Queen's Economics Department Working Paper No. 1132

A Quantile Based Test of Protection for Sale Model

Susumu Imai
Queen's University

Hajime Katayama
University of Sydney

Kala Krishna
Pennsylvania State University and NBER

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

8-2007

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

July 25, 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 provides the following prediction: in the quanitle regression of the protection measure on the inverse import penetration ratio divided by the import demand elasticity, its coefficient should be positive at the quantile close to one. We examine this prediction using the data from Gawande and Bandyopadhyay (2000). The results do not provide any evidence favoring the PFS model.

1. Introduction

Recently there has been much interest in political economy aspects of trade policy. This growing interest is in part triggered by the 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) found that the data on trade protection are consistent with predictions by the PFS model. In particular, their results show that as predicted by this framework, protection is positively related to the import penetration ratio for politically unorganized industries, while negatively related for politically organized ones.

An important issue in these empirical studies is how to classify industries into politically organized and unorganized ones. When classifying industries, Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) have encountered the following problem: while only politically organized industries are assumed to make campaign contributions in the PFS model, their data indicate that all industries make Political Action Committees' (PAC) contributions. Thus, if they were to follow the assumption in the model, all industries would be classified as politically organized. To overcome this problem, Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) used some simple rules for classification. However, those rules are somewhat arbitrary and subject to the possibility of misclassification.

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 can lobby for both trade and domestic policies. In their model, it is possible that some industries are politically unorganized for trade policies but 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.

This paper proposes a new approach to testing the PFS model. Unlike most previous studies, our approach does not require classification of industries into organized and unorganized ones. This is important both because of the above mentioned problems in such a classification and because political contribution data itself is not available for most countries. In this manner, our approach can expand the realm of application of such models.

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 the above argument within the framework of recent work on quantile regressions and quantile IV's. To empirically test this implication, we use estimation techniques such as quantile regression (Koenker and Bassett, 1978) and instrumental variable quantile regression (Chernozhukov and Hansen, 2004a; 2004b, 2006). Our results suggest that there is no strong evidence in favor of the PFS model. The point estimates indicate that the inverse import penetration ratio is negatively related to the protection measure at high quantiles, which is the exact opposite of what the PFS model predicts. Importantly, this evidence is robust to a number of sensitivity analyses.

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 the 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.[1] By restricting agents to bids that are "regret free" equilibrium bids have the same curvature as welfare, and a unique equilibrium can be obtained.[2] The equilibrium outcome, thus, is as if the government was maximizing weighted social welfare ($W(p)$ where 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. Thus, 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 and the welfare of agents in

[1] 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".

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

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 employed in industry j, wage is unity, $\frac{N_j}{N}$ is the share of workers employed in the jth industry, 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[3]

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

where so $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, α_i is the share of labor 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 $W(p)$ with respect to p_j gives

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

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

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

[3]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 employment share 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

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

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 $z_j = \frac{x_j(p_j)}{m_j(p_j)}$ and $e_j = -m_j'(p_j)\frac{p_j}{m_j(p_j)}$. This is the basis of the key estimating equation. Note that protection is predicted to be positively related to $\frac{z_j}{e_j}$ 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

For the key equation to be estimable, an error term is added in a linear fashion:

$$\frac{t_j}{1+t_j} = \gamma\frac{z_j}{e_j} + \delta I_j \frac{z_j}{e_j} + \varepsilon_j. \tag{2.1}$$

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 (2.1) can be modified as follows:

$$\frac{t_j}{1+t_j} = Max\left\{\gamma\frac{z_j}{e_j} + \delta I_j\frac{z_j}{e_j} + \epsilon_j, 0\right\}. \tag{2.2}$$

The PFS model provides the following well-known predictions on the coefficients on $\frac{z_j}{e_j}$ and $I_j\frac{z_j}{e_j}$: $\gamma < 0$, $\delta > 0$ and $\gamma + \delta > 0$.[4] To test these predictions, equations (2.1) and (2.2)(hereafter 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) use 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

[4] 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 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 quota model that has no PFS element, one will obtain the parameter estimates consistent with the PFS hypothesis.

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. The idea is that in organized industries, an increase in contributions would likely occur when import penetration increased.

