
Susumu Imai        Hajime Katayama
Concordia University and CIREQ        University of Sydney

Kala Krishna
Pennsylvania State University and NBER

July 2005

Abstract

This paper asks whether the results obtained from using the standard approach to testing the influential Grossman and Helpman “protection for sale (PFS)” model of political economy might arise from alternative specifications and argues that the standard approach may be less of a test than previously thought.

Keywords: Common agency. Political economy, Protection for sale..
JEL: F13, D72, F17.
1. Introduction

Over the past decade, the Grossman and Helpman (1994) model of “Protection for Sale” (PFS) has become the most influential one, both theoretically and empirically, in the political economy of trade. Goldberg and Maggi (1999), Gawande and Bandyopadhyay (2000) were the first papers that took the model to the data and found estimates consistent with the model. More recently, researchers have extended the original PFS model in various directions. For example, Facchini et al. (2004) constructed and estimated the quota version of the PFS model instead of the original tariff version. This is important since in the U.S. data, the non-tariff barrier coverage ratio is commonly used as the proxy for tariff protection. Bombardini (2004) incorporates firm size into the protection equation\(^1\). Krishna and Gawande (2004) incorporate foreign lobbies as well as domestic ones while Krishna and Gawande (2005) add lobbying of both upstream and downstream producers. The models are then tested using data on contributions from foreign firms and final and intermediate good producers respectively.

While such extensions seem to provide much evidence in support of the PFS model, there are a number of reasons for concern.\(^2\) First, there remain some deep problems in terms of what the PFS model (and in fact many others) explain. A robust empirical regularity seems to be that NTB coverage is positively and significantly related to the change in import penetration.\(^3\) However, the PFS model (as well as that of Findlay and Wellisz (1982) and Hillman (1982)) yields the counter intuitive result that protection is higher the lower the import penetration.

\(^1\)However, there are some problems with pinning down a unique equilibrium in her model.

\(^2\)Most of these concerns have been voiced in the literature, see Rodrik (1995), Helpman (1995), and Gawande (2003) for more on much of what follows.

\(^3\)See Trefler (1993).
Sectors with a lower import penetration would tend to be those where a comparative advantage exists and this does not fare well with intuition and data which suggest exactly the opposite: that it is the sectors where comparative advantage is low that imports are a threat. Moreover, estimates of the weight on welfare relative to contributions is very high or that political economy factors matter little. Given that contributions are small relative to their effects on firm profits and welfare, one would expect a reasonably high weight on contributions relative to welfare. However, contributions in Goldberg and Maggi (1999) are only used to indicate whether an industry is organized or not. They are used to see if lobbying expenditure follows predicted patterns in Gawande and Bandyopadhyay (2000).

Second, other than a few exceptions, most of the empirical literature does not try and choose between various models. Since many of the models predict similar things in some dimensions, it is vital to (a) pick out those aspects where they differ and focus on testing for these predictions rather the ones where they do not differ and (b) test all predictions of a model, not a select few. Not doing this leaves consumers of the literature reading a guide written by a reviewer with ADD (Attention Deficit Disorder).

Third, as in any empirical enterprise, the data is far from perfect. To begin

---

4 Goldberg and Maggi (1999) argue that there is no inconsistency between their results and these as protection enters additively in these models and interactively with political organization in theirs. See Gawande and Krishna (2003) for a fine survey of the empirical work.

5 See Rodrik (1995) whose survey was the first to bring this out and is well worth reading even today.

6 Ansolobehere et al. (2002) argue forcefully against thinking of contributions as buying policy.

7 Maybe the problem that our intuition on where protection occurs is basically dynamic, while most such models are static. One exception is the work of Brainard and Verdier (1993) which explains the persistence of protection by pointing out that protection today raises production relative to imports and this in turn raises protection tomorrow.

8 Prominent examples include Eicher and Osang (2002), Gawande (1998). Of course, the classic work of Baldwin (1985) also tries to do so but in less tightly structured settings.
with contributions data is often not available outside the US. In addition the elasticity estimates commonly used, those of Sheills, Deardorff and Stern are from the 80’s and at the three digit level of aggregation. Moreover, the estimates do not seem good enough for this purpose: half the estimates are of the wrong sign or insignificant. More recent estimates at a disaggregated level need to be used since testing political economy models, in particular, should be done at as disaggregated a level as possible.

