Queen's Economics Department Working Paper No. 1114

Protection for Sale or Surge Protection?

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

12-2006

Protection for Sale or Surge Protection?

Susumu Imai* Hajime Katayama†

Queen's University University of Sydney

Kala Krishna‡

Pennsylvania State University and NBER

January 2008

Abstract

This paper asks whether the results obtained from using the standard approach to testing the influential Grossman and Helpman "protection for sale" model of political economy might arise from a simpler setting. A model of imports and quotas with protection occuring in response to import surges, but only for organized industries, is simulated and shown to provide parameter estimates consistent with the protection for sale framework. This suggests that the standard approach may be less of a test than previously thought.

JEL Classification: F13, F14

Keywords: Protection for Sale; Lobbying; Political Economy

*Corresponding author. Dunning Hall, 94 University Avenue, Queen's University, Kingston, Ontario, Canada, K7L 3N6; e-mail: imais@econ.queensu.ca; phone: +1 613 533 2258; fax: +1 613 533 6668

†Room 363, H04 – Merewether Building, The University of Sydney, NSW 2006, Australia; e-mail: h.katayama@econ.usyd.edu.au

‡Department of Economics, The Pennsylvania State University, 523 Kern Graduate Building, University Park, PA 16802; e-mail: kmk4@psu.edu

1 Introduction

The Grossman and Helpman (1994) model of "Protection for Sale" (PFS) has become the most influential one in the political economy of trade over the past decade. It is an equilibrium model of protection which explicitly accounts for lobbying from industries to influence government. Politically organized industries are assumed to propose campaign contribution bid functions that specify the relationship between campaign contribution and tariff. Given those bids from industries, the government chooses the tariffs so as to maximize its objective function, a weighted sum of the campaign contribution and the welfare of the voters. The PFS model provides a clear-cut prediction on the relationships between the level of protection and the import penetration ratio: protection is positively related to the import penetration for politically unorganized industries, but negatively related for politically organized ones. A number of studies have tested this prediction. The first few studies are Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) (GM and GB respectively from here on). While they used the U.S. data, Mitra et al. (2002) and McCalman (2004) used Turkish and Australian data, respectively. Findings in these studies are consistent with the key prediction made by the PFS framework.

Recently, researchers have extended the original PFS model in various directions by incorporating firm size (Bombardini, 2004), foreign and domestic lobbies (Gawande and Krishna, 2004), lobbying of both upstream and down stream producers (Gawande and Krishna, 2005), and labor unions and labor immobility (Matschke and Sherlund, 2006). While the original model accounts for tariffs, its quota version was also constructed and estimated (Facchini et al., 2006). These extensions typically leave the basic predictions of the PFS model unchanged and seem to provide more evidence in support of the PFS model.

Despite much evidence favoring the PFS model, the extent to which past studies did a stringent job of testing the PFS model is an open question. This results from the fact that most past studies did not formally test the PFS model.[1] Past studies simply estimated the protection equation derived by the PFS model and examined whether the signs of the key coefficients follow the pattern predicted by the model. However, such an estimation exercise was typically conducted in the absence of a well-specified alternative model.[2] Therefore, it is possible that some alternative model, not the PFS model, may be the correct one.

This paper first shows that a simple setting, where government provides protection for politically organized industries when imports exceed a trigger level (which we call the "Surge Protection" model), is also consistent with the estimates in the literature. In particular, we simulate a simple equilibrium model of domestic consumption and imports, where imports in industries are subject to an exogenously and uniformly set threshold: if import demand exceeds this threshold either because of supply or demand shocks, then there is pressure for protection. This pressure is more likely to get transformed into actual protection if the industry is organized. In fact, we take an extreme position here and assume that the industry is protected, i.e., the threshold becomes a quota, only if it is organized. Political organizations are set exogenously and randomly. Obviously, in this simple model, there is no strict PFS effect. Parameters are set so that the simulated data roughly match the basic statistics of the actual data. Then, we estimate the key equation

[1] This problem was noticed by Goldberg and Maggi (1999); they mentioned that "(s)trictly speaking, we do not test the G-H model, because we do not have a well-specified alternative hypothesis" (p.1135).

[2] Notable exceptions are Eicher and Osang (2002) and Gawande (1998) who formally tested the PFS model. However, in our view, the results are far from satisfactory. Eicher and Osang (2002) is a good example to make our point. They compared the tariff equation derived by the PFS model and that of the Tariff Function approach by using the Davidson-McKinnon non-nested hypothesis test, concluding that the results are in favor of the PFS model. While this kind of formal approach could be very helpful in making model comparisons, we believe the simplistic approach traditionally being followed can be more misleading than helpful. Even though the tariff equation which they estimated is sufficient for the estimation of the structural parameters, it is a small part of the entire PFS model or the Tariff Function model. Hence, testing the tariff equation only could lead to misleading results; to correctly execute the non-nested model specification tests, one needs to impose all the restrictions of the model on the data. This involves the full solution of the model, which is difficult for the PFS model, and to the best of our knowledge, has not been done in the literature.

of the PFS model on the artificial data following the procedures by GM and GB. We find that coefficient estimates are consistent with the PFS paradigm. Furthermore, similar findings are obtained in the analogous tariff-setting version of the model. Our results therefore suggest that estimation of the protection equation, even though sufficient for all the structural parameters, is not enough to test the validity of the PFS model against alternatives such as the simple Surge Protection (SP) model.

To make our point as clear as possible, we use the SP model because to us it seems a simple way to model the institutional side of trade policy. In all countries, membership in the GATT/WTO restricts the ability of countries to protect domestic industries except under certain circumstances, for example, as a safeguard measure, or under anti-dumping law, or in the 80's as a voluntary export restraint. Though injury has to be shown, such institutional measures would allow protection more easily when industries are threatened with competition, i.e., when imports surge. Moreover, it is also likely to be easier for organized industries to obtain such protection as it involves jumping through some hoops and because they can more easily overcome the usual free rider problems. This idea is captured in the simplest form by the SP model. We use artificial data for estimation, thereby abstracting ourselves from the intricacies of the actual data. In that sense, the paper is in the same spirit as the "counterfactual estimation" in Keller (1998), where he created an artificial trade pattern that was not related to R&D spillovers and "verified" the model of international R&D spillovers.

Our results help explain some puzzle in previous work. Most studies that we are aware of found that political economy factors seem to matter little; the estimate of the weight on contribution relative to welfare placed by the government is typically very low.[3] However,

[3] The estimated low weight on contributions might be partly attributed to the fact that data on contributions is not actually used in the estimation procedures of most previous studies. The only paper we know that actually used contribution data directly is Kee et al. (2005). They assumed that lobbies have a first mover advantage over government as is the norm in this literature, and looked at foreign lobbying

given that contributions are small relative to their effects on firm profits and welfare, one would expect the government to put a reasonably high weight on contributions, because in the PFS model, equilibrium contributions by a group keep the government as well off as in the absence of the lobby group, i.e., just compensate the government.[4] Our results simply suggest that the supposedly low values for the weight on contributions obtained by past studies can be thought of as just a misinterpretation of the parameter estimates; a simpler model than the PFS framework yields similar estimated coefficients, but without the strict PFS interpretation.

Given that findings in the literature favor the PFS model *and* the SP model, one question naturally arises: which model is more plausible? It is important to note that the PFS model is not supported by a recent study by Imai et al. (2008). They use quantile regression and IV quantile regression to examine the relationship between the level of protection at various conditional quantiles and the inverse import penetration ratio. Using the data from Gawande and Bandyopadhyay (2000), they provide evidence that at low quantiles of protection measure, the estimated coefficient on the inverse import penetration ratio is zero, and it becomes negative at higher quantiles.[5] Their results are inconsistent with the PFS model, as at high quantiles of protection where industries should be mostly politically organized, protection should be positively related to inverse import penetration ratio.

To examine whether the SP model is empirically plausible, this paper follows Imai et al.

in the US for preferential access (which reduces tariffs to zero or leaves them unchanged) with world prices given. As a result, the welfare cost to the US is the loss of tariff revenue. This loss is, in essence, compared to the contributions received to obtain a weight on contributions relative to welfare. Their results suggest that the government seems to value contributions five times more than welfare: a vast difference from the results using either the GM or GB approach!

[4] See Rodrik (1995) for an early survey of political economy models in trade and Gawande and Krishna (2003) for a recent one of the empirical work in the area.

[5] From now on, we commit a minor abuse of definition to denote "quantile" for "conditional quantile given inverse import penetration ratio."

(2008) and runs the quantile regression and IV quantile regression on the data simulated from the model. The estimated coefficients are zero at low quantiles and negative at higher quantiles, having the same sign as those of the actual data. We thus conclude that the SP model better describes the relationship between the protection and inverse import penetration ratio at various quantiles than the PFS model.

The paper proceeds as follows. The PFS model is briefly laid out in the next section. Section 3 develops the SP model, which we calibrate to broadly match the data. We then generate data from it. Section 4 runs the standard regressions on the simulated data and shows that the standard results are obtained despite the absence of any strict PFS effects. Section 5 verifies that our results go through even with a tariff version of the SP model. Section 6 conducts some robustness checks. Section 7 then explains why this is happening. Section 8 examines and discusses which model is more consistent with the data. Section 9 concludes.