Note that both these two procedures are questionable. The procedure used in Goldberg and Maggi (1999) implicitly assumes that all the contributions are directed towards influencing trade policies. Moreover, any non zero 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. The procedure used by Gawande and Bandyopadhyay (2000) might have a potential identification problem, since a function of the import penetration is used to classify industries and the import penetration divided by the exchange rate is concurrently used as a regressor in the PFS equation. That is, the positive coefficient on the interaction term of the political organization and inverse import penetration ratio could be because the protection measure is a nonlinear function of the inverse import penetration ratio and the exchange rate.

Recently, Cadot et. al. (2006) propose a different approach that that does not require any data on political organization. Instead of deriving the political organization dummy in an ad-hoc manner, they propose to recover it as a by-product of the estimation process. Specifically, they initially set the political organization dummy to zero for every industry. Then, they estimate the PFS equation and obtain the error terms. If the error term of an industry is greater than some threshold value, its political organization dummy is set to be one. The idea is that such industries do not fit the unorganized category.

Using the generated political organization dummies, they again estimate the PFS equation and obtain the error terms. They repeat the procedures of generating the political organization dummies and estimating the PFS equation until the parameters converge. Their method is attractive since information that has been used to classify industries (e.g., contributions) is unavailable in many countries. However, their approach by construction creates a positive correlation between the error term and the generated political organization dummies, which cannot be overcome by any conventional instruments.

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 similar to that of Cadot, Grether and Olarreaga (2006) (C-G-O for short) in the sense that the approach allows us to test the PFS model without using data on political contributions, directly as in G-M or indirectly as in G-B or iteratively as in C-C-O, to construct an organization dummy. However, our approach substantially differs from theirs: instead of classifying industries as organized or not in some manner, our estimation procedure relies heavily on the relationship between observables implied by the PFS model.

Equation (2.2) and the restrictions on the coefficients have at least two implications. First, as has been discussed in the literature, $\frac{z_j}{e_j}$ has a negative effect on the level of protection for politically unorganized industries while it has a positive effect for politically organized industries. Second, given $\frac{z_j}{e_j}$, politically organized industries have higher protection. These implications lead to the following claim: given $\frac{z_j}{e_j}$, high protection industries are more likely to be politically organized and thus effect of an the increase in $\frac{z_j}{e_j}$ on protection tends to be that of politically organized industries. The relevant proposition, and proof, can be found in Appendix 1. The proposition essentially states that in the quantile regression of $\frac{t_j}{1+t_j}$ on $\frac{z_j}{e_j}$, the coefficient on $\frac{z_j}{e_j}$

should be close to at the quantile close to $\gamma + \delta > 0$ at the quantiles close to $\tau = 1$.

Let $T_j = \frac{t_j}{1+t_j}$ and $Z_j = \frac{z_j}{e_j}$. Using quantile regression (Koenker and Bassett, 1978), we estimate the following equation:

$$Q_T(\tau|Z_j) = \alpha(\tau) + \beta(\tau) Z_j/10000. \tag{3.1}$$

where τ denotes quantile, $T_j = \frac{t_j}{1+t_j}$ and $Z_j = \frac{z_j}{e_j}$, and $Q_T(\tau|Z_j)$ is the conditional $\tau-$th quantile function of T. If the PFS model is correct, it is expected that $\beta(\tau)$ converges to $(\gamma + \delta) > 0$ as the quantile, τ, approaches its highest level of unity from below.[5] We use part of the data used in Gawande and Bandyopadhyay (2000)[6]. The data consist of 242 four-digit SIC industries in the U.S. See Gawande and Bandyopadhyay (2000) for a description of the variables.

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 endogeneity, the main prediction of the PFS model in terms of

[5] The estimation results are presented in Table 1. The estimation is done by using a MATLAB code written by Christian Hansen (available at http://faculty.chicagogsb.edu/christian.hansen/research).

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

our quantile approach does not change. The relevant proposition (proposition 2), an analogue of proposition 1, is presented in Appendix $1B$. 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):

$$P\left(T_j \leq \alpha\left(\tau\right) + \beta\left(\tau\right) Z_j/10000 | W_j\right) = \tau \tag{4}$$

where W is a set of instrumental variables.

Importantly, nowhere in equations (3) and (4) 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 whereby we can avoid biased results due to misclassification.