Fourth, replication of the major results in the area turns out to be extremely hard to do. The data used by Goldberg and Maggi is not available as it was lost. Gawande is extremely generous with his data. Unfortunately, neither we nor Bombardini (2004) could exactly replicate his results quantitatively, though we were able to do so qualitatively: more on this below. In personal communication, Gawande pointed out that the exact set of instruments used was important in getting the results, which is, perhaps, cause for concern. All of this makes one wonder how much of what researchers are getting is due to choosing a set of regression results that validates existing empirical work? How much of it is due to the same forces at work in a variety of setups that result in estimates that look like support for the model? How much of the work supposedly supporting the PFS model actually is testing for results that are common in a variety of different models, not just the PFS one? For example, to what extent should firm size effects

---

9A way around this, using an iterative procedure, is proposed by Cadot et al. (2005). As a by product they find that their estimates of the weight on social welfare are lower than that on contributions! However, how well their procedure, performs is not as yet clear.

10They estimate import elasticity industry by industry by using OLS or 2SLS. Obviously, OLS is subject to endogeneity and measurement error bias. 2SLS as executed by them has is problematic because the industry by industry sample size is very small and 2SLS has potentially serious finite sample bias. Furthermore, Shiells et. al. control for tariffs in their elasticity estimation but not for the non-tariff barrier. Hence, if researchers use the Shiells et. al. estimates, the reverse causality from non-tariff barrier to the import elasticity, which could arise with aggregation in the industry data, cannot be controled for.
matter as in Bambardini (2004) or protection on being lower when there are organized downstream users of the industry’s output be seen as a validation of the PFS model? Would similar predictions not arise in other models? All of this makes one suspect that the model is, perhaps, not being subjected to the right kind of test in much of this work.

These concerns should not be taken negatively: these are hard questions to tackle. Rather, they should be taken as an attempt to refocus attention on the key issues and so guide future work.

A key insight provided by the PFS model, which differs from that of other models, and so is a vital one to test, and seems borne out by the data, is the prediction that organized industries have one sign for a key coefficient and unorganized ones have another. In this paper we ask whether this (compelling) result could be illusionary. We provide a small contribution to the literature by showing how results like this could arise even in the absence of the kind of behavior posited in the PFS literature. We simulate a simple equilibrium model of domestic consumption and imports, where imports in the politically organized sector are subject to an exogenously and uniformly set quota. There are no quotas if a sector is not organized. Political organization is set exogenously and randomly. Obviously, in this simple model of quotas, there is no protection for sale effect. Parameters are set so the simulated data roughly match the basic statistics of the actual data. Then, we estimate the protection equation on the artificial data following the procedures by Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000). We obtain coefficient estimates that are consistent with the protection for sale paradigm though there is none in the model! We then explain where our estimates are coming from.

The paper proceeds as follows. The PFS model is laid out in the next section. Section 3 then develops a simple model of imports with no lobbying which we
calibrate to broadly match the data. We then generate data from it. Section 4 then runs the standard regressions on the simulated data and shows that the standard results are obtained despite the absence of any PFS effects. Section 5 then explains why this is happening. Section 6 suggests and runs some further robustness tests using Gawande’s data. Section 7 concludes.

2. The PFS Model and Its Estimation

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 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 to the Some sectors 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 $\alpha$ 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). Governments then choose what to do to maximize their own objective function. In this way, the government is the common agent all principals (orga-
nized lobbies) are trying to influence. Such games are known to have a continuum of equilibria.\textsuperscript{11} By restricting agents to bids that are “truthful”, so that their bids have the same curvature as their welfare, a unique equilibrium is obtained.\textsuperscript{12} The equilibrium outcome, thus, is as if the government was maximizing weighted social welfare with a greater weight on the welfare of organized sectors. Thus, equilibrium tariffs can be found by maximizing

\[ G(p) = \alpha W(p) + \sum_{j \in J_0} W_j(p), \]

where \( J_0 \) is the set of politically organized firms and the welfare of agents in sector \( j \) is

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

where \( \pi_j(p_j) \) is producer surplus in sector \( j \), \( l_j \) is labor employed in sector \( j \), wage is unity, \( \frac{N_j}{N} \) is the share of workers employed in the \( j \) sector while \( T(p) + S(P) \) 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.

\textsuperscript{11}Given the bids of all other lobbies, each lobby wants a particular outcome to occur, namely, the one where it obtains the greatest benefit relative to 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.

\textsuperscript{12}For a detailed discussion of this concept, see Bernheim and Whinston (1986).
Differentiating $W_i(p)$ with respect to $p_j$ gives\(^\text{13}\)

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

where so $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, $\alpha_j$ is the share of labor employed in sector $j$, $m'_j(p_j)$ is the derivative of the demand for imports, and $x_j(p_j) = \pi'_j(p_j)$ denotes supply of sector $j$. 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

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

Now $\sum_{i \in J_0} \alpha_i = \alpha_L$, the employment share of organized sectors and $\sum_{i \in J_0} \delta_{ij} = I_j$ is unity if $j$ is organized and zero otherwise. Thus, the above is the same as

$$x_j(p_j)(I_i - \alpha_L) + (p_j - p_j^*)m'_j(p_j)(a + \alpha_L) = 0.$$ 

Using the fact that $(p_j - p_j^*) = (1 + t_j)p_j^*$, the above equation can be rewritten as

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

\(^{13}\)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 $j$. 

8
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 positive. Moreover, the coefficients on \( \frac{z_j}{e_j} \) and \( I_j \frac{z_j}{e_j} \) can be used to infer the weight on welfare placed by government.\(^{14}\)

Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000) add an error term to the above model to permit estimation:

\[
\frac{t_j}{1 + t_j} = \gamma \frac{x_j}{m_j e_j} + \delta I_j \frac{x_j}{m_j e_j} + \varepsilon_j.
\]

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. Both Gawande and Bandyopadhyay (2000) and Goldberg and Maggi (1999) used the coverage ratios for non-tariff barriers as \( t_j \) instead of the tariff itself. Gawande and Bandyopadhyay (2000) estimated a variant of equation (2.1) together with the other equations which determine the political contribution and the inverse import penetration ratio. Their protection for sale equation also accounts for tariffs on intermediate goods and adds as explanatory variables the tariff and NTBs on intermediates goods used by the industry. As shown in Grossman and Helpman (1994), protection for the final good is increasing in that for intermediate inputs used. To consistently estimate the above equation (since the inverse import penetration ratio and the import elasticity could be endogenous), they used a nonlinear IV estimation technique proposed by Kelejian (1971).

Goldberg and Maggi (1999) explicitly considered the corner solution of the

\(^{14}\)An even stronger prediction is that \( z_j \) and \( e_j \) do not enter separately once their ratio is controlled for.
protection measure on the RHS. Using full information maximum likelihood, they estimated the following system of equations. First, the “true level of protection”, the latent variable $t_j^*$, is related to organization and $\frac{x_j}{m_j}$.\footnote{Note that $e_j$ is moved to the left hand side to alleviate concerns about its endogeniety.}

\[
\frac{t_j^* e_j}{1 + t_j^*} = \gamma \frac{x_j}{m_j} + \delta I_j x_j + \varepsilon_j. 
\tag{2.2}
\]

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

\[
t_j = \begin{cases} 
\frac{1}{\mu} t_j^* & \text{if } 0 < t_j^* < \mu \\
0 & \text{if } t_j^* \leq 0 \\
1 & \text{if } t_j^* \geq \mu 
\end{cases} 
\tag{2.3}
\]

where $\mu$ is exogenously set at the value 1, 2, or $3^{16}$. Domestic production to import ratios are related to a variety of factors in

\[
\frac{x_j}{m_j} = \xi_1 Z_{1j} + u_{1j}. \tag{2.4}
\]

\[
I_j^* = \xi_2 Z_{2j} + u_{2j} \\
I_j = \begin{cases} 
1 & \text{if } I_j^* > 0 \\
0 & \text{if } I_j^* \leq 0 
\end{cases} \tag{2.5}
\]

\footnote{Though there is no reason for $\mu$ not to be less than unity as quotas may be barely binding.}
where $I^*_j$ is a latent variable for political organization. The key variables $\gamma$ and $\delta$ have the predicted sign and are significant at the 5% level. No matter what level of $\mu$ is used, the estimate of $a$ is high (over 49) and $\alpha_L$ is close to unity (over .95), though as expected, a high $\mu$ reduces the estimates of $a$ and $\alpha_L$. A high $\mu$ raises true tariffs and this, in turn, is consistent with a higher $\delta$ and lower $\gamma$ and hence, lower weight on welfare and a lower degree of organization.\footnote{Solving gives $a = \frac{\gamma + 1}{\delta - \gamma}$, $\alpha_L = \frac{-\gamma}{\delta - \gamma}$, where $\delta > 0$ and $\gamma < 0$.}