2 The PFS Model and Its Estimation

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. The PFS model is characterized by a freely traded

Ricardian numeraire (i.e., unlike the other sectors which use mobile labor and a specific factor, only labor is used in production of the numeraire and the world price ties down the domestic price) which insures that sectors are not linked in production via the market for the only mobile factor. It is this combination that allows the "general equilibrium" model to be treated essentially as a sum of partial equilibrium sectors.[6] Some industries are politically organized (i.e., form lobby groups), 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 the contributions made to it and social welfare and puts a relative weight of α on social welfare, $W(\mathbf{p})$ where $\mathbf{p}$ is the domestic price and equals the tariff vector plus the world price $\mathbf{p}^*$[7].

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 that all principals (i.e., lobby groups) are trying to influence. Such games are known to have a continuum of equilibria.[8] By restricting agents to bids that are "truthful" so that their bids have the same curvature as their welfare, a unique equilibrium is obtained.[9] The equilibrium outcome is as if the government was maximizing the weighted sum of the welfare of politically unorganized industries and that of the organized industries with a greater weight on the

[6] We are grateful to the anonymous referee for providing some important insights of the PFS model.

[7] For the remainder of the paper, vectors are written in bold letters.

[8] 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 by an agent at other outcomes affect the optimal choices of other lobbies and as their behavior affects her in return, multiplicity arises naturally. Uniqueness is obtained by pinning down the bids at all outcomes to yeild the same payoff as at the desired one.

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

welfare of organized industries. Thus, equilibrium tariffs can be found by maximizing

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

where J_0 is the set of politically organized industries and the welfare of agents in industry j is

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

where $\pi_j(p_j)$ is producer surplus in industry j, l_j is labor employed in industry j (wage is unity in equilibrium and hence labor income equals l_j), N_j/N is the share of owners of industry j specific factor in total population, while $T(\mathbf{p}) + S(\mathbf{p})$ is the sum of tariff revenue and consumer surplus in the economy.

This is the great charm of the PFS model: not only does it cleanly model where both the demand and the supply of protection are coming from, but the results can be derived from a simple maximization exercise! Small wonder it is so popular.

Differentiating $W_i(\mathbf{p})$ with respect to p_j gives[10]

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

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, α_i is the share of industry i specific factor owners, $m_j'(p_j)$ is the derivative of the demand for imports, and $x_j(p_j) = \pi_j'(p_j)$ denotes the supply of industry j. Differentiating $W(\mathbf{p})$ with respect to p_j gives

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

[10] 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.

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

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

or equivaently,

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

where $\sum_{i\in J_0}\alpha_i = \alpha_L$, the share of owners of specific factors in organized industries, and $\sum_{i\in J_0}\delta_{ij} = I_j$ is unity if j is organized and zero otherwise. Now, using the fact that $(p_j - p_j^*) = t_j p_j^*$ where t_j is defined to be the tariff rate**,** the above equation can be rewritten as

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

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 following key equation, which we call the protection equation:

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

Note that $\gamma = \left[-\alpha_L/\left(\alpha + \alpha_L\right)\right] < 0$, $\delta = 1/\left(\alpha + \alpha_L\right) > 0$, and $\gamma + \delta > 0$; protection is positively related to z_j/e_j if the industry is politically organized, but otherwise negatively related to it. An even stronger prediction is that z_j and e_j do not enter separately once their ratio is controlled for.

2.2 Estimation Methods in GM (1999) and GB (2000)

GM and GB added an error term to equation (2) to permit estimation:

$$\frac{t_j}{1+t_j} = \gamma\frac{z_j}{e_j} + \delta I_j\frac{z_j}{e_j} + \epsilon_j. \tag{3}$$

The error term, ϵ_j, is interpreted as the composite of variables potentially affecting protection that may have been left out, and the measurement error of the dependent variable. Both GM and GB used the coverage ratios for non-tariff barriers (NTBs) as t_j instead of the tariff itself. GB estimated a variant of equation (3) together with the other equations which determine the political contribution and the inverse import penetration ratio. Their protection equation also accounts for protection on intermediate goods and includes as explanatory variables the tariff and NTBs on intermediates goods used by the industry. As z_j and I_j may be endogenous in equation (3), GB used a nonlinear IV estimation technique proposed by Kelejian (1971).

GM explicitly considered the corner solution of the protection measure on the LHS. Using full information maximum likelihood, they estimated the following system of equations. First, the "true level of protection" in industry i, t_i^*, is modelled as follows:[11]

$$\frac{t_i^* e_i}{1 + t_i^*} = \gamma z_i + \delta I_i z_i + \epsilon_i. \tag{4}$$

The true protection level is assumed to be a multiple of the coverage ratio which lies between zero and unity (to account for the boundedness of the coverage ratio in the data):

$$\begin{aligned} t_i &= \tfrac{1}{\mu} t_i^* \quad && \text{if} \quad 0 < t_i^* < \mu, \\ &= 0 && \text{if} \quad t_i^* \leq 0, \\ &= 1 && \text{if} \quad t_i^* \geq \mu, \end{aligned} \tag{5}$$

where μ is exogenously set at the value 1, 2, or 3.[12] Domestic production to import ratios

[11] Note that e_i is moved to the left hand side to alleviate concerns about its endogeneity.

[12] There is no reason for μ not to be less than unity as quotas may be barely binding.

are related to a variety of factors in

$$z_i = \varsigma_1' \mathbf{R}_{1i} + u_{1i}, \tag{6}$$

and whether the industry is politically organized is modelled as

$$I_i = D(\varsigma_2' \mathbf{R}_{2i} + u_{2i} > 0), \tag{7}$$

where $D(\cdot)$ is an indicator function, and $\mathbf{R}_{1i}$ and $\mathbf{R}_{2i}$ are vectors of exogenous variables.

Both GM and GB found that the key parameters γ and δ have the predicted signs and are statistically significant. They also found that the estimate of α is very high (70 in GM and 3715 in GB), suggesting that the government puts a very low weight on contributions.

3 A Simple Model of Imports (Surge Protection Model)

We now develop a simple model of imports that we will simulate. Our model has to have several features to match the key statistics of the data. First, in the data some industries are politically organized and others are not. In our model we simply assume political organization is randomly determined. Second, in the data some politically organized industries are protected by quota and others are not. To capture that in a simple way, we assume that the import quota will bind for politically organized industries whose equilibrium imports exceed some threshold level.

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:

$$\ln x_{ij}^{Hd} = ahd_1 + ahd_2 \ln p_{ij}^{H} + xhd_i + uhd_{ij}. \tag{8}$$

Similarly, the supply of the same good is given by

$$\ln x_{ij}^{Hs} = ahs_1 + ahs_2 \ln p_{ij}^{H} + xhs_i + uhs_{ij}. \tag{9}$$

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 normal distributions though the parameters of the distribution differ. For all i, xhd_i (xhs_i) has mean 0 and standard deviation σ_{xhd} (σ_{xhs}). Similarly, for all ij, uhd_{ij} (uhs_{ij}) has mean 0 and standard deviation σ_{uhd} (σ_{uhs}). Equilibrium satisfies

$$x_{ij}^{Hd} = x_{ij}^{Hs} = x_{ij}^{H}. \tag{10}$$

Similarly, let import demand be given by

$$\ln x_{ij}^{Md} = amd_1 + amd_2 \ln p_{ij}^{M} + xmd_i + umd_{ij}, \tag{11}$$

and supply by

$$\ln x_{ij}^{Ms} = ams_1 + ams_2 \ln p_{ij}^{M} + xms_i + ums_{ij}. \tag{12}$$

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 deviations σ_{xmd}, σ_{xms}, σ_{umd}, and σ_{ums}, respectively. Equilibrium satisfies

$$x_{ij}^{Md} = x_{ij}^{Ms} = x_{ij}^{Me}. \tag{13}$$

We assume that there are n_t industries and each industry has n_j subindustries. Each subindustry ij is politically organized with probability Po_i. We allow for some variation in the political organization probability across industries: $Po_i = 0.9$ with probability 0.3, $Po_i = 0.8$ with probability 0.2, $Po_i = 0.7$ with probability 0.2, and $Po_i = 0.1$ with probability 0.3. This is done to ensure that there is sufficient variation in the numbers of subindustries that are politically organized within industries. If we had only one probability of political organization for every industry, say .6, the fraction of industries that are politically organized will be clustered around .6. We simulate political organization by generating a $(0, 1)$ uniformly distributed random variable u_{pi}, and generate independently another $(0, 1)$ uniformly distributed random variable u_{oij}. If $u_{pi} \leq 0.3$, then $I_{ij} = 1$ if $u_{oij} \leq 0.1$ and $I_{ij} = 0$ otherwise. If $0.3 < u_{pi} \leq 0.5$, then $I_{ij} = 1$ if $u_{oij} \leq 0.7$ and $I_{ij} = 0$ otherwise. If $0.5 < u_{pi} \leq 0.7$, then $I_{ij} = 1$ if $u_{oij} \leq 0.8$ and $I_{ij} = 0$ otherwise. Finally, if $0.7 < u_{pi}$, then $I_{ij} = 1$ if $u_{oij} \leq 0.9$ and $I_{ij} = 0$ otherwise.

We simulate the output and prices of each industry by first drawing n_t industry (import) demand and supply shocks and for each industry, drawing n_j subindustry (import) demand and supply shocks. Then, given these shocks and parameters of the (import) demand and supply equations, we compute the equilibrium price and quantities for each subindustry ij.

We now introduce a uniform quota level $\hat{Q}$ for politically organized subindustries. 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 industries can actually make use of these provisions perhaps because they can overcome the usual free rider problems.