4. A Brief Description of the Data

We use part of the data used in Gawande and Bandyopadhyay (2000) (hereafter also referred to as GB). 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. Table 2 provides the sample statistics of the variables. As is clear from Table 2, 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 (3). 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, all quantiles) in favor of the one-sided alternative that $\beta(\tau) > 0$. Moreover, the point estimates indicate that contrary to the PFS prediction, $\beta(\tau)$ 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 corner solution ($T = 0$) greatly affect 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 in Gawande and Bandyopadhyay (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 addition, 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; Biesebroeck et al., 2004) estimate a system of equations: equation (2) 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.

As the table indicates, $\beta(\tau)$ starts from zero at the quantile $\tau = 0.4$ (since there are a large number group of unprotected industries for whom the coverage ratio is zero) and **decreases** as τ goes from 0.4 to 0.9. Note that this is the **opposite** of what the PFS model predicts, casting doubt on the validity of the PFS model. It is fair to say that our argument here relies on the point estimates. The estimated standard errors are rather large and none of the $\beta(\tau)$'s is significantly different from zero. If the reader is not satisfied with our argument based on the point estimates, the evidence should be interpreted as suggesting that there is no strong evidence in favor of the PFS model (This applies to also to evidence from the instrumental variable quantile regression presented in the next subsection). Two aspects of the results are worth mentioning. First, α and β are estimated to start from zero at the 0.4 quantile, suggesting the corner solution ($T_j = 0$) greatly affect 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 in Gawande and Bandyopadhyay (2000), Bombardini (2005), and others are likely to be subject to bias due to such corners. To address this issue, several studies (e.g., Goldberg and Maggi (1999), Biesebroeck et. al., (2004)) estimate a system of equations: equation (2.2), explicitly allowing for such truncation of protection at zero, as well as an import penetration equation, and an equation for political organization. On the other hand, the assumption of normality of the error terms is usually made and this may affect the estimation results. In contrast, our estimation results are unlikely to be subject to the corner solution problem, since we focus mainly on the higher quantiles where the effect of corner solution is minimal. Second, our results are not driven by the parametric assumption on the error term; the quantile regression does not require them.

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; $\beta\left(\tau\right)$ decreases (for the most part) from zero to a negative value with the increase in , contrary to what the PFS model predicts. α and β are found to be be zero at the and quantiles, again suggesting the importance of corner solution.

5.2. IV Quantile Regression Results

Table 3 presents the estimation results of equation (4). Our choice of instruments is guided by GB 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. 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 and 2 of Table 3.1. 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. (columns 3 and 4).

We further control for capital-labor ratio in equation (4). 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. The estimation results are presented in Table 3.2. $\beta(\tau)$ again are estimated to be zero at quantiles $\tau = 0.1$, and are negative at higher quantiles, except at the 90% quantile when instrument set 2 is used.

Another potential source of bias is when the political organization dummy is econometrically endogenous. That is, when the error term of the equation determining the political organization is correlated with the error term of the protection equation (4). In this paper, we are less concerned about it for the following three reasons. First, in GM correlation comes from the measurement of the political organization dummy coming from the campaign contribution, which is likely to be correlated with the protection measure. Since we do not use the political organization dummy, we are not subject to it. Secondly, GM and others have shown that the results do not change when they control for the endogeneity of the political organization. Thirdly, as long as the error term of the equation determining political organization and that of the protection equation are positively correlated, which is the likely direction of correlation, if there is any, then out 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 with positive correlation, 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 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.

Although we use a subset of GB's instruments, our results may be driven by too many instruments. Thus, we further estimate equation (4) 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. The results are presented in Table 4. The results suggest that having many instruments affect 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 all negative at high quantiles, and significant when all three IV's are used and capital labor ratio is included in the set of controls, which is inconsistent with the PFS's prediction.

6. Discussion

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

$$\varepsilon_j = w_j + (1 - I_j)\,\zeta_j$$

then politically unorganized industries would have error terms with much higher variance. 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, it is worth pointing out that Goldberg and Maggi (1999) did test for heteroskedasticity and the null hypothesis of homoskedasticity could not be rejected.[7]

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.

The third possibility is that the PFS model is indeed not the one that explains the data well. If so, what model would better explain the data? Here, we consider an alternative model, the "Surge Protection" model (hereafter, the SP model) recently proposed by Imai et al. (2006). The SP model is meant to loosely replicate the institutional setup in the United States. It is a simple non-optimizing model where politically organized industries can obtain a limit on imports if imports increase above a specified threshold. The idea is that today most countries have signed the GATT and joined the WTO. In doing so, they have bound their tariffs and committed to limits on their ability to change trade policy. As a result, the main scope for trade policy lies in the safeguard or escape clause realm where temporary protection may be afforded an industry that is under stress and organized enough to lobby for

[7] If equation (**??**) is indeed the error structure, then the PFS equation is modified to be:

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

Importantly, the modified equation has an additional term $1 - I_j$ with a random coefficient ς_j. Then, the original lobbying model needs to be substantially modified so that the reduced form of the PFS equation 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.

protection. The detail of the SP model is provided in Appendix 2 and further details are explicated in Imai et al. (2006).

To examine the validity of the SP model, we conduct the following exercise. First, artificial data are simulated from a calibrated version of the SP model. The parameters in the model are set in the exactly same manner as in Imai et al. (2006); the simulated data reasonably match the aggregate statistics of the U.S. data, as illustrated in Imai et al. (2006). We then estimate equations (3) and (4) on the simulated data using quantile and IV quantile regressions, respectively . We ask whether the parameter estimates from the simulated data resemble those reported in the previous section. If the SP model is valid, then the patterns exhibited in the former are expected to be similar with those in the latter.

In the original SP model, all politically organized subindustries are assumed to have a uniform level of quota, (See Appendix 2). Since this is rather a strong assumption, we slightly extend the SP model by allowing the quota to be stochastically determined. Specifically, we add some randomness to the quota, i.e.,

$$\hat{Q}_{ij} = \widehat{Q} + \varsigma, \varsigma\tilde{\ }N(0,1)$$

where is the quota level for politically organized subindustries ij. Using simulated data from this model (the modified SP model), we estimate equations (3) and (4) again. The quantile regression results are presented in columns 1 and 2 of Table 5. The coefficients on Z are found to be zero at lower quantiles, and thereafter decrease with quantile, which is consistent with the results of the actual data. A similar pattern is observed for the IV quantile regression (Table 6 Column 1 and 2). It is also noteworthy that the size of β's is by and large similar with that obtained from the actual data with one instrument (Table 4).

The results overall suggest that the feature of the SP model is more consistent with the actual data than the PFS model. The intuition behind the negative coefficient estimate of the SP model is simple. A surge in imports, which increases the import penetration ratio, tends to result in the quota being binding, which corresponds to an increase in the NTB coverage ratio. Hence, the negative relationship is observed between the inverse import penetration ratio and the NTB coverage ratio.

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. 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.

We also examined the validity of the SP model proposed in Imai et al. (2006). Using simulated data arising from the SP model, we run the same regressions as in those for the actual data. The estimated coefficients are more in line with those of the actual data. The findings overall seem to suggest that the SP model is consistent with the data while the PFS is not so.

Table 1: Definition of the Variables

Variable	IV	Description
t		All NTB coverage ratios at the 4-digit SITC level
z		Consumption in 1983 /Total Imports
$INTERMTAR$		Average tariff on intermediate goods use
$INTERMNTB$		Average NTB coverage of intermediate goods use
e	1	Absolute import elasticity after correcting for measurement errors
$ln\left(e\right)$	2	Log of absolute import elasticity after correcting for measurement errors
$ln\left(HERF\right)$	3	Log of Hefindahl index of firm concentration

Table 1 Continued

$ln\,(DOWNSTREAMHR)$	4	Log of percentage of an industry's shipment used as intermediate goods in others
$ln(DOWN$ $-STREAMHERF)$	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	U.S. post-Tokyo round ad valorem tariffs (Ratio)
$PERMELAST$	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, resource intensive
$(K/L)_3$	16	Capital-Labor ratio, general manufacturing
$(K/L)_4$	17	Capital-Labor ratio, capital intensive

Table 2: Quantile Regression[8]

$(Q_T(\tau|Z_j) = \alpha(\tau) + \beta(\tau) Z_j/10000)$

	(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	

[8]Note: This table provide the estimation results of equation (3). Standard errors are in parenthese. "GB Controls" indicate whether $INTERMTAR$ and $INTERMNTB$ are controlled for. For the definition of these variables, see Table 1.

Table 3.1: IV Quantile Regression[9]

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.003)	0.000 (0.407)	0.000 (0.002)	0.000 (0.270)
0.2	0.000 (0.012)	0.000 (0.402)	0.000 (0.011)	0.000 (0.369)
0.3	−0.025 (0.011)	−0.370 (0.357)	−0.026 (0.011)	−0.370 (0.287)
0.4	−0.028 (0.009)	−0.200 (0.621)	−0.029 (0.009)	−0.200 (0.421)
0.5	−0.031 (0.023)	−0.270 (1.395)	−0.026 (0.023)	−0.270 (1.091)
0.6	−0.053 (0.023)	−0.080 (2.153)	−0.053 (0.024)	−0.080 (1.184)
0.7	−0.044 (0.015)	−0.130 (2.403)	−0.044 (0.014)	−0.130 (1.611)
0.8	−0.046 (0.016)	0.020 (2.722)	−0.046 (0.014)	0.020 (1.826)
0.9	−0.002 (0.044)	−0.230 (3.572)	−0.001 (0.042)	−0.230 (3.383)
GB Controls	Yes		Yes	
K/L	No		No	
Instruments	Set 1		Set 2	

[9]Note: This table provide 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 3.2 IV Quantile Regression

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.013)	0.000 (0.592)	0.000 (0.013)	0.000 (0.582)
0.2	−0.037 (0.018)	−0.050 (0.496)	−0.036 (0.017)	−0.110 (0.944)
0.3	−0.037 (0.012)	−0.050 (0.391)	−0.036 (0.010)	−0.190 (0.338)
0.4	−0.060 (0.017)	−0.020 (0.798)	−0.043 (0.017)	−0.140 (0.359)
0.5	−0.043 (0.030)	−0.250 (1.241)	−0.060 (0.033)	−0.250 (0.828)
0.6	−0.100 (0.033)	−0.540 (2.085)	−0.103 (0.035)	−0.270 (1.602)
0.7	−0.080 (0.028)	−0.120 (2.388)	−0.080 (0.031)	−0.120 (2.031)
0.8	−0.059 (0.022)	−0.160 (2.794)	−0.059 (0.024)	−0.160 (2.387)
0.9	−0.037 (0.054)	−0.250 (3.077)	−0.070 (0.064)	4.190 (3.444)
GB Controls	Yes		Yes	
K/L	Yes		Yes	
Instruments	Set 1		Set 2	

Table 4: IV Quantile Regression [10]

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.013)	0.000 (0.566)	0.000 (0.017)	0.000 (0.667)
0.2	0.000 (0.017)	0.000 (0.755)	−0.040 (0.037)	0.690 (5.422)
0.3	−0.018 (0.047)	−4.270 (15.28)	−0.042 (0.061)	3.340 (14.78)
0.4	−0.024 (0.033)	2.290 (10.25)	−0.034 (0.033)	−0.270 (4.210)
0.5	−0.027 (0.018)	−0.140 (1.205)	−0.042 (0.060)	−2.910 (13.43)
0.6	−0.032 (0.034)	−4.740 (10.23)	−0.070 (0.076)	−6.210 (18.40)
0.7	−0.043 (0.027)	−3.890 (7.039)	−0.060 (0.043)	−3.400 (8.382)
0.8	−0.040 (0.022)	−2.910 (4.023)	−0.057 (0.064)	−7.380 (15.69)
0.9	0.111 (0.047)	−9.590 (8.541)	0.089 (0.114)	−10.53 (21.70)
GB Controls	Yes		Yes	
K/L	No		Yes	
Instruments	Scientists		Scientists	

[10] GB Controled for,with strong instruments only Note: This table provide the estimation results of equation (4). Standard errors are in parenthese. They are calculated by 200 bootstrap resampling. Both INTERMTAR and INTERMNTB are controled for. “K/L” indicate whether (K/L)i, (i =1,2,3,4) are controled for. “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 4 Continued

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.014)	0.000 (0.878)	0.000 (0.018)	0.000 (0.744)
0.2	0.000 (0.018)	0.000 (1.171)	−0.038 (0.020)	0.000 (0.958)
0.3	−0.030 (0.025)	0.870 (4.928)	−0.043 (0.033)	0.440 (3.955)
0.4	−0.032 (0.024)	−0.600 (4.548)	−0.043 (0.024)	−0.010 (1.162)
0.5	−0.037 (0.026)	−1.300 (6.201)	−0.055 (0.035)	−0.840 (5.087)
0.6	−0.037 (0.033)	−4.370 (10.02)	−0.078 (0.047)	−3.290 (10.04)
0.7	−0.040 (0.033)	−5.350 (9.584)	−0.047 (0.072)	−6.680 (16.39)
0.8	−0.002 (0.045)	−9.450 (11.36)	0.053 (0.094)	−12.89 (18.19)
0.9	0.098 (0.046)	−8.690 (8.081)	0.095 (0.069)	−12.14 (10.84)
GB Controls	Yes		Yes	
K/L	No		Yes	
Instruments	Managers		Managers	

Table 4 Continued

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.013)	0.000 (0.687)	0.000 (0.017)	0.000 (0.713)
0.2	0.000 (0.017)	0.000 (0.916)	−0.029 (0.087)	−7.490 (23.21)
0.3	−0.023 (0.043)	−4.070 (12.70)	−0.034 (0.046)	−2.270 (8.778)
0.4	−0.027 (0.032)	−2.570 (8.367)	−0.041 (0.035)	−0.780 (4.799)
0.5	−0.038 (0.024)	−1.260 (5.271)	−0.057 (0.036)	−1.570 (5.868)
0.6	−0.051 (0.024)	−0.510 (7.201)	−0.077 (0.041)	−4.040 (8.927)
0.7	−0.041 (0.038)	−7.720 (11.80)	−0.060 (0.043)	−3.550 (10.07)
0.8	−0.031 (0.031)	−5.280 (8.244)	−0.054 (0.048)	−6.450 (10.61)
0.9	0.098 (0.053)	−8.690 (11.19)	0.087 (0.092)	−7.870 (15.12)
GB Controls	Yes		Yes	
K/L	No		Yes	
Instruments	CROSSELI		CROSSELI	

Table 4 Continued

τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.004)	0.000 (0.857)	0.000 (0.014)	0.000 (1.054)
0.2	0.000 (0.013)	0.000 (2.763)	−0.027 (0.012)	−11.90 (4.915)
0.3	−0.018 (0.043)	−17.81 (5.865)	−0.032 (0.012)	−4.810 (4.567)
0.4	−0.021 (0.015)	−7.160 (4.612)	−0.036 (0.019)	−4.740 (4.228)
0.5	−0.019 (0.025)	−7.180 (5.405)	−0.040 (0.037)	−8.320 (5.529)
0.6	−0.033 (0.023)	−8.500 (5.132)	−0.075 (0.039)	−4.150 (4.335)
0.7	−0.043 (0.020)	−3.890 (3.976)	−0.056 (0.037)	−3.610 (3.873)
0.8	−0.029 (0.037)	−4.970 (4.058)	−0.057 (0.060)	−7.360 (4.680)
0.9	0.079 (0.070)	−5.780 (4.700)	0.090 (0.113)	−10.90 (5.084)
GB Controls	Yes		Yes	
K/L	No		Yes	
Instruments	All 3 IV's		All 3 IV's	

Table 5: Quantile Regression Estimates of Surge Protection Model

	SP		Modified SP	
τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
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)	−11.333 (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)

Table 6: Quantile IV Regression Estimates of Surge Protection Model

	SP		Modified SP	
τ	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
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)	−15.330 (7.286)	0.000 (0.002)	−2.931 (0.292)
0.4	0.042 (0.004)	−17.000 (2.153)	0.006 (0.670)	−7.210 (0.718)
0.5	0.043 (0.001)	−14.240 (1.613)	0.018 (1.921)	−8.934 (0.891)
0.6	0.044 (0.000)	−11.540 (1.349)	0.038 (1.836)	−9.866 (0.985)
0.7	0.044 (0.000)	−9.450 (1.141)	0.053 (2.162)	−10.832 (1.082)
0.8	0.045 (0.000)	−7.430 (1.138)	0.073 (2.618)	−11.078 (1.109)
0.9	0.045 (0.003)	−5.390 (1.074)	0.110 (3.342)	−10.531 (1.078)

8. Appendix 1A: Quantile Regression

Let $T_j = \frac{t_j}{1+t_j}$ and $Z_j = \frac{z_j}{e_j}$.

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 has a smooth* *density function which has support that is bounded from above and below, (3) ϵ 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*

$$Q_T(\tau|Z_j) = F_{\epsilon}^{-1}(\tau') + (\gamma + \delta)Z_j \tag{8.1}$$

where

$$\tau' = \frac{\tau - P(I_j = 0)}{P(I_j = 1)}. \tag{8.2}$$

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

$$P\left(T_j \leq \overline{T}|Z_j\right) = P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) P(I_j = 0) + P\left(\epsilon_j \leq \overline{T} - (\gamma + \delta)Z_j\right) P(I_j = 1). \tag{8.3}$$

Let

$$\overline{T} = F_{\epsilon}^{-1}(\tau') + (\gamma + \delta)Z_j \tag{8.4}$$

where

$$\tau' = \frac{\tau - P(I_j = 0)}{P(I_j = 1)}, \text{ or } \tau = P(I_j = 0) + \tau' P(I_j = 1). \tag{8.5}$$

From equation (8.5), 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

$$P\left(\epsilon_j \leq \overline{T} - \gamma Z_j\right) = 1. \tag{8.6}$$

Substituting equations (8.4), (8.5), and (8.6) into (8.3), we obtain

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

Therefore, for τ sufficiently close to 1,

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

■

We make two remarks on the assumptions. First, we assume that ϵ 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 is independent of both Z_j and and I_j (assumption 3). This is rather a strong assumption and will be relaxed when quantile IV's are discussed.

When we introduce IV's we show that $\beta(\tau) \to (\gamma + \delta) > 0$ as $\tau \nearrow 1$.

Assume the model is as follows:

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

where

$$Z_j = g(W_j, v_j).$$

W_j is an instrument vector and v_j is a random variable independent of W_j. 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. 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*,*

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

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

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,

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

From the definition of τ'*, for* $\tau \nearrow 1$*,* $\tau' \nearrow 1$ *as well. Hence, for* τ *sufficiently close to* 1*, we have* τ' *close enough to* 1 *such that*

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

Hence,

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

Therefore,

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

It follows that for τ *sufficiently close to* 1,

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

■

9. Appendix 2: Surge Protection Model.

In what follows, we detail the SP model. Our procedure follows Imai et. al. (2006) and is explained in Appendix 2. First, consider the domestic and foreign goods equilibrium without quota. For each industry i and subindustry j, there are two types of goods: domestic and foreign goods. To make matters simple, we assume that each good's demand depends only on its own price and random shocks and that home is the only source of demand. Let x_{ij}^H be the equilibrium quantity of home goods in industry i subindustry j, and let p_{ij}^H be its equilibrium price.

The equilibrium is described by the demand and supply equations. The demand for industry i subindustry j of the home good depends on a constant, the price of the good, and random terms as follows:

$$\ln x_{ij}^{Hd} = ahd_1 + ahd_2 \ln p_{ij}^H + xhd_i + uhd_{ij}.$$

Similarly, the supply of the same good follows the supply equation:

$$\ln x_{ij}^{Hs} = ahs_1 + ahs_2 \ln p_{ij}^H + xhs_i + uhs_{ij}.$$

The random terms xhd_i and xhs_i are industry specific demand and supply shocks, and hence, common across all subindustries, while uhd_{ij} and uhs_{ij} are subindustry specific demand and supply shocks and are idiosyncratic to each subindustry. All shocks are assumed to be i.i.d. with mean zero normal distributions with standard errors σ_{xhd}, σ_{xhs}, σ_{uhd}, and σ_{uhs}, respectively. Equilibrium satisfies

$$x_{ij}^{Hd} = x_{ij}^{Hs} = x_{ij}^H.$$

Similarly, let import demand be given by

$$\ln x_{ij}^{Md} = amd_1 + amd_2 \ln p_{ij}^M + xmd_i + umd_{ij}$$

and supply by

$$\ln x_{ij}^{Ms} = ams_1 + ams_2 \ln p_{ij}^M + xms_i + ums_{ij}.$$

As before, the random terms xmd_i, xms_i, umd_{ij}, and ums_{ij} are industry and subindustry specific demand and supply shocks. They are distributed i.i.d. normally with means zero and standard errors σ_{xmd}, σ_{xms}, σ_{umd}, and σ_{ums} respectively. Equilibrium satisfies

$$x_{ij}^{Md} = x_{ij}^{Ms} = x_{ij}^{Me}.$$

We assume that there are $n_t = 250$ industries and each industry has $n_j = 6$ subindustries. Each subindustry ij is politically organized with probability Po_i.

We simulate the output and prices of each subindustry by first drawing n_t industry demand and supply shocks xmd_i and xms_i for $i = 1, ..., n_t$ and for each industry i, drawing n_s subindustry demand and supply shocks umd_{ij} and ums_{ij} for $j = 1, ..., n_s$. Then, given these shocks and parameters of the demand and supply equations, we compute the equilibrium price and quantities for each subindustry ij.

We then simulate the political organization for each subindustry and introduce a uniform quota level $\hat{Q}$ for all politically organized subindustries. That is, the quota becomes binding in subindustry ij if the equilibrium output for the foreign goods exceeds $\hat{Q}$. Let d_{ij}^q be the indicator for a binding quota. That is, if x_{ij}^{Me} for subindustry ij exceeds $\hat{Q}$, then actual imports, x_{ij}^M, equal $\hat{Q}$ and $d_{ij}^q = 1$. Otherwise, $x_{ij}^M = x_{ij}^{Me}$ and $d_{ij}^q = 0$. One way of interpreting this is that there is a trigger level of imports, $\hat{Q}$, above which the relevant agency would restrict imports if asked, but only politically organized agencies ask for such protection. In other words, that there are provisions for preventing a surge of imports, but only organized subindustries can actually make use of these provisions perhaps because they can overcome the usual free rider problems.

Next we aggregate subindustry output to the industry level. Total industry equilibrium output is computed as

$$X_i^H = \sum_{j=1}^{n_j} x_{ij}^H$$

for home goods and

$$X_i^M = \sum_{j=1}^{n_j} x_{ij}^M$$

for foreign goods.

We then generate the variables that we use in the estimation as follows. First, we compute the coverage ratio C_i of industry i to be:

$$C_i = \frac{\sum_{j=1}^{n_j} x_{ij}^M d_{ij}^q}{X_i^M}.$$

That is, coverage ratio is the fraction of industry output i where quota is binding. Furthermore, the inverse import penetration ratio, z_i, for industry i is the ratio of domestic production to imports or

$$\frac{X_i^H + X_i^M}{X_i^M} = 1 + z_i.$$

We also derive the political organization dummy of industry i, I_i, as:

$$\begin{aligned} I_i &= 1 \; if \sum_{j=1}^{n_j} I_{ij} > \frac{n_j}{2} \\ &= 0 \;\; otherwise. \end{aligned}$$

10. References

Bombardini, M. (2005) "Firm Heterogeneity and Lobby Participation." *Mimeo. University of British Columbia.*

Cadot Olivier, Jean-Marie Grether and Marcelo Olarreaga (2006) "Endogenous Tariffs in a Common-Agency Model: A New Empirical Approach Applied to India." *mimeo. World Bank*

Chernozhukov, V. and Hansen, C. (2006) "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics, Vol.* 132*, pp.* $491-525$.

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

Gawande Kishore and Usree Bandyopadhyay (2000) "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection." *Review of Economics and Statistics*, *Vol.* 82, *No.* 1, *pp.* $139-152$.

Goldberg Pinelopi K. and Giovanni Maggi (1999) "Protection for Sale: An Empirical Investigation." *American Economic Review*, *Vol.* 89, *No.* 5, *pp.* $1135-1155$.

Grossman Gene and Elhanan Helpman (1994) "Protection for Sale." *American Economic Review*, *Vol.* 84, *No.* 4, *pp.* $833-50$.

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

Matschke Xenia (2006) "Costly Revenue-Raising and the Case for Favoring Import-Competing Industries" *University of Connecticut Working Paper.*

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