### 3. A Simple Model of Imports

We now develop the simple model of imports we will simulate. 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 all of the foreign goods are produced abroad and consumed domestically and that each good’s demand depends only on its own price and random shocks. Let $q^H_{ij}$ be the equilibrium quantity of home goods in industry subindustry $j$, and let $p^H_{ij}$ 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:

$$\ln q^H_{ij} = a_{ij} + a_{ij} \ln p^H_{ij} + x_{ij} + u_{ij}.$$  

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

$$\ln q^H_{ij} = a_{ij} + a_{ij} \ln p^H_{ij} + x_{ij} + u_{ij}.$$
The random terms $xhd_i$ and $xhs_i$ are industry specific demand and supply shocks, and hence, common across all sub industries, while $uhd_{ij}$ and $uhs_{ij}$ are sub industry specific demand and supply shocks and are idiosyncratic to each subindustry. All shocks are assumed to be iid with normal distributions though the parameters of the distribution differ. Thus, for all $i$, $xhd_i$ has mean $\mu_{xhd}$ and standard deviation $\sigma_{xhd}$, while $xhs_i$ has mean $\mu_{xhs}$ and standard deviation $\sigma_{xhs}$. Similarly, for all $ij$, $uhd_{ij}$ has mean 0 and standard deviation $\sigma_{uhd}$, while $xhs_i$ has mean 0 and standard deviation $\sigma_{uhs}$. Equilibrium satisfies

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

Similarly, let import demand be given by

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

and supply by:

$$\ln q_{ij}^{Ms} = amd_1 + amd_2 \ln p_{ij}^M + xms_i + ums_{ij}.$$

As before, the random terms $xmd_i$, $xms_i$, $umd_{ij}$, and $ums_{ij}$ are industry and subindustry specific demand and supply shocks. Their parameters are given by $(\mu_{xmd}, \sigma_{xmd})$, $(\mu_{xms}, \sigma_{xms})$, $(0, \sigma_{umd})$, and $(0, \sigma_{ums})$ respectively. Equilibrium satisfies

$$q_{ij}^{Md} = q_{ij}^{Ms} = q_{ij}^M.$$

We assume that there are $n_t = 100$ industries and each industry has $n_j = 6$ subindustries. Each subindustry $ij$ is politically organized with probability
We allow for some variation in the political organization probability across industries: \( P_{oi} = .9 \), \( P_{oi} = .9 \) with probability \( .2 \), \( P_{oi} = .8 \) with probability \( .3 \), \( P_{oi} = .6 \), \( P_{oi} = .6 \) with probability \( .3 \), \( P_{oi} = .1 \) with probability \( .1 \) and \( P_{oi} = .1 \) with probability \( .1 \). 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 0.6, the number of industries that are politically organized will be clustered around 0.6.

We simulate the output and prices of each industry by first drawing \( n_t \) industry demand and supply shocks \( xmd_i \) and \( xms_i \) for \( i = 1, ..., n_t \) and for each industry \( i \), draw \( n_j \) subindustry demand and supply shocks \( umd_{ij} \) and \( ums_{ij} \) for \( j = 1, ..., n_j \). Then, given these shocks and parameters of the 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 all subindustries. That is, the quota becomes binding in industry \( ij \) 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 \( q_{ij}^{Me} \) for subindustry \( ij \) exceeds \( \hat{Q} \), then actual imports, \( q_{ij}^M \), equals \( \hat{Q} \) and \( d_{ij}^q \) = 1. Otherwise, \( q_{ij}^M = q_{ij}^{Me} \) and \( d_{ij}^q = 0 \).

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

\[
Q_i^H = \sum_{j=1}^{n_j} q_{ij}^H
\]

for home goods and

\[
Q_i^M = \sum_{j=1}^{n_j} q_{ij}^M
\]
for foreign goods.

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

$$C_i = \frac{\sum_{j=1}^{n_j} q_{ij} q_{ij} q_{ij}}{Q_i^M}.$$ 

That is, coverage ratio is the fraction of industry output $i$ where quota is binding. Furthermore, the inverse import penetration ratio for industry $i$ is obtained as:

$$IP_i = \frac{Q_i^M + Q_i^H}{Q_i^M} = 1 + \frac{Q_i^H}{Q_i^M}.$$ 

We also derive the political organization dummy $d_{p,i}$ of industry $i$ as:

$$d_{p,i} = 1 \text{ if } \sum_{j=1}^{n_j} d_{p,ij} > \frac{n_j}{2}$$

$$= 0 \text{ otherwise.}$$

In other words, we call industry $i$ politically organized if more than half of its sub-industries are politically organized.

We chose the parameters of the above 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: $ahd_1 = 4.0$, $ahd_2 = -1.2$, $ahs_1 = 3.0$, $ahs_2 = 1.3$, $amd_1 = 1.0$, $amd_2 = -1.5027$, $ams_1 = 1.0$, $ams_2 = 1.1$.

The import demand elasticity, i.e., $-amd_2$, is set at the mean of the industry import demand elasticity. Furthermore,
In Table 1, we compare the simulation of the model to the data. The model matches the average political organization, NTB, log output/import ratio and the standard error of log output import ratio reasonably closely. Notice that we did not vary the import demand elasticity because, together with the uniform quota level, it would generate correlation between the import demand elasticity and the NTB coverage ratio in the simulation, which we wanted to purge from the model.

<table>
<thead>
<tr>
<th>Table 1: Summary statistics of the simulation and the data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Political organization frequency</td>
</tr>
<tr>
<td>NTB positive</td>
</tr>
<tr>
<td>Average log output/import ratio</td>
</tr>
<tr>
<td>Std. error of log output/import ratio</td>
</tr>
</tbody>
</table>

4. Estimating the Model Using Simulated Data

Since the two major papers use slightly different techniques we consider both.

\[
\mu_{hd} = \mu_{hs} = \mu_{md} = \mu_{ms} = .02 \\
\sigma_{xhd} = \sigma_{xhs} = 2.0, \sigma_{xmd} = \sigma_{xms} = .64 \\
\sigma_{uhd} = \sigma_{uxhs} = \sigma_{umd} = \sigma_{ums} = .3.
\]
4.1. OLS-IV Regression

To replicate the IV estimation done by Gawande and Bandyopadhyay (2000), we generated 100 data points from our simple model and estimated the following equation by three stage least squares. Note that we scale variables by dividing by 10,000 so that estimated parameters move upwards as done by Gawande and Bandyopadhyay (2000).

\[
\frac{C_i}{1 + C_i} \cdot amd_2 = \beta_0 + \gamma \frac{IP_i}{10000} + \delta d_{p,i} \frac{IP_i}{10000} + u_i.
\]

We report the OLS regression results, 3 stage least squares where the instruments are the shocks: \(xhd_i, xhs_i, xmd_i, xms_i\) (3SLS 1). We also run another 3 stage least squares where the instruments included the above exogenous variables, their square terms, and interactions (3SLS 2). The results are shown in Table 2. Standard errors are shown in the parentheses. The estimates with * signs are those that are significant with a 95% confidence interval.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>3SLS 1</th>
<th>3SLS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.2807</td>
<td>0.3642</td>
<td>0.3009</td>
</tr>
<tr>
<td></td>
<td>(0.0293)*</td>
<td>(0.0814)*</td>
<td>(0.0349)*</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>-36.63</td>
<td>-224.1</td>
<td>-67.16</td>
</tr>
<tr>
<td></td>
<td>(11.57)*</td>
<td>(102)*</td>
<td>(23.30)*</td>
</tr>
<tr>
<td>(\delta)</td>
<td>38.25</td>
<td>261.4</td>
<td>71.95</td>
</tr>
<tr>
<td></td>
<td>(11.60)*</td>
<td>(121)*</td>
<td>(25.47)*</td>
</tr>
</tbody>
</table>
Notice that in all cases above, the coefficients on the inverse import penetration ratio are significantly negative, the coefficients of the inverse import penetration ratio times the political organization dummy are significantly positive, and the sums of these two coefficients are positive, just as the protection for sale theory predicts. This is in spite of the fact the data comes from a simple model where the quota is set exogenously at a uniform level in all sub-industries, the import elasticity is set constant across all industries, and political organization is completely exogenous to the system. That is, it is fair to say that the simulated data comes from a model that has no protection for sale.

4.2. Maximum Likelihood Estimation

Next we follow Goldberg and Maggi (2000) and assume the error terms of the equations (2)-(6) are jointly normally distributed. That is, \((\epsilon_i, u_1, u_{21}) \sim N(0, \Sigma)\). We use full-information maximum likelihood to estimate the parameters of the model, where the instruments are the exogenous demand and supply shocks: \(Z_i = (xhd_i, xhs_i, xmd_i, xms_i)\). The estimation results are shown in Table 3.

| Table 3: Maximum Likelihood Results |  
|------------------------------------|---

<table>
<thead>
<tr>
<th>Specification 1</th>
<th>Specification 2</th>
<th>Specification 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff for Constant</td>
<td>0.1812</td>
<td>0.3568</td>
</tr>
<tr>
<td></td>
<td>(0.0923)</td>
<td>(0.141)*</td>
</tr>
<tr>
<td>Coeff for $d_{p,j}$</td>
<td></td>
<td>0.9258</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.173)*</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-0.01713</td>
<td>-0.02268</td>
</tr>
<tr>
<td></td>
<td>(0.00937)</td>
<td>(0.00914)*</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.02144</td>
<td>0.02573</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.00958)*</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-493.71</td>
<td>-490.89</td>
</tr>
</tbody>
</table>

Specification 1 estimates the original equation system (2.2)-(2.5) taking $\mu = 1$. Specification 2 adds a constant to equation (2.2), and specification 3 further adds the political organization dummy to the RHS of equation (2.2). In all specifications, the coefficient of the inverse import penetration ratio is negative and that of the product of the political organization dummy and the inverse import penetration ratio is positive. They are both significantly different from zero in specification 2. They are both significant at the 90% significance level in specification 1 as well. In both specifications 1 and 2, the sum of the coefficients of the terms that include inverse import penetration ratio is positive, which is in line with the results of Goldberg and Maggi (2000). Again, we obtain these results, consistent with the protection for sale model of Grossman and Helpman (1994), even though in the simulated model there is no protection for sale effect.
5. 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 general equilibrium model of protection, which does not even have any aggregation over sub-industries. Suppose that the demand for and supply of home goods has no random component:

\[
\ln q_i^{Hd} = ahd_1 + ahd_2 \ln p_i^H \\
\ln q_i^{Hs} = ahs_1 + ahs_2 \ln p_i^H.
\]

Then, home goods equilibrium quantity is:

\[
\ln q_i^H = \frac{ahd_2 ahs_1 - ahs_2 ahd_1}{ahd_2 - ahs_2}.
\]

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

\[
\ln q_i^H = \frac{ahd_2 ahs_1 - ahs_2 ahd_1}{ahd_2 - ahs_2} = 0.
\]

For imported goods in the same industry, however, demand and supply are random. Furthermore, let

\[
\ln q_i^{Md} = amd_1 + amd_2 \ln p_i^H + xmd_i \\
\ln q_i^{Ms} = ams_1 + ams_2 \ln p_i^H + xms_i.
\]
Thus, the equilibrium of the foreign goods market is

\[ \ln q_i^M = \frac{am_2 a m_1 - a m_2 a m_1}{am_2 - a m_2} + \frac{am_2 a m_1 - a m_2 a m_1}{am_2 - a m_2} . \]

We set the parameters so as to set the foreign goods equilibrium to be as follows.

\[ \ln q_i^M = -1.0 + 2.0 X_i , \]

where \( X_i \) is assumed to be uniformly distributed on \([0,1]\). Also, we set the uniform quota level, \( \hat{Q} = 1 \) so \( \ln \hat{Q} = 0 \). As before, organization is random and the protection occurs if the quota is binding and the sector is organized. There is, of course, no PFS.

Then, the coverage ratio, \( C_i \), the ratio of trade under quota to total trade is:

\[ C_i = 0 \text{ so } \frac{C_i}{1+C_i} = 0 \text{ if } \ln(q_i^M) = -1.0 + 2.0 X_i < 0 \text{ and } \]

\[ C_i = 1 \text{ so } \frac{C_i}{1+C_i} = .5 \text{ if } \ln(q_i^M) = -1.0 + 2.0 X_i \geq 0 . \]

Now suppose that half the industries are not organized. The inverse import penetration ratio for them is:

\[ II P_i = \frac{q_i^H + q_i^M}{q_i^M} = \frac{1}{q_i^M} + 1 = \frac{1}{e^{-1.0 + 2.0 X_i}} + 1 . \]

The inverse import penetration ratio for the other half of the industries, which
are politically organized, is:

\[ IIP_i = \begin{cases} 
\frac{1}{e^{-1.0 + 2.0X_i}} + 1, & \text{if } \ln(q_i^M) = -1.0 + 2.0X_i < 0, \\
2, & \text{if } \ln(q_i^M) = -1.0 + 2.0X_i \geq 0.
\end{cases} \]

Now for concreteness, number the industries that are not organized by integers between 1 and 1000, and number the industries that are organized by integers between 1001 and 2000. Industries with a high index and which are organized are going to have the quota bind if it is invoked and are going to have the quota invoked. Instead of drawing \( X_i \) from the uniform distribution over the unit interval, for industry \( i = 1, \ldots, 1000 \), we set \( X_i = \frac{i}{1000} \) and for industry \( i = 1001, \ldots, 2000 \), we set \( X_i = \frac{i-1000}{1000} \).

Figure 1 plots the import shock \( X_i \) and the import quantity. Notice that for industry \( i = 1001, \ldots, 2000 \), which are politically organized, the quota binds and import quantity equals the quota when \( X_i \) is large (industries 1500 to 2000).

Figure 2 plots the protection measure. The coverage ratio is positive only for industries that are politically organized and whose quota is binding, i.e., industry 1500 to 2000. Their protection measure is .5. Thus, the protection measure in Figure 2 is what we need to fit.
Figure 1: Import shock and import quantity

Figure 2: Protection Measure
Figure 3 plots the inverse import penetration ratio, $IIP_i$. As we can see, the inverse penetration ratio is high for industries with small imports and low for industries with large imports. The inverse import penetration ratio is constant for industries with a high index 1500 to 2000 because of the binding quota.