The variables that we use in the estimation are constructed as follows. We first aggregate subindustry output to the industry level. Total industry equilibrium outputs for home goods and foreign goods are computed as $X_i^H = \sum_{j=1}^{n_j} x_{ij}^H$ and $X_i^M = \sum_{j=1}^{n_j} x_{ij}^M$, respectively. Then, we compute the coverage ratio of industry i, C_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 where quota is binding. The inverse import penetration ratio for industry i, z_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 follows:

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

In other words, we call industry i politically organized if more than half of its subindustries are politically organized.

We choose the parameters of the model so that the simulation is reasonably close to the actual data in several dimensions. The parameters of the home goods demand and supply equations are: $ahd_1 = 3.6$, $ahd_2 = -1.3$, $ahs_1 = 3.0$, $ahs_2 = 1.4$, $amd_1 = 1.0$, $amd_2 = -1.5027$, $ams_1 = 1.0$, and $ams_2 = 1.0$. The import demand elasticity, i.e., $-amd_2$, is set at the mean of the industry import demand elasticity from the estimation of Shiells et al. (1986). Furthermore, we set $\sigma_{xhd} = \sigma_{xhs} = 2.0$, $\sigma_{xmd} = \sigma_{xms} = 0.48$, $\sigma_{uhd} = \sigma_{uhs} = 0.2$, and $\sigma_{umd} = \sigma_{ums} = 0.05$.

Table 1: Summary Statistics

	Simulation	Data
Political organization frequency	0.626	0.680
NTB positive	0.541	0.533
Average log output/import ratio	2.354	2.783
Std. deviation of log output/import ratio	1.347	1.620

In Table 1, we compare the simulation of the model to the data used in GB. The summary statistics of the simulation is based on 1000 artificially generated industries, each having 6 subindustries.[13] The model matches the average political organization, NTB coverage ratio, log output/import ratio, and the standard deviation of log output import ratio reasonably closely.

4 Estimating the Protection Equation Using Simulated Data

Using the data simulated from the SP model, we now estimate the PFS protection equation following the procedures of both GB and GM.

[13] Note that the simulation size is set to be much larger than the sample size of the GB's data, 242. This avoids finite sample variation of the sample average; the average over large sample would represent the stochastic model more accurately than that of 242 sample.

4.1 OLS-IV Regression

To replicate the IV estimation done by GB, we estimate the following equation by three stage least squares:

$$\frac{C_i}{1+C_i} \mid amd_2 \mid = \beta_0 + \gamma \frac{(1+z_i)}{10000} + \delta I_i \frac{(1+z_i)}{10000} + u_i, \tag{14}$$

where we scale variables by dividing by 10000 as done by GB. Importantly, we use $1+z_i$, not z_i, as GB and GM use consumption (which equals domestic production plus imports in the standard homogeneous good model) relative to imports, not production relative to imports. Thus, they in effect use $(1+z_i)$ and we follow their lead for comparability. Note that due to the presence of the interaction term, $I_i(1+z_i)$, this choice of variable results in some mis-specification which could affect the estimates of γ and δ as well as β_0. The impact of using one versus the other turns out to be quite small in GM but larger in our model. We discuss more on this when presenting our maximum likelihood results.

We use two sets of instrumental variables. The first set of instruments includes the exogenous home demand and supply shocks and political organization shocks: xhd_i, xhs_i, and u_{pi} (3SLS1). The second set includes these three instruments, their square terms, and interactions (3SLS2). The simulation sample size (i.e., the number of industries) is chosen to be 200 which is close to that used in both GB and GM. Each industry is assumed to have 6 subindustries.

Columns 3-5 in Table 2 present the estimation results of OLS, 3SLS1, and 3SLS2, respectively. All the parameter estimates as well as their standard errors are the average of 10 simulation/estimation exercises. Notice that in all the estimates, γ is negative and significant, δ is positive and significant, and the sum of the two coefficients is positive, just as the PFS model predicts. While our parameter estimates are an order of magnitude

Table 2: OLS-IV Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GB	GM	OLS	3SLS1	3SLS2	OLS	3SLS1	3SLS2
β_0	-0.042		0.315	0.312	0.310	0.306	0.291	0.304
	(0.017)		(0.025)	(0.038)	(0.027)	(0.011)	(0.017)	(0.012)
γ	-3.088	-0.0093	-28.1	-95.8	-47.3	-26.6	-101.0	-58.8
	(1.532)	(0.0040)	(7.91)	(22.9)	(12.3)	(3.56)	(11.2)	(5.68)
δ	3.145	0.0106	35.2	141.4	66.8	35.3	155.6	77.1
	(1.575)	(0.0053)	(9.05)	(31.3)	(15.8)	(3.96)	(14.9)	(7.10)
α	3715	70.43	335.15	73.11	166.50	282.28	63.64	128.98
α_L	0.9819	0.883	0.8308	0.6761	0.6854	0.7540	0.6494	0.6981
Nobs	242	107	200	200	200	1000	1000	1000

Note: Standard errors are in parentheses. In columns (3)-(5), the results are the average of 10 simulation/estimation exercises. GB is from the first column in Table 3A (p.145). GM is from the first column in Table 1 (p.1145).

larger than those of GB (column 1), they are close to those of GM (column 2).[14] These estimates are also close to those estimated for the simulated data following the procedure of GM, as will be prsented later.

As the sample size of 200 is a bit small, the IV estimates may be subject to small sample bias. To see if there is a significant bias in the mean, we also run the same simulation/estimation exercise once with the simulation sample size of 1000. The results are reported in columns 6-8 of Table 2. The estimates do differ as expected, but again follow the patterns predicted by the PFS model. It should be stressed that all of these results are obtained in spite of the fact that the data comes from a simple model where the quota is set exogenously at a uniform level in politically organized subindustries, the import elasticity is set constant across all industries, and political organization is completely exogenous to the system – a much less restrictive model than the PFS model.

[14] Note that we need to divide our coefficients by 10000 to make them comparable to GM's.

4.2 Maximum Likelihood Estimation

Next we follow GM and assume the error terms of equations (4), (6), and (7) are jointly normally distributed: $(\epsilon_i, u_{1i}, u_{2i}) \sim N(0, \Sigma)$.[15] We use full-information maximum likelihood to estimate the parameters of the model. The instruments for $(1 + z_i)$ are the exogenous home demand and supply shocks and political organization shocks: $\mathbf{R}_{1i} = (xhd_i, xhs_i, u_{pi})$. Also, the instruments for the political organization dummy are the exogenous demand and supply shocks as well as the political organization shocks: $\mathbf{R}_{2i} = (xhd_i, xhs_i, xmd_i, xms_i, u_{pi})$. Again, we conduct 10 simulation/estimation exercises with the sample size of 200 and then take the average of those results.

Table 3: Maximum Likelihood Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GM	GM-A3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
β_0		−0.2545		0.178	−0.435		0.127	−0.393
		(0.241)		(0.070)	(0.127)		(0.033)	(0.059)
β_I		0.3851			0.951			0.820
		(0.347)			(0.156)			(0.071)
γ	−0.0093	−0.0092	−0.007	−0.010	−0.003	−0.012	−0.015	−0.004
	(0.0040)	(0.0044)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
δ	0.0106	0.0089	0.011	0.012	0.003	0.019	0.020	0.006
	(0.0053)	(0.0089)	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
α	70.43		88.46	83.62	396.4	51.57	48.56	163.03
α_L	0.883		0.651	0.867	1.033	0.633	0.748	0.656
l			−972.0	−968.6	−950.2	−4755.9	−4749.2	−4685.7
Nobs	107	107	200	200	200	1000	1000	1000

Note: Standard errors are in parentheses. β_I is the coefficient on I_i. l is the log-likelihood. In columns (3)-(5), the results are the average of 10 simulation/estimation exercises. GM is from the first column in Table 1 (p.1145). GM-A3 is from Table A3 (p.1152).

Column 3 of Table 3 presents the result when taking $\mu = 1$ (Model 1). γ is negative and significant, δ is positive and significant, and their sum is positive, which is totally consistent with the PFS model and also in line with those of GM (column 1). In Model

[15] For nomalization, the variance of u_{2i} is set to one.

2, we add a constant term to equation (4). The estimates of γ and δ are nearly the same as before, while the intercept is found to be positive and significant. Similar results are obtained when the sample size is increased to 1000 (see columns 6-7).

Model 3 further adds the political organization dummy to Model 2, allowing the intercept to differ for organized and unorganized industries. As column 5 indicates, the intercepts differ in sign: positive for organized industries and negative for unorganized ones. γ and δ are found to be insignificant. However, they both become significant when the sample size increases to 1000 (column 8). Overall, the results are consistent with the PFS model: γ is negative and significant, δ is positive and significant, and their sum is positive. These findings are robust to the inclusion of the constant term and the political organization dummy.

One small inconsistency is worth noting. In Model 3, it is found that the constant term is negative and the coefficient on the political organization dummy is positive. If the true model is equation (4) but we use $z_i^* = 1 + z_i$ instead of z_i, then the estimating equation can be expressed as follows:

$$\begin{aligned}\frac{t_i^* e_i}{1+t_i^*} &= \gamma\left(z_i^* - 1\right) + \delta I_i \left(z_i^* - 1\right) + \epsilon_i \\ &= -\gamma - \delta I_i + \gamma z_i^* + \delta I_i z_i^* + \epsilon_i.\end{aligned}$$

Because $\gamma < 0$ and $\delta > 0$, the constant term should be estimated to be positive and the coefficient on the political organization dummy to be negative. However, that is not the case either in our results or GM's and in this sense, one could argue that our and GM's results are not in line with the PFS model.

It is also noteworthy that there are some discrepancies between our and GM's findings. First, the constant term and the coefficient on the political organization dummy are significant in Model 3 (columns 5 and 8), while they are not in GM (column 2). GM's small sample size (107), which results in large standard errors, may explain this discrepancy. Another possible reason is that the simulated data and the actual data differ in the variations in the dependent variable and the import penetration ratio. More variation exists in the dependent variable in the data than in our simulated data, as we keep e constant. The data on the inverse import penetration ratio in GM is clustered away from origin, since US, a large country, has a low ratio of imports to domestic production in most industries. As a result, greater variation in the dependent variable is being explained by a small variation in the explanatory ones in GM, leading to the insignificance of the constant term in GM.

The second discrepancy is that adding the constant and dummy for organization affects their estimates of γ and δ less than ours. Why does this difference arise? If the true model is

$$\frac{t_i}{1+t_i} e_i = \beta_0 + \beta_I I_i + \gamma\left(1+z_i\right) + \delta I_i \left(1+z_i\right) + \epsilon_i,$$

then the omission of I_i will result in overestimate of δ (as occurs in Table 3) due to the positive correlation between I_i and $I_i\left(1+z_i\right)$. The correlation is smaller in the data (0.105) than the one in the simulation (0.325), which explains why our results change more than those of GM. The higher correlation in the simulated data is because of the uniform quota level; when quota is binding for a subindustry, its import is constant. We find it noteworthy that even in the simple setting, we obtain a significant and negative value for γ and a significant and positive value for δ despite allowing for different intercepts for organized and unorganized industries as required by the PFS model when $1+z_i$ is used instead of z_i!

5 Tariffs instead of Quotas

Although most of the empirical work estimating the PFS model use NTB's as proxies for tariffs, there are some notable exceptions such as McCalman (2004) who used Australian data on tariffs. In this section, we simulate a simple equilibrium model of trade with exogenously determined tariff levels, which has the same spirit as our equilibrium model with quotas. We solve the model and estimate the protection equation using the simulated data. As before, equations (8)-(10), and (11)-(13) define the demand, supply, and equilibrium for domestic goods and imports respectively.

The parameterization is the same as in the quota case except for the inclusion of the uniform tariff t in the import demand equation. We set a uniform import tolerance level $\hat{Q}$ for politically organized subindustries. We assume that if the equilibrium output for the foreign goods exceeds $\hat{Q}$, then government imposes an uniform tariff $t = 0.1$. Otherwise, the tariff is set to be 0. Let d_{ij}^t be the indicator that takes on the value of one if the tariff is positive. That is, $d_{ij}^t = 1$ if x_{ij}^{Me} exceeds $\hat{Q}$ and $d_{ij}^t = 0$ and $x_{ij}^M = x_{ij}^{Me}$, otherwise. For subindustries with positive tariffs, the demand equation becomes as follows:

$$\ln x_{ij}^{Md} = amd_1 + amd_2 \ln\left[(1+t)\, p_{ij}^M\right] + xmd_i + umd_{ij}.$$

Equilibrium under the tariff is computed by equalizing subindustry demand and supply. The output, $(1 + z_i)$, and political organization for each industry are computed by aggregating over subindustries, just as in the quota model. The industry level tariff is the simple average of the subindustry tariffs, $t_i = \sum_{j=1}^{n_i} t_{ij}/n_i$.

Generating data from the model, we estimate the following equation by OLS and 3SLS:

$$\frac{t_i}{1+t_i}\, |amd_2| = \beta_0 + \gamma \frac{(1+z_i)}{10000} + \delta I_i \frac{(1+z_i)}{10000} + u_i.$$

Table 4: Regression Results: A Model with Tariffs

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	3SLS1	3SLS2	OLS	3SLS1	3SLS2
β_0	0.051	0.049	0.049	0.049	0.046	0.048
	(0.004)	(0.006)	(0.004)	(0.002)	(0.003)	(0.002)
γ	-4.927	-16.04	-7.753	-4.504	-16.72	-6.792
	(1.310)	(3.84)	(2.02)	(0.585)	(1.90)	(0.948)
δ	5.482	24.34	11.16	5.399	26.77	13.05
	(1.540)	(5.40)	(2.68)	(0.669)	(1.62)	(1.22)
α	2285.4	427.4	984.9	1872.3	373.0	765.5
α_L	0.939	0.659	0.675	0.844	0.625	0.674
Nobs	200	200	200	1000	1000	1000

Note: Standard errors are in parentheses. In columns (1)-(3), the results are the average of 10 simulation/estimation exercises.

Estimation results are presented in Table 4. The results in columns 1-3 are the average of 10 simulation/estimation exercises with the sample size of 200, while those in columns 4-6 are based on the sample size of 1000. Estimates are fully consistent with the PFS model, although the simple tariff model from which we generate data is quite different from the PFS one. Qualitatively similar results (available upon request) are obtained when we estimate the model by the maximim likelihood.

Our findings have an important implication for a puzzle in the literature: α is estimated to be "too high" (i.e., the estimated weight on contributions is "too low"). Note that as Table 4 shows, our estimates of α are high just as those in the literature. This is also observed in the model with quotas (see Tables 2-3). As parameters in our models do not have any strict PFS interpretation, our results suggest that the strict PFS interpretation being put on the parameter estimates in previous studies may be misplaced; the supposedly low value for the weight on contributions may be merely a misinterpretation of the parameter estimates.

6 Some Robustness Checks

It is possible that our main results are driven by the choice of parametrization we made. To see how our results are affected by changes in model specification and parameterization, we conduct some robustness checks by relaxing several assumptions in the simulation.[16] First, we relax the assumption that the import demand elasticity, amd_2, is constant across industries; we now assume that it is normally distributed with mean 1.5027 and standard error 0.3705, which are the sample mean and standard deviation of the import price elasticities in the GB's data. We simulate the SP model and then estimate equation (14) with the same instruments as those in earlier simulation/estimation exercises. The OLS, 3SLS1, and 3SLS2 results are presented in columns 1-3 of Table 5, respectively. Regardless of the estimation methods, all the coefficient estimates are consistent with the PFS hypothesis, i.e., $\gamma < 0$, $\delta > 0$, and $\gamma + \delta > 0$. Note that the estimates of δ are lower than those obtained for the constant import price elasticity (see columns 6-8 of Table 2). This is as expected because higher elasticity leads to higher variation of imports and together with the uniform quota level, would result in higher probability of protection, which would generate negative correlation between z/e and the NTB coverage ratio for politically organized industries.

We further allow for correlation between the home goods demand and imports which we observe in the data. Specifically, we set correlation between home goods demand shock and import demand shock to be 0.636.[17] As columns 4-6 of Table 5 show, our main results do not change; all of the results are in support of the PFS hypothesis. We also checked for the robustness of our results to changes in the following parameters; we either increased or decreased a subset of the parameters by 20% (home goods demand, home goods supply,

[16] We are grateful to an anonymous referee for suggesting us to do the robustness checks done in this section.

[17] This number is taken from the correlation between log consumption and log import in data used by Facchini et al. (2006).

Table 5: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	3SLS1	3SLS2	OLS	3SLS1	3SLS2
β_0	0.191	0.189	0.186	0.175	0.138	0.164
	(0.007)	(0.012)	(0.008)	(0.007)	(0.011)	(0.008)
γ	-12.5	-74.1	-30.8	-14.8	-58.6	-39.7
	(2.65)	(8.74)	(4.40)	(3.67)	(10.3)	(5.88)
δ	17.8	113.9	49.3	37.0	132.0	82.7
	(3.05)	(11.9)	(5.74)	(3.94)	(13.1)	(7.11)
α	562.6	87.15	202.21	269.66	75.31	120.44
α_L	0.7049	0.6508	0.6236	0.4006	0.4435	0.4797
Nobs	1000	1000	1000	1000	1000	1000

Note: Standard errors are in parentheses. In columns (1)-(3), the import demand elasticity is assumed to be normally distributed with mean 1.5027 and standard error 0.3705. In columns (4)-(6), home goods demand shock and import demand shock are allowed to be correlated as well (the correlation coefficient is 0.636).

foreign goods demand and foreign goods supply parameters). In all those perturbations, the parameter estimates were consistent with the PFS hypothesis. [18]

7 Why the Simulation Results?

Why is it that we spuriously estimate a protection for sale effect from the simulated data? In this section, we try to explain the reason by using an even simpler model of protection, which does not have any aggregation over subindustries. Suppose that the demand for and supply of home goods have no random component:

$$\ln X_i^{Hd} = ahd_1 + ahd_2 \ln p_i^H,$$

$$\ln X_i^{Hs} = ahs_1 + ahs_2 \ln p_i^H.$$

[18] Those results are available upon request.

Then, the home goods equilibrium quantity is:

$$\ln X_i^H = \frac{ahd_2 ahs_1 - ahs_2 ahd_1}{ahd_2 - ahs_2}.$$

We choose the parameters so as to set the home goods equilibrium quantity to unity. That is,

$$\ln X_i^H = \frac{ahd_2 ahs_1 - ahs_2 ahd_1}{ahd_2 - ahs_2} = 0.$$

For imported goods in the same industry, however, demand and supply have random components. Let

$$\ln X_i^{Md} = amd_1 + amd_2 \ln p_i^H + xmd_i,$$
$$\ln X_i^{Ms} = ams_1 + ams_2 \ln p_i^H + xms_i.$$

Then, the equilibrium of the foreign goods market is

$$\ln X_i^M = \frac{amd_2 ams_1 - ams_2 amd_1}{amd_2 - ams_2} + \frac{amd_2 xms_i - ams_2 xmd_i}{amd_2 - ams_2}.$$

The parameters are set in such a way that the foreign goods equilibrium is as follows:

$$\ln X_i^M = -1.0 + 2.0 U_i,$$

where U_i is assumed to be uniformly distributed on [0,1]. This gives the desired level of imports. Also, we set the uniform quota level, $\hat{Q} = 1$, so $\ln \hat{Q} = 0$. As before, political organization is random and there is a .5 chance of being organized. Protection occurs if the quota is binding and the industry is organized. In this setting, if $\ln X_i^M \geq 0$ and the industry is organized, the coverage ratio (C_i) and the protection measure $(C_i/(1 + C_i))$

are 1 and 0.5, respectively; otherwise, they are both zero.

Since the probability of being organized is .5, with a large enough number of industries, half of them will be organized and half will not. For the half that are not organized, the consumption to import ratio is:

$$\frac{X_i^H + X_i^M}{X_i^M} = 1 + z_i = 1 + \frac{1}{e^{(-1.0+2.0U_i)}}.$$

For the other half of the industries, which are politically organized, it is:

$$\begin{aligned} 1 + z_i &= 1 + \frac{1}{e^{(-1.0+2.0U_i)}} \text{ if } \ln X_i^M < 0, \\ &= 2 \text{ otherwise.} \end{aligned}$$

Now consider that we have drawn 2000 industries. For a large enough sample size, in any realization, roughly half will be organized. To illustrate the intuition, we take exactly half to be organized. Number the non-organized (organized) industries by integers between 1 (1001) and 1000 (2000) with a higher index given to the industry with a larger U_i. Only industries with an index above 1000 will ever get protection. As the number of draws gets large enough, we would expect to see a uniform empirical distribution of the realizations of U_i. To capture this in figures below, we place one firm at each integer. That is, we assume that industry i has $U_i = i/1000$ for $i = 1, ..., 1000$, and $U_i = (i - 1000)/1000$ for $i = 1001, ..., 2000$. Industries with an index higher than or equal to 1500 will have the quota invoked and be binding while industries with an index below the cutoff, while organized, never have the quota invoked.

Figure 1 plots the U_i and the import quantity. Notice that for industry $i = 1001, ..., 2000$, which are politically organized, the quota binds and import quantity equals the quota when

U_i is large (industries 1500 to 2000).

(Figure 1 in here)

Figure 2 plots the protection measure, $C_i/\left(1+C_i\right)$. It is 0.5 for industries with index 1500 to 2000 which are politically organized and whose quota is binding.

(Figure 2 in here)

Figure 3 plots 1+ z_i. As we can see, this is high for industries with small imports and low for those with large imports. It is constant for industries with index 1500 to 2000 because of the binding quota.

(Figure 3 in here)

We next plot the $1+z_i$ times the political organization dummy in Figure 4. Notice that for industries 1 to 1000, it is zero because they are never politically organized.

(Figure 4 in here)

Now we fit the protection measure in Figure 2 by using $(1+z_i)$ (Figure 3) and $I_i(1+z_i)$ (Figure 4) by OLS. We obtain

$$\left(\widehat{\frac{C_i}{1+C_i}}\right) = \underset{(0.0160)}{0.3728} - \underset{(0.0072)}{0.1571}(1+z_i) + \underset{(0.0035)}{0.0921} I_i\left(1+z_i\right),$$

where standard errors are in parentheses. Note that $\widehat{\gamma}$ and $\widehat{\delta}$ have opposite signs as in the PFS model. There seems to be a positive correlation between protection and $(1+z_i)$ for politically organized industries but a negative one for unorganized industries. This is what the regression is picking up.

We can confirm this insight by looking at the regression results from a different angle, i.e., by using the partitioned regression. Let RIP_i be the component of $(1 + z_i)$ that is orthogonal to $I_i(1 + z_i)$. It is obtained by regressing $(1 + z_i)$ on the constant term and $I_i(1 + z_i)$ and then taking the residual. The thin line in Figure 5 plots this orthogonal component.

(Figure 5 in here)

Due to the properties of the partitioned regression, the coefficients of the OLS regression of the protection measure on the orthogonal component gives the coefficient on $(1 + z_i)$ back. As can be seen from the figure, the orthogonal component of $(1 + z_i)$ is negatively correlated with the protection measure, which is the reason for the negative coefficient of the $(1 + z_i)$ in the original OLS.

Similarly, let $RIIP_i$ be the component of $I_i(1 + z_i)$ that is orthogonal to $(1 + z_i)$. It is dipicted as the thin line in Figure 6.

(Figure 6 in here)

As is clear from the figure, the orthogonal component of $I_i(1 + z_i)$ is positively correlated with the protection measure. This accounts for the positive coefficient of $I_i(1 + z_i)$ in the original regression.

The qualitative aspects of the results do not change when we use IV estimation with U_i and U_i^2 as instruments. γ and δ are estimated to be -0.610 and 1.008 along with the standard errors of 0.280 and 0.645, respectively. In this case, not only are the signs right but $\widehat{\gamma} + \widehat{\delta} > 0$, which is even more consistent with the PFS model.

Conventional studies in trade estimating the political economy effects use NTBs as a proxy for tariff protection measures, even though NTBs could be better interpreted as quotas. Our results suggest that the real reason behind evidence in support of the PFS model could be the difference between the quota being binding and non-binding. That is, $\delta I_i (1 + z_i)$ with $\delta > 0$ fits well for the industries under quota (1500 to 2000) and industries that are not politically organized (1 to 1000), but does not fit well for industries that are politically organized but not under quota (1001 to 1499). On the other hand, $\gamma(1+z_i)$ with $\gamma < 0$ fits well for politically organized industries since those with high equilibrium imports face binding quotas, but fits very poorly for those that are not politically organized. Hence, it is natural that combining both would give the best fit, and these results correspond to the signs obtained by GM, GB, and others. Similar interpretations can be offered for the tariff version.

8 Which Model is More Plausible?

We have shown that (1) data simulated from the SP model are reasonably close to actual data and (2) estimation of the protection equation using the simulated data provides coefficients that are consistent with those predicted by the PFS framework. This suggests that numerous past findings in favor of the PFS model are also in favor of the SP model. Clearly, estimation of the PFS protection equation is insufficient to conclude the validity of the PFS model.

Given our results along with past findings in the literature, it is important to ask which model is (more) correct.[19] Recently, Imai et al. (2008) present some quantile regression and IV quantile regression results where the dependent variable is the protection measure and the RHS variable is the inverse import penetration ratio. They argue that the signs of the estimated coefficients at high quantiles are not consistent with those predicted by the

[19] We thank the associate editor for raising the issue.

PFS model. On the other hand, in this paper, we show that the coefficient estimates are indeed in line with the prediction of the SP model.

Imai et al. (2008) recently propose a new approach to testing the PFS model. An important feature of their approach is that it does not require the classification of industries into organized and unorganized ones[20]. Their approach relies heavily on observables and 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, one should expect a positive relationship between the inverse import penetration ratio and the protection measure. Imai et al. (2008) 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). Their proposition (Proposition 1) essentially states that in the quantile regression of $t/(1+t)$ on z/e, the coefficient on z/e should be close to $(\gamma+\delta)$ at the quantile close to 1. To empirically examine this, they use quantile regression (Koenker and Bassett, 1978) and estimate the following equation:

$$Q_T(\tau|Z) = \alpha(\tau) + \beta(\tau) Z/10000, \tag{15}$$

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 to 1 from below. Using the data from Gawande and Bandyopadhyay (2000), Imai et al. (2008) find that the null hypothesis of $\beta(\tau) = 0$ cannot be rejected at high quantiles (in fact, all quantiles) in favor of the one-sided alternative

[20] Imai et al. (2008) also argue that the conventional classificaton of political organization by Goldberg and Magee (1999) and Gawande and Bandyopadhyay (2000) are inconsistent with the PFS model, and relying on those arbitrary classification could result in wrong coefficient estimates of the protection equation.

of $\beta(\tau) > 0$. Moreover, the point estimates indicate that contrary to the PFS prediction, $(\gamma + \delta)$ are all negative at high quantiles and decrease as goes from 0.4 to 0.9. The results do not provide any evidence favoring the PFS model.

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. Imai et al. (2008) therefore allow for the potential endogeneity of Z. They formally show that even in the presence of this endogeneity, the main prediction of the PFS model in terms of their quantile approach does not change. See Imai et al. (2008) for the relevant proposition (Proposition 2). To test the prediction in the presence of possible endogeneity of Z, they estimate the following equation by using IV quantile regression (Chernozhukov and Hansen, 2004a; 2004b; 2006):

$$P(T \leq \alpha(\tau) + \beta(\tau) Z/10000|W) = \tau, \tag{16}$$

where W is a set of instrumental variables. As in the quantile regression, Imai et al. (2008) find that the null hypothesis of $\beta(\tau) = 0$ in favor of the one-sided alternative cannot be rejected. The point estimates are not favorable for the PFS model, either; even after correcting for the endogeneity of Z, the estimate of at the highest quantile is not positive as required by the PFS model. Imai et al. (2008) also show the robustness of their findings; regardless of which instrument they use and whether they 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, zero at low quantiles and negative at any other quantiles, which is inconsistent with the PFS's prediction.

Next, we examine the empirical validity of the SP model. We conduct the following exercise: using simulated data from the SP model, we estimate equations (15) and (16). We ask whether the parameter estimates from the simulated data resemble those reported in Imai et al. (2008). If the SP model is valid, then the patterns exhibited in the former are expected to be similar to those in the latter. The results are presented in Table 6. The coefficients on Z are found to be zero at lower quantiles and thereafter negative, which is consistent with those in Imai et al. (2008).

Table 6: Quantile Regression Results

	SP		Imai et al. (2008)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.000)	0.000 (0.000)	0.000 (0.004)	0.000 (0.056)
0.2	0.000 (0.000)	0.000 (0.000)	0.000 (0.005)	0.000 (0.079)
0.3	0.000 (0.000)	0.000 (0.223)	0.000 (0.006)	0.000 (0.091)
0.4	0.003 (0.029)	0.020 (2.639)	0.000 (0.006)	0.000 (0.097)
0.5	0.139 (0.055)	−3.190 (5.087)	0.002 (0.006)	−0.003 (0.099)
0.6	0.311 (0.049)	−6.320 (6.061)	0.028 (0.006)	−0.046 (0.098)
0.7	0.399 (0.023)	−4.332 (5.466)	0.077 (0.010)	−0.126 (0.095)
0.8	0.443 (0.013)	−1.374 (3.928)	0.157 (0.026)	−0.258 (0.094)
0.9	0.498 (0.012)	−1.048 (2.569)	0.308 (0.040)	−0.505 (0.089)

Note: We estimated on the 250 simulated industries. The results are the averates of 100 simulation/estimation exercises. Standard errors are in parenthese. The results in the last column is from Imai et al. (2008) Table 2, column 5.

Similar results are obtained for the IV quantile regression (Table 7). It is also noteworthy that the size of β's is by and large similar with that obtained by Imai et al. (2008) when they use three instruments highly correlated with Z.

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. A similar mechanism has been considered in the literature. Findlay and Wellisz (1982), for example, argued that an increase in import in an industry triggers lobbying efforts which result in higher protection. That is, like in the SP model, a negative

Table 7: IV Quantile Regression Results

	SP		Imai et al. (2008)	
τ (quantile)	$\alpha(\tau)$	$\beta(\tau)$	$\alpha(\tau)$	$\beta(\tau)$
0.1	0.000 (0.000)	0.000 (0.000)	0.000 (0.004)	0.000 (0.000)
0.2	0.000 (0.000)	0.000 (0.000)	0.000 (0.009)	−0.000 (0.000)
0.3	0.000 (0.001)	0.000 (0.196)	−0.018 (0.018)	−17.81 (23.37)
0.4	0.011 (0.046)	−1.105 (5.373)	−0.021 (0.019)	−7.160 (12.05)
0.5	0.137 (0.060)	−3.965 (6.421)	−0.019 (0.026)	−7.180 (14.49)
0.6	0.308 (0.049)	−7.575 (6.698)	−0.033 (0.026)	−8.500 (13.08)
0.7	0.405 (0.024)	−6.430 (6.048)	−0.043 (0.027)	−3.890 (6.624)
0.8	0.445 (0.013)	−2.640 (4.877)	−0.029 (0.049)	−4.970 (7.401)
0.9	0.495 (0.014)	−1.135 (3.529)	0.079 (0.070)	−5.780 (7.446)

Note: Standard errors are in parenthese. The results in the last column is from Imai et al. (2008) Table 3, column (11).

relationship is predicted between the change in inverse import penetration ratio and the NTB coverage ratio. This prediction is indeed borne out in the data; Trefler (1993) found that the relationship between import penetration ratio and protection is insignificant but a change in import penetration has a positive and significant effect on protection.

9 Conclusion

In this paper, we showed that the usual tests of the PFS model are actually also consistent with a simpler model, which we call "Surge Protection" model, where protection tends to occur when imports surge and the industry is organized. We then showed that the Surge Protection model is more consistent with the data when we look at the relationship between the level of protection at various conditional quantiles and inverse penetration ratio divided by the import demand elasticity. The relationship is found to be negative at higher quantiles in the SP model, which is consistent with that in the data, whereas it is predicted to be positive by the PFS model. An important implication our results provide is that there is no puzzle regarding the high weight on welfare generated by the "tests" of the PFS model, since the Surge Protection model does not allow γ and δ to be used to construct a weight on welfare placed by the government.

References

[1] Bernheim, B.D., Whinton, M., 1986. Menu auctions, resource allocation, and economic influence. Quarterly Journal of Economics 101, 1-31.

[2] Bombardini, M., 2004. Firm heterogeneity and lobby participation. Mimeo, MIT.

[3] Chernozhukov, V., Hansen, C., 2004a. Instrumental Variables Quantile Regression. Mimeo, MIT.

[4] Chernozhukov, V., Hansen, C., 2004b. The effects of 401(k) participation on the wealth distribution: an instrumental quantile regression analysis. Review of Economics and Statistics 86, 735-751.

[5] Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effects model. Journal of Econometrics 132, 491-525

[6] Eicher, T., Osang, T., 2002. Protection for sale: an empirical estimation: comment. American Economic Review 92, 1702-1710.

[7] Facchini, G., Van Biesbroeck, J., Willmann, G., 2006. Protection for sale with imperfect rent capturing, Canadian Journal of Economics 39, 845-873.

[8] Findlay, R., Wellisz, S., 1982. Endogenous tariffs and the political economy of trade restrictions and welfare. In: Bhagwati, J. (Ed.), Import Competition and Response. University of Chicago Press.

[9] Gawande, K., 1998. Comparing theories of endogenous protection: Bayesian comparison of Tobit models using Gibbs sampling output. Review of Economics and Statistics 80, 128-140.

[10] Gawande, K., Bandyopadhyay, U., 2000. Is protection for sale? evidence on the Grossman-Helpman theory of endogenous protection. Review of Economics and Statistics 82, 139-152.

[11] Gawande, K., Krishna, P., 2003. The political economy of trade policy: empirical approaches. In: Harrigan, J., Choi, E.K. (Eds.), The Handbook of International Trade, Basil Blackwell.

[12] Gawande, K., Krishna, P., 2004. Foreign lobbies and US trade policy. NBER Working Paper No. 10205.

[13] Gawande, K., Krishna, P., 2005. Lobbying competition over US trade policy. NBER Working Paper No. 11371. Also in Handbook of International Trade, Basil Blackwell.

[14] Goldberg, P., Maggi, G., 1999. Protection for sale: an empirical investigation. American Economic Review 89, 1135-1155.

[15] Grossman, G., Helpman, E., 1994. Protection for sale. American Economic Review 84, 833-850.

[16] Imai, S., Katayama, H., Krishna, K., 2008. A quantile-based test of protection for sale model. Mimeo, Queen's University.

[17] Kee, H.L., Olarreaga, M., Silva, P., 2005. Market access for sale. Journal of Development Economics, forthcoming.

[18] Koenker, R.W., Bassett, G.W., 1978. Regression Quantiles, Econometrica 46, 22-50.

[19] Kelejian, H., 1971. Two-stage least squares and econometric systems linear in parameters but nonlinear in the endogenous variables. Journal of the American Statistical Association 66, 373-374.

[20] Keller, W., 1998. Are international R&D spillovers trade-related? analyzing spillovers among randomly matched trade partners. European Economic Review 42, 1469-1481.

[21] Matschke, X., Sherlund, S., 2006. Do labor issues matter in the determination of U.S. trade policy? an empirical reevaluation. American Economic Review 96, 405-421.

[22] McCalman, P., 2004. Protection for sale and trade liberalization: an empirical investigation. Review of International Economics 12, 81-94.

[23] Mitra, D., Thomakos, D., Ulubasoglu, M., 2002. Protection for sale in a developing country: democracy versus dictatorship. Review of Economics and Statistics 84, 497-508.

[24] Rodrik, D., 1995. Political economy of trade policy. In: Grossman, G., Rogoff, K. (Eds.), Handbook of International Economics, vol. 3. North-Holland, Chapter 28.

[25] Shiells, C.R., Stern, R.M., Deardorff, A.V., 1986. Estimates of the elasticities of substitution between imports and home goods for the United States. Weltwirtschaftliches Archiv 122, 497-519.

[26] Trefler, D., 1993. Trade liberalization and the theory of endogenous protection. Journal of Political Economy 101, 138-160.

Figure 1: Import Shock U(i) and Import Quantity

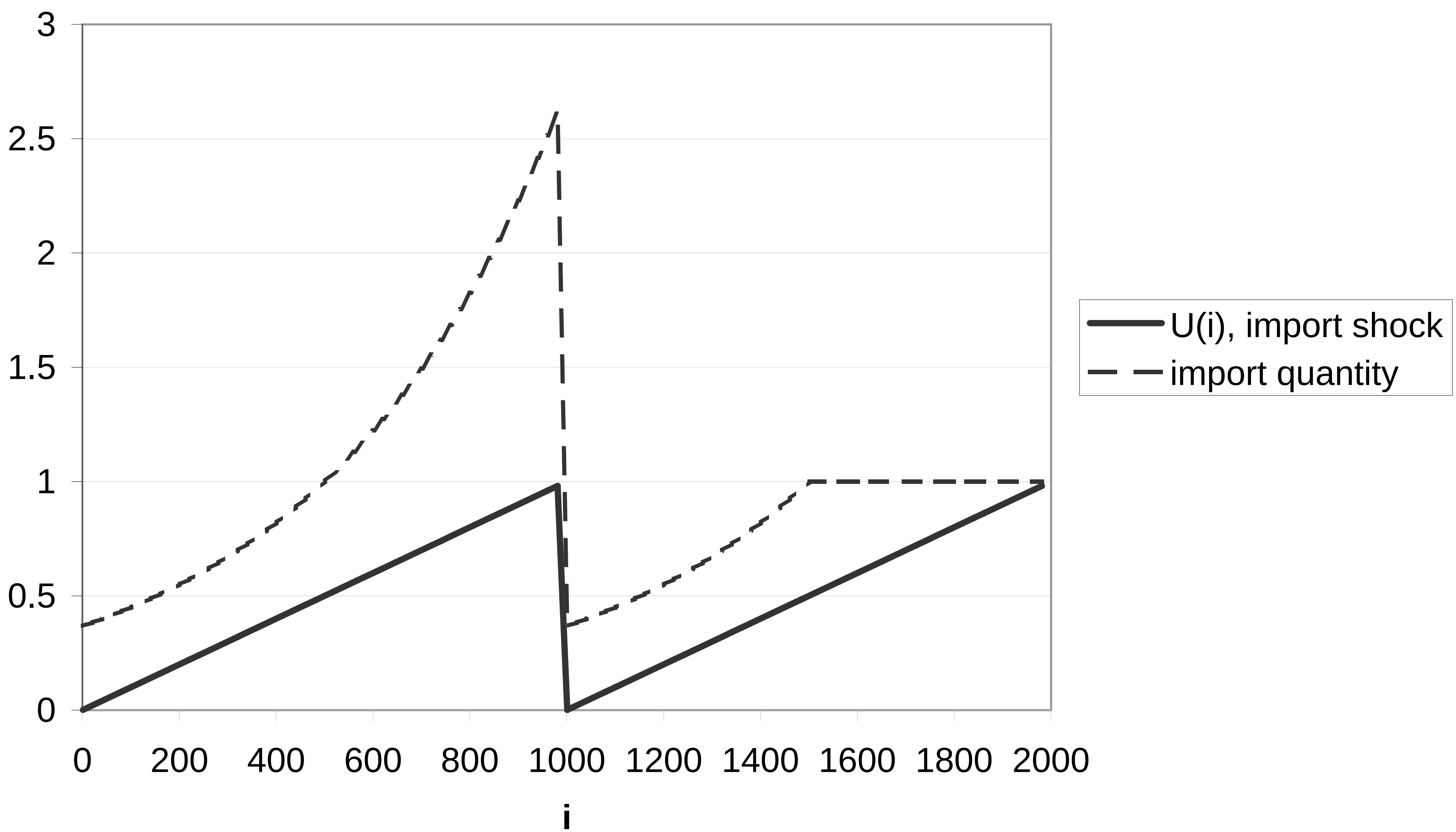

Figure 2: Protection Measure: Simulation

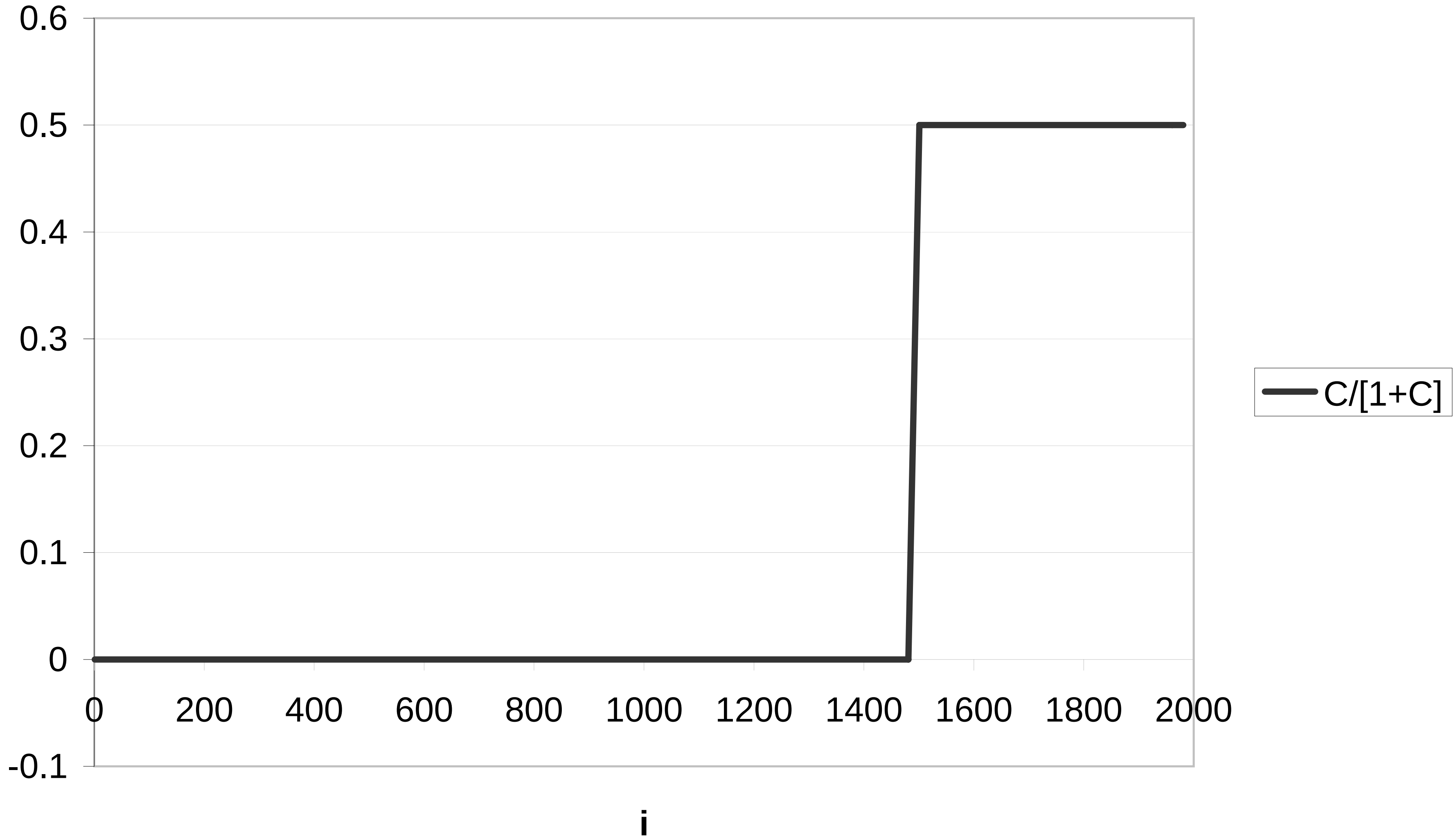

Figure 3: Inverse Import Penetration Ratio

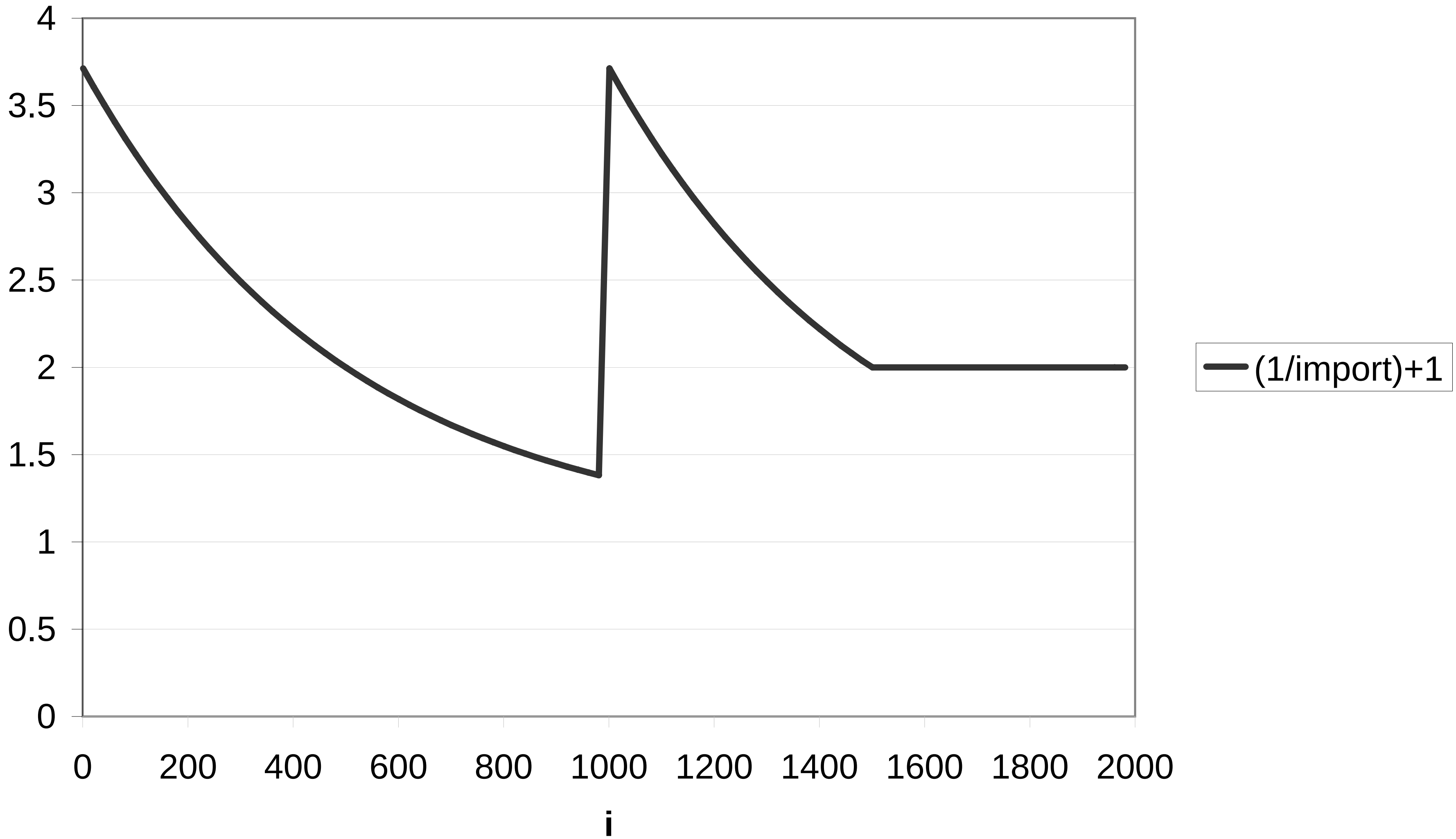

Figure 4: Politically Organized Inverse Import Penetration Ratio

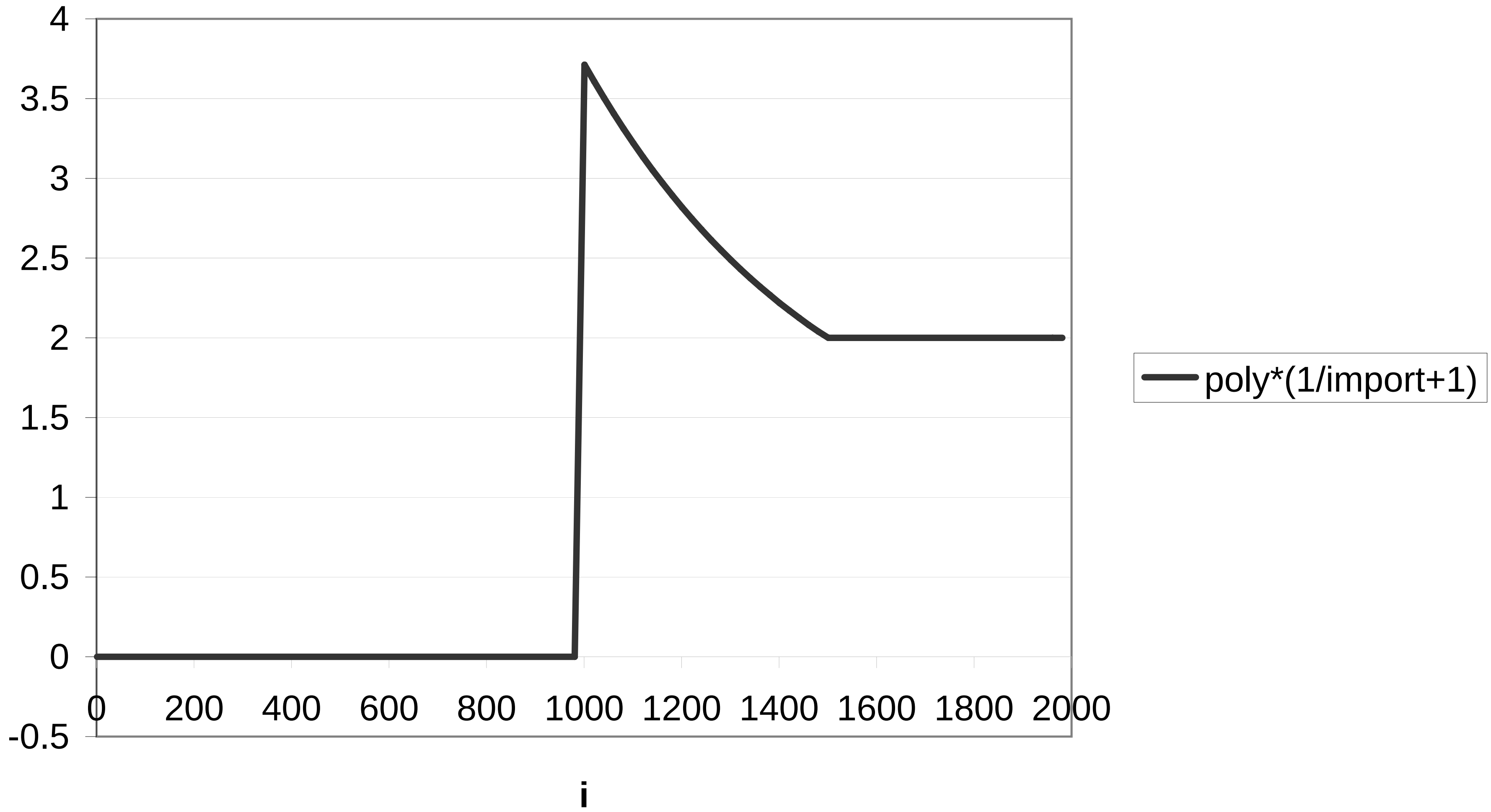

Figure 5: Orthogonal Component of Inverse Import Penetration Ratio

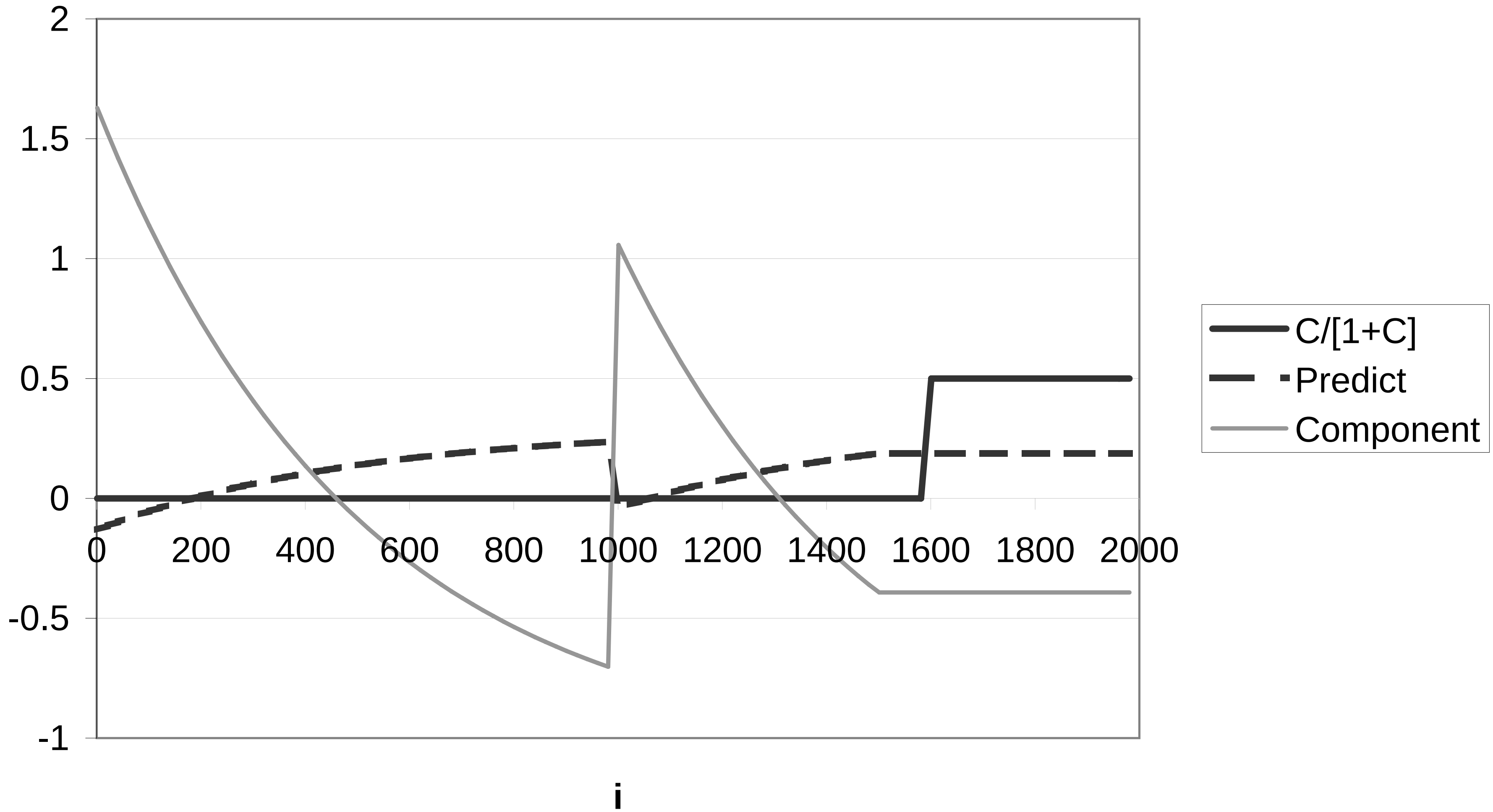

Figure 6: Orthogonal Component of Politically Organized Inverse Import Penetration Ratio