

# Protection for Sale Under Monopolistic Competition: An Empirical Investigation

Pao-Li Chang<sup>\*†</sup>

School of Economics and Social Sciences

Singapore Management University, Singapore 178903

Myoung-Jae Lee

Department of Economics

Korea University, Seoul, Korea

August 7, 2006

## Abstract

This paper proposes a general empirical framework to estimate the protection-for-sale model, where the protection regime shifts according to a sector's market structure (perfectly or monopolistically competitive). We base the protection structure on Grossman and Helpman (1994) for the subset of perfectly competitive sectors and on Chang (2005) for the subset of monopolistically competitive sectors. The two protection regimes are simultaneously estimated with joint constraints. The results of the J-test consistently reject the homogeneous (perfect competition) protection-for-sale model often adopted in previous literature and suggest a direction of improvement toward the proposed heterogeneous protection structure model.

*Keywords:* endogenous trade policy; campaign contribution; monopolistic competition; intraindustry trade; import penetration

*JEL classification:* F12; F13

---

<sup>\*</sup>Tel: +65-6828-0830; fax: +65-6828-0833. *E-mail address:* plchang@smu.edu.sg (P.-L. Chang).

<sup>†</sup>The paper benefits from the comments by seminar participants at the Department of Economics, University of Kiel, Germany, and the Institute of Economics, Academia Sinica, Taiwan. We also wish to thank Jonathan Eaton at the Asia Pacific Trade Seminars (APTS) for his suggestions. This paper is funded by the Office of Research, Singapore Management University.

# 1. INTRODUCTION

Politics has been recognized by economists as an crucial element of trade policy making. How it interacts with the economics of trade policy and to what extent it matters, however, remain an open question and continue to be investigated by theoretical and empirical research. Among them, the paper by Grossman and Helpman (1994) (henceforth G-H) emerged as a popular theory. In G-H, tariff formation is modeled as the outcome of a menu auction (Bernheim and Whinston, 1986). In an economy of perfectly competitive sectors, sectoral interest groups bid for favorable trade policy by implicitly promising campaign contributions which are a function of potential trade policies; the government in turn optimizes by choosing the trade policy that maximizes the sum of its campaign contribution receipts and weighted aggregate welfare. The G-H model was later extended by Chang (2005) to the setting of monopolistically competitive sectors, which arguably represent a significant portion of industrial production and international trade (Helpman, 1999). Chang (2005) shows that the endogenous protection structure for monopolistically competitive sectors differs systematically from that for perfectly competitive sectors. First of all, the benchmark welfare-maximizing import tariff is strictly positive, which is in contrast with the free trade prediction under perfect competition. Second, as in G-H, lobbying efforts of competing interest groups jointly raise import protection levels in organized sectors (represented by a lobby) and lower them in unorganized sectors (not represented by a lobby) from the benchmark welfare-maximizing levels. However, the endogenous tariff level in unorganized sectors will not fall below zero. This is contrary to G-H where unorganized sectors will be penalized by negative protection. Third, the level of import protection varies inversely with the degree of import penetration, regardless of whether or not the sector is organized. In other words, sectors with higher import penetration will receive lower protection. This negative relationship applies only to organized sectors in G-H; in unorganized sectors, higher import penetration leads to higher protection (or more precisely, less negative protection).

The last theoretical prediction is of most interest to empirical researchers. The general perception before the G-H model was that a sector badly injured by imports would tend to receive higher protection. This perception was supported by most of the empirical work conducted before G-H; however, their findings were often derived from *ad hoc* regression models without theoretical underpinnings. The G-H model challenged this conventional view, as it predicts the opposite for organized sectors. Several authors (Goldberg and Maggi, 1999; Gawande and Bandyopadhyay,

2000; Eicher and Osang, 2002) have hence estimated the G-H model and tested this prediction. In general, their findings support the G-H model. The results of Chang (2005) offer a picture completely at odd with the conventional view, as the import protection level actually falls with import penetration, regardless of whether or not the sector is organized. This difference in findings between G-H and Chang (2005) suggests that the prediction of the G-H-type model in this regard will depend on the nature of market structure. Hence, a correctly-specified empirical model relating to the G-H-type model should allow for heterogeneous responses of protection to import penetration across sectors of different market structures. This has not been considered by previous cross-sectional G-H-type empirical studies.

Difference in market structure not only implies difference in the signs of protection response to import penetration, the absolute magnitude of response also differs. In G-H, the magnitude of response in organized and unorganized sectors can be expressed as two functions of the two political parameters – the government’s weight on aggregate welfare relative to campaign contributions, and the economy’s degree of political representation in the trade policy area. As a result, the two political parameters can be identified. In the model of Chang (2005), the responses can be similarly expressed as two functions of the two political parameters. However, