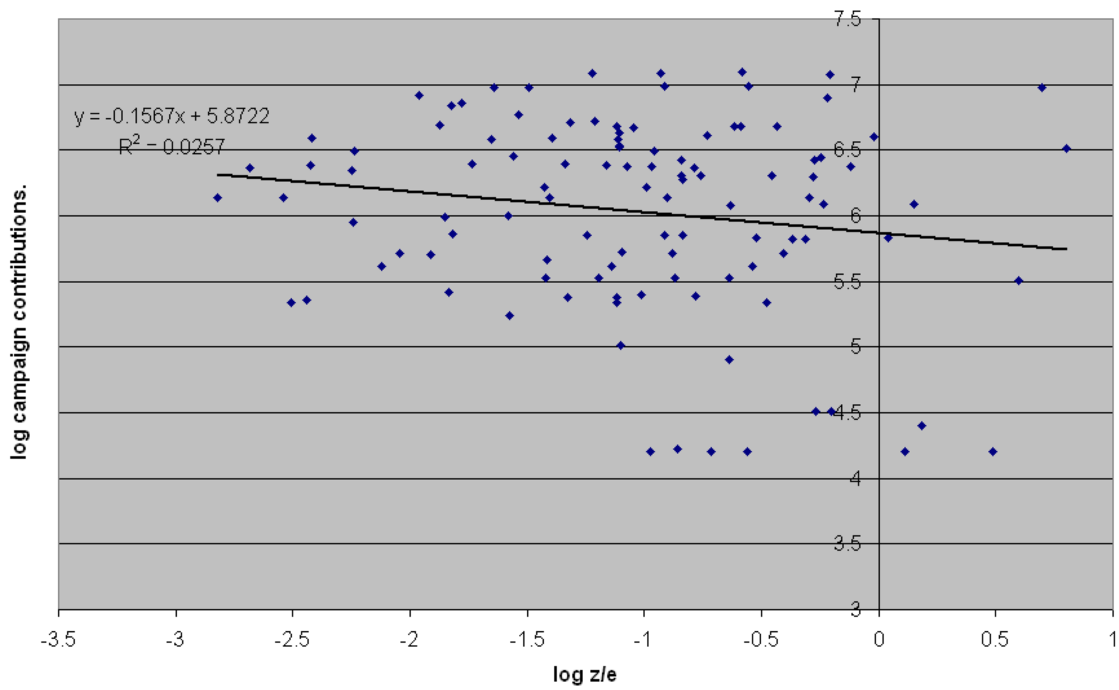


Figure 4: PFS Protection Equation — $t/(1+t) = \beta + \gamma \frac{z}{e} + \delta I \frac{z}{e} + u$

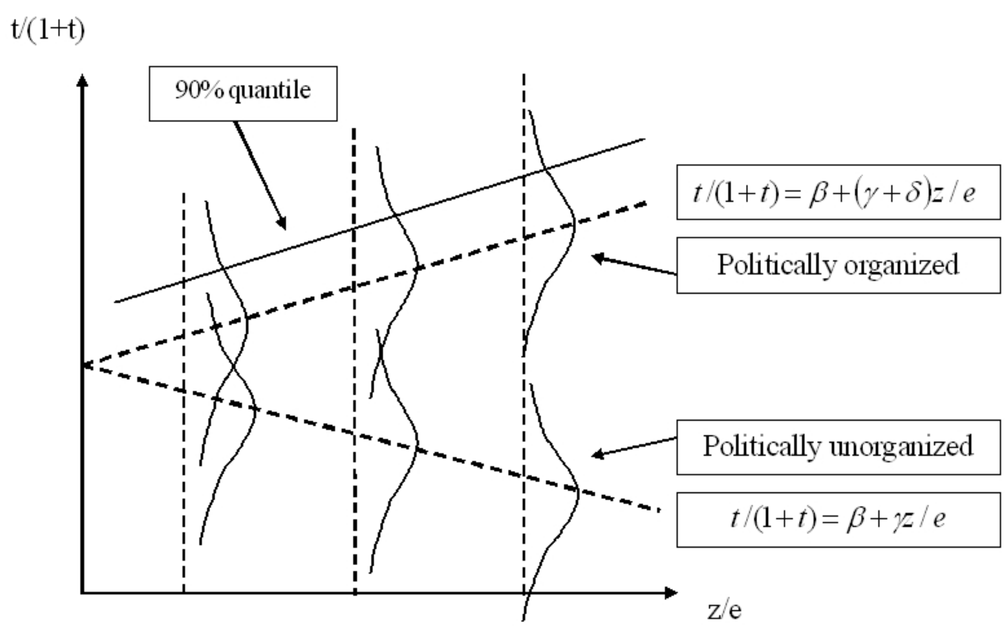


Figure 5: Constructing Costs

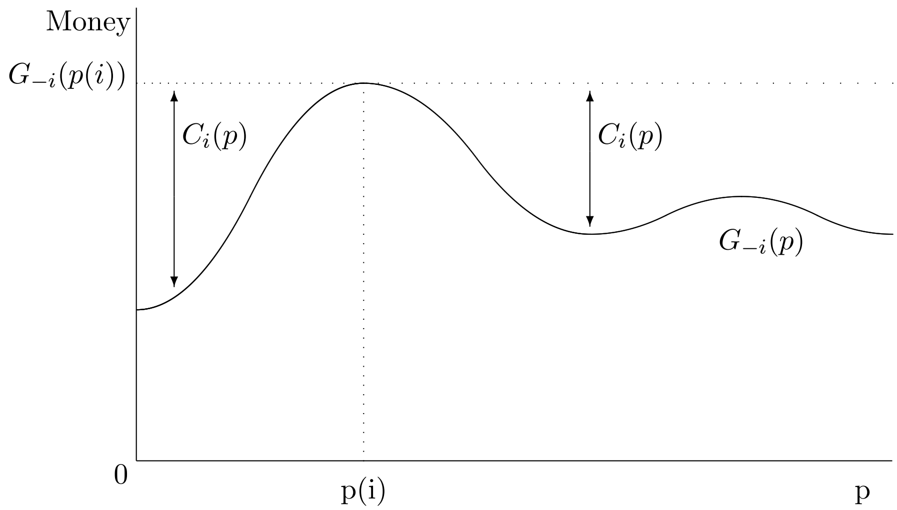
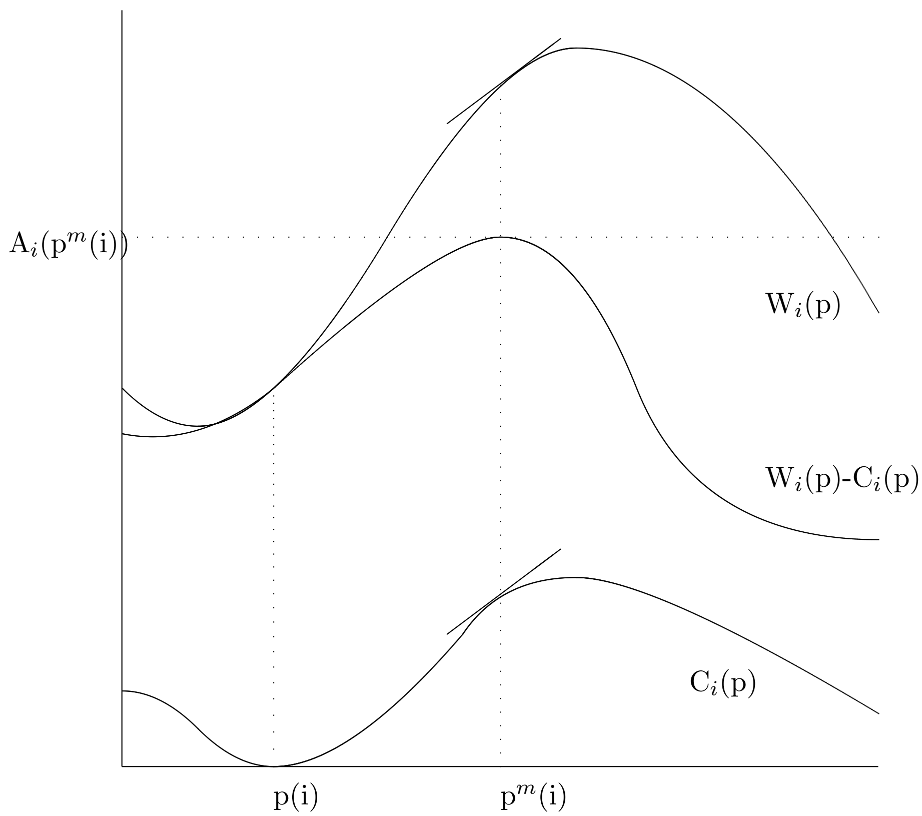


Figure 6: Regret Free Bids



Appendix 1: Equilibrium Campaign Contribution

Deriving Costs

The objective function of the government is denoted by $G(\mathbf{p})$. It is made up of social welfare (which has a weight α given to it) plus the contributions or bribes the government receives from lobbies. Lobby group j in J_0 submits contribution schedule $B_j(p)$. Hence the government objective function is

$$G(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j \epsilon J_0} B_j(\mathbf{p})$$

where the set J_0 consists of the sectors that are organized. Let

$$G_{-i}(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j \neq i, j \epsilon J_0} B_j(\mathbf{p}).$$

This is the objective function of the government when lobby group i does not enter the picture. In this event, the government would choose $\mathbf{p}(\mathbf{i})$ and get $G_{-i}(\mathbf{p}(\mathbf{i}))$. For this reason, Figure A1 depicts $G_{-i}(\mathbf{p})$ as having its maximum at $\mathbf{p}(\mathbf{i})$.

If lobby i wants $\mathbf{p}$ chosen, all it has to do is offer enough to compensate the government for choosing $\mathbf{p}$ rather than $\mathbf{p}(\mathbf{i})$, or offer $C_i(\mathbf{p}) = G_{-i}(\mathbf{p}(\mathbf{i})) - G_{-i}(\mathbf{p})$. Of course, $C_i(\mathbf{p}(\mathbf{i})) = 0$ as drawn in Figure 5. Thus, $C_i(\mathbf{p})$ is the minimum that needs to be offered to get $\mathbf{p}$ to be (weakly) chosen and so can be thought of as the *cost* of $\mathbf{p}$ for i.

The Desired Outcome

Lobby group i has welfare $W_i(\mathbf{p})$. It wants to maximize its net welfare or

$$A_i(\mathbf{p}) = W_i(\mathbf{p}) - C_i(\mathbf{p}).$$

This maximum occurs as depicted at $\mathbf{p}^m(i)$ as depicted in Figure 6. Note that $W_i(\mathbf{p}) - C_i(\mathbf{p})$ is tangent to $W_i(\mathbf{p})$ at $\mathbf{p} = \mathbf{p}(\mathbf{i})$ as $C_i(\mathbf{p}(\mathbf{i})) = 0$. It lies below it elsewhere.

Now given the contribution functions of all other lobby groups, there are any number of ways for lobby group i to get $\mathbf{p}^m(i)$ chosen by the government. All it has to do is offer a little more than $C_i(\mathbf{p})$ at $\mathbf{p} = \mathbf{p}^m(i)$ and anything weakly below $C_i(\mathbf{p})$ everywhere else. However, as this is a game, what it offers will affect what others want the government to choose and the bribes they offer. This in turn will affect the equilibrium. It is for this reason that such games have a continuum of equilibria.

Choosing a Contribution Function

Suppose lobby i offered contributions (subject to these being non negative) at $\mathbf{p} \neq \mathbf{p}^m(i)$ so that it was as well off if as it is at $\mathbf{p}^m(i)$ were any other point was chosen. After all, at the "right price" any outcome can be acceptable! In this manner, its contribution function keeps it *"regret free"*, at least locally. In other words, it bids $max(0, B_i^*(\mathbf{p}))$ where $W_i(\mathbf{p}^m(i)) - C_i(\mathbf{p}^m(i)) = W_i(\mathbf{p}) - B_i^*(\mathbf{p})$ or rearranging terms

$$\begin{aligned} B_i^*(\mathbf{p}) &= W_i(\mathbf{p}) - [W_i(\mathbf{p}^m(i)) - C_i(\mathbf{p}^m(i))] && (9) \\ &= W_i(\mathbf{p}) - A_i(\mathbf{p}^m(i)) && (10) \end{aligned}$$

where $A_i(\mathbf{p}^m(i)) = K_i$ a constant. $B_i^*(\mathbf{p})$ will lie weakly below $C_i(\mathbf{p})$ since offering $C(\mathbf{p})$ would reduce the lobby's net welfare below $W_i(\mathbf{p}^m(i)) - C_i(\mathbf{p}^m(i))$. This contribution function thus has the shape of $W_i(\mathbf{p})$ where it lies above $A_i(\mathbf{p}^m(i))$ in Figure 6. Note that as shown in Figure 6, in the region close to $\mathbf{p}^m(i)$, $W_i(\mathbf{p})$ must lie above $A_i(\mathbf{p}^m(i))$ so contributions are positive, and least locally, the curvature of the equilibrium bid is the same as that of welfare.

7.1 Implications for Equilibrium

Restricting lobbies to contributions that are "regret free" as done above does two things. First, as explained above, it pins down the contribution functions and gives a unique equilibrium. Second, it yields the useful property that the bids have the same curvature as welfare as is evident from equation (9). In effect, lobbies bid their welfare function less a constant! However, since government chooses $\mathbf{p}$ (the domestic price) to maximize the sum of α weighted social welfare and total contributions, it in effect maximizes the sum of α weighted social welfare and the aggregate welfare of all organized sectors. In other words, the equilibrium outcome of this game is the $\mathbf{p}$ that maximizes

$$Z(\mathbf{p}) = \alpha W(\mathbf{p}) + \sum_{j \epsilon J_0} W_j(\mathbf{p}) + \sum_{j \epsilon J_0} K_j$$

where the K_j's are constants. The equilibrium outcome, thus, is as if the government was maximizing the sum of welfare with greater weight placed on the welfare of organized industry groups! Thus, equilibrium tariffs in this relatively complicated setting can be characterized by performing a simple maximization exercise!

However, the model has predictions, other than those on the equilibrium tariff levels, which are usually not incorporated into the estimation. For example, the contribution function in equilibrium keeps the government indifferent between the outcome in the absence of lobby i participating at all, and the equilibrium outcome, $\mathbf{p}^E$ or

$$0 = [Z(\mathbf{p(i)}) - (W_i(\mathbf{p(i)}) + K_i)] - \left[Z(\mathbf{p}^E)\right].$$

Recall, $W_i(\mathbf{p(i)}) + K_i = 0$, since i can get $\mathbf{p(i)}$ chosen by contributing nothing, so in equilibrium

$$Z(\mathbf{p(i)}) = Z(\mathbf{p}^E).$$

Moreover, as $(W_i(\mathbf{p(i)}) + K_i) = 0$ and $\left(W_i(\mathbf{p}^E) + K_i\right) = B_i^*(\mathbf{p}^E)$,

$$\alpha W(\mathbf{p(i)}) + \sum_{j \epsilon J_0, j \neq i} (W_j(\mathbf{p(i)}) + K_j) = \alpha W(\mathbf{p}^E) + \sum_{j \epsilon J_0, j \neq i} \left(W_j(\mathbf{p}^E) + K_j\right) + B_i^*(\mathbf{p}^E)$$

Hence, if the outcome is $\mathbf{p(i)}$ in the absence of lobby $i's$ participating, and is $\mathbf{p}^E$ or the equilibrium price vector when lobby i does participate, then lobby i pays the difference in $\alpha W(\mathbf{p}) + \sum_{j \epsilon J_0, j \neq i} W_j(\mathbf{p})$ evaluated at these two points.

$$B_i^*(\mathbf{p}^E) = \alpha W(\mathbf{p(i)}) + \sum_{j \epsilon J_0, j \neq i} W_j(\mathbf{p(i)}) - \left[\alpha W(\mathbf{p}^E) + \sum_{j \epsilon J_0, j \neq i} W_j(\mathbf{p}^E)\right]$$

Thus, if lobbying by a group i results in distortions that result in a large loss in $\alpha W(\mathbf{p}) + \sum_{j\epsilon J_0, j\neq i} W_j(\mathbf{p})$, then equilibrium contributions must be large. Of course, if the outcome with lobby i not participating is not very different in welfare terms from that when it does, then equilibrium contributions could be small.

Appendix 2: Quantile Regression

Proposition 1 (***Quantile Regression***) *Assume that (1) Z_j is bounded below by a positive number, i.e. there exists $\underline{Z} > 0$ such that $Z_j \geq \underline{Z}$, (2) ϵ_j has a smooth density function which has support that is bounded from above and below, (3) ϵ_j is independent of both Z_j and and I_j, and (4) $\delta > 0$. Then, for τ sufficiently close to 1, τ quantile conditional on Z_j can be expressed as*

$$Q_T(\tau|Z_j) = F_\epsilon^{-1}(\tau') + (\gamma+\delta)Z_j \tag{11}$$

where

$$\tau' = \frac{\tau - P(I_j = 0)}{P(I_j = 1)}. \tag{12}$$

Proof. For any $0 < \tau < 1$, for any $\overline{T} > 0$,

$$P\left(T_j \leq \overline{T}|Z_j\right) = P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) P(I_j = 0) + P\left(\epsilon_j \leq \overline{T} - (\gamma+\delta)Z_j\right) P(I_j = 1). \tag{13}$$

Let

$$\overline{T} = F_\epsilon^{-1}(\tau') + (\gamma+\delta)Z_j \tag{14}$$

where

$$\tau' = \frac{\tau - P(I_j = 0)}{P(I_j = 1)}, \text{ or } \tau = P(I_j = 0) + \tau' P(I_j = 1). \tag{15}$$

From equation (15), we can see that for $\tau \nearrow 1$, $\tau' \nearrow 1$ as well. Hence, for τ sufficiently close to 1, we have τ' close enough to 1 such that

$$F_\epsilon^{-1}(\tau') + \delta Z_j \geq F_\epsilon^{-1}(\tau') + \delta \underline{Z} > F_\epsilon^{-1}(1).$$

Hence,

$$\overline{T} = F_\epsilon^{-1}(\tau') + (\gamma+\delta)Z_j > F_\epsilon^{-1}(1) + \gamma Z_j$$

and

$$P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) \geq P\left(\epsilon_j \leq F_\epsilon^{-1}(1)\right) = 1$$

which results in

$$P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) = 1. \tag{16}$$

Substituting equations (14), (15), and (16) into (13), we obtain

$$\begin{aligned} P\left(T_j \leq \overline{T}|Z_j\right) &= P(I_j = 0) + P\left(\epsilon_j \leq F_\epsilon^{-1}\left(\tau'\right)\right) P\left(I_j = 1\right) \\ &= P(I_j = 0) + \tau - P\left(I_j = 0\right) = \tau. \end{aligned}$$

Therefore, for τ sufficiently close to 1,

$$Q_T\left(\tau|Z_j\right) = \overline{T} = F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j.$$

■

We make two remarks on the assumptions. First, we assume that ϵ_j has bounded support (assumption 2). This assumption is reasonable since the protection measure is usually derived from the NTB coverage ratio (e.g., Goldberg and Maggi, 1999; Gawande and Bandyopadhyay, 2000) and therefore it is clearly bounded above and below. Second, we assume that ϵ_j is independent of both Z_j and and I_j (assumption 3). This is rather a strong assumption and will be relaxed next. In particular, we allow Z_j to be correlated with ϵ_j.

Assume the model is as follows:

$$\begin{aligned} T_j^* &= \gamma Z_j + \epsilon_j \text{ if } I_j = 0 \\ T_j^* &= \left(\gamma + \delta\right) Z_j + \epsilon_j \text{ if } I_j = 1 \end{aligned}$$

where

$$Z_j = g\left(W_j, v_j\right).$$

W_j is an instrument vector and v_j is a random variable independent of W_j. We will show that $\beta\left(\tau\right) \to \left(\gamma + \delta\right) > 0$ as $\tau \nearrow 1$.

Let us define u_j as follows:

$$\epsilon_j = E\left[\epsilon_j|v_j\right] + u_j, \quad u_j \equiv \epsilon_j - E\left[\epsilon_j|v_j\right],$$

where u_j is assumed to be i.i.d. distributed. For the sake of simplicity, we assume that both u_j and $E\left[\epsilon_j|v_j\right]$ are uniformly bounded, hence so is ϵ_j. Furthermore,

$$T_j = \max\left\{T_j^*, 0\right\}.$$

Then, for $I_j = 0$ the model satisfies the assumptions A1-A5 of Chernozhukov and Hansen (2006). Similarly for $I_j = 1$. Therefore, from Theorem 1 of Chernozhukov and Hansen (2006), it follows that

$$P\left(T \leq F_\epsilon^{-1}(\tau) + \gamma Z_j|W_j\right) = \tau \text{ for } I_j = 0,$$

and

$$P\left(T \leq F_\epsilon^{-1}(\tau) + (\gamma + \delta) Z_j|W_j\right) = \tau \text{ for } I_j = 1.$$

Proposition 2 (***Quantile IV***) *Assume that Z_j is bounded below by a positive number, i.e. there exists $\underline{Z} > 0$ such that $Z_j \geq \underline{Z}$. Then, for τ sufficiently close to* 1*,*

$$P\left(T \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma + \delta) Z_j|W_j\right) = \tau,$$

where

$$\tau' = \frac{\tau - P\left(I_j = 0\right)}{P\left(I_j = 1\right)}.$$

Proof.

$$\tau' = \frac{\tau - P\left(I_j = 0\right)}{P\left(I_j = 1\right)}, \text{ or } \tau = P\left(I_j = 0\right) + \tau' P\left(I_j = 1\right).$$

Then,

$$
\begin{aligned}
& P\left(T_j \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j | W_j\right) \\
= \; & P\left(\epsilon_j + \gamma Z_j \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j | W_j\right) P\left(I_j = 0\right) \\
& + P\left(\epsilon_j + (\gamma+\delta) Z_j \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j | W_j\right) P\left(I_j = 1\right) \\
= \; & P\left(\epsilon_j \leq F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) P\left(I_j = 0\right) + P\left(\epsilon_j \leq F_\epsilon^{-1}\left(\tau'\right) | W_j\right) P\left(I_j = 1\right) \\
= \; & P\left(\epsilon_j \leq F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) P\left(I_j = 0\right) + \tau' P\left(I_j = 1\right)
\end{aligned}
$$

From the definition of τ', for $\tau \nearrow 1$, $\tau' \nearrow 1$ as well. Because ϵ is uniformly bounded, for τ sufficiently close to 1, we have τ' close enough to 1 such that

$$F_\epsilon^{-1}\left(\tau'\right) + \delta \underline{Z} > F_\epsilon^{-1}(1).$$

Hence,

$$P\left(\epsilon_j \leq F_\epsilon^{-1}\left(\tau'\right) + \delta Z_j | W_j\right) = 1.$$

Therefore,

$$P\left(T_j \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j | W_j\right) = P(I_j = 0) + \tau' P\left(I_j = 1\right) = \tau.$$

It follows that for τ sufficiently close to 1,

$$P\left(T \leq F_\epsilon^{-1}\left(\tau'\right) + (\gamma+\delta) Z_j | W_j\right) = \tau.$$

■