We next plot the inverse import penetration ratio times the political organization dummy in Figure 4, i.e., $I_{p,j} \times IIP_i$.

Notice that for industries 1 to 1000, the politically organized dummy times the inverse demand penetration ratio is zero because they are never politically organized. Let us try to fit the protection measure in Figure 2 by using inverse import penetration ratio in Figure 3 and politically organized dummy times the inverse import penetration ratio in Figure 4. That is, we run an OLS regression where the dependent variable is the protection measure and the RHS variables are constant term, the inverse import penetration ratio and the politically organized inverse import penetration ratio. That is,
Then, the coefficient estimates are:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0.3728</td>
<td>0.0160</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.1571</td>
<td>0.00718</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0921</td>
<td>0.00345</td>
</tr>
</tbody>
</table>

Again, note the opposite signs of $\beta_1$ and $\beta_2$ as in the PFS model. Figure 5 plots the dependent variable and the model prediction.
There seems to be a positive correlation between protection and inverse import penetration for politically organized sectors but a negative one between protection and inverse import penetration for non organized sectors. This is what the regression is picking up.

We can confirm the above insight by looking at the regression results in a different angle, i.e., by using the partitioned regression. Let $RIP_t$ be the component of the inverse import penetration ratio that is orthogonal to the politically organized inverse import penetration ratio. We obtain it by regressing the constant term and $I_{p,j} \times IIP_i$ and taking the residual. The blue line in Figure 6 plots the orthogonal component.

Due to the properties of the partitioned regression, the coefficients of the OLS regression of the orthogonal component on the protection measure gives the coefficient on the inverse import penetration ratio back. As can be seen from the graph, the blue is the orthogonal component of the inverse import penetration
ratio, which clearly is negatively correlated with the protection measure, which is the reason behind the negative coefficient of the inverse import penetration ratio in the original OLS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0.1253</td>
<td>0.00443</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.1571</td>
<td>0.00792</td>
</tr>
</tbody>
</table>

The red line is the prediction by the constant term and the orthogonal component. Similarly, let $RIIP_i$ be the component of the politically organized inverse import penetration ratio that is orthogonal to the inverse import penetration ra-
tio. We obtain it by regressing the constant term and $IIP_i$ on $I_{p,j} \times IIP_i$ on and taking the residual. The blue line in Figure 7 plots the orthogonal component.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>$0.1253$</td>
<td>$0.00421$</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$0.09205$</td>
<td>$0.00361$</td>
</tr>
</tbody>
</table>

Again, the coefficients of the OLS regression of the orthogonal component on the protection measure give the coefficient on the politically organized inverse import penetration ratio back. As can be seen from the graph, the blue is the orthogonal component of the politically organized inverse import penetration ratio, which clearly is positively correlated with the protection measure, which is the
reason behind the positive coefficient of the inverse import penetration ration in the original OLS.

The qualitative aspects of the above results do not change if we used IV estimation instead of OLS. In the next table, we show the estimation results of the same equation where we use the inverse import penetration ratio and its square term as instruments. In this case, $\beta_1 + \beta_2 > 0$, which is even more consistent with the Protection for Sale model than the OLS results.

Conventional empirical studies in trade estimating the political economy effects use non-tariff barriers as a proxy for tariff protection measures, even though non-tariff barriers could be better interpreted as quotas. The above results show that the real reason behind the results in support of Grossman-Helpman protection for sale models could be the difference between the quota being binding and non-binding. That is, $\beta_2 IP_i \times IIP, \beta_2 > 0$ fits well to the industries under quota (1500 to 2000) and industries that are not politically organized, but does not fit well to industries that are politically organized but not under quota (1001 to 1499).

On the other hand, $\beta_1 IIP_i, \beta_1 < 0$ fits well to the industries that are politically organized since those with high equilibrium imports face binding quotas, but fits very poorly those that are not politically organized. Hence, it is natural that combining both would give the best fit, and those results coincide to the signs obtained by Goldberg-Maggi and others.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0.3195</td>
<td>0.166</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.6102</td>
<td>0.280</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>1.008</td>
<td>0.645</td>
</tr>
</tbody>
</table>
6. Robustness Checks for PFS

The original Grossman and Helpman model imposes a strong structural model restriction the data. There has been some work done checking the robustness of the Protection for Sale results with respect to the changes in the model specifications. Examples include the original Goldberg and Maggi (1999), Gawande and Bandyopadhyay (2000), Facchini et al. (2004), and others. In this section, we further examine the model specification issue, in particular on the robustness of the assumption that in equation (1) only the inverse import penetration matters and not imports themselves or domestic production, once inverse import penetration has been controlled for. We use part of the data used by Gawande and Bandyopadhyay (2000) to estimate various specifications of the Protection for Sale equation. The results are summarized in Table 4. We used IV’s similar to Bombardini (2004) first, to replicate her results. Specification 1 is the replication of the basic results of Bombardini (2004). The size of the coefficients differs from that of Gawande and Bandyopadhyay (2000): theirs are around — SUSUMU NUMBERS). We hope to fully replicate their results once we obtain more information from them. In specification 2, we added the political organization dummy on the RHS. In both cases, the coefficients on the G-H import term and the interaction term of the political organization dummy and the G-H import term have the expected sign, and are either significantly different from zero or close to being significant. If we add log of import value and log of consumption value to the RHS, then even though the signs remain the same, the G-H coefficients are no longer significant or close to being significant at the 95% confidence interval, whereas log consumption values are significant in both specifications 3 and 4. This suggests that the strong

\[\text{Their original working paper version had alternative specifications that included production and imports separately as well as their ratio, but this did not survive in the published version.}\]
functional form assumptions behind G-H protection for sale equation may not be supported in the data.

Table 4
Robustness Checks
<table>
<thead>
<tr>
<th></th>
<th>Specification 1</th>
<th>Specification 2</th>
<th>Specification 3</th>
<th>Specification 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.04161</td>
<td>-0.03399</td>
<td>-0.2683</td>
<td>-0.2715</td>
</tr>
<tr>
<td></td>
<td>(0.0176)*</td>
<td>(0.0232)</td>
<td>(0.0626)*</td>
<td>(0.0654)*</td>
</tr>
<tr>
<td>$z_i$</td>
<td>-3.077</td>
<td>-3.397</td>
<td>-2.905</td>
<td>-2.766</td>
</tr>
<tr>
<td></td>
<td>(1.55)*</td>
<td>(1.68)*</td>
<td>(1.69)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>$I_i z_i$</td>
<td>3.044</td>
<td>3.378</td>
<td>2.883</td>
<td>2.745</td>
</tr>
<tr>
<td></td>
<td>(1.58)</td>
<td>(1.71)*</td>
<td>(1.65)</td>
<td>(1.84)</td>
</tr>
<tr>
<td>intermtar</td>
<td>0.7855</td>
<td>0.8226</td>
<td>1.119</td>
<td>1.107</td>
</tr>
<tr>
<td></td>
<td>(0.244)*</td>
<td>(0.256)*</td>
<td>(0.253)</td>
<td>(0.263)*</td>
</tr>
<tr>
<td>intermtb</td>
<td>0.3598</td>
<td>0.3495</td>
<td>0.3520</td>
<td>0.3554</td>
</tr>
<tr>
<td></td>
<td>(0.0629)*</td>
<td>(0.0663)*</td>
<td>(0.0612)</td>
<td>(0.0646)*</td>
</tr>
<tr>
<td>$d_{p,i}$</td>
<td>-0.01042</td>
<td></td>
<td>0.003527</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0206)</td>
<td></td>
<td>(0.0208)</td>
<td></td>
</tr>
<tr>
<td>Log imp</td>
<td></td>
<td>0.01136</td>
<td>0.01173</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00786)</td>
<td>(0.263)</td>
<td></td>
</tr>
<tr>
<td>Log cons.</td>
<td></td>
<td>0.02005</td>
<td>0.1989</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00935)*</td>
<td>(0.00941)*</td>
<td></td>
</tr>
</tbody>
</table>
7. Conclusion

In this paper, we examined the model specification issue of the conventional structural estimation of Grossman-Helpman Protection for Sale models. To do so, we generated artificial data from a simple equilibrium model of trade where the political organization is purely random and government imposes a quota on politically organized industries uniformly whenever it is binding and where all industries have the same import demand elasticity. That is, we constructed the model so that there was no protection for sale effect. Parameters of the model are selected so that the basic statistics of the simulated data match the actual data reasonably closely. When we estimated the simulated model, we found all the coefficients to be consistent with the protection for sale model.

We then conducted a similar exercise using a simpler model without any industry aggregation to see why our simple model delivered all the protection for sale results. We find that the main reason behind the spurious results is the use of NTB coverage ratio as the proxy for tariff protection. That is, the results can come from a model where quotas can be obtained and these could be either binding or non-binding if imposed. Their imposition depends on organization: politically organized sectors get them, others do not\textsuperscript{19}.

\textsuperscript{19}Facchini, et al. (2004) constructed and estimated the quota version of the Protection for Sale model. However, their model does not incorporate demand and supply shocks, hence the quota was always binding. They attached error terms to the reduced form equation to allow for variation in the data. Hence, the reduced form turned out to be similar to the original protection for sale equation, with some minor changes.
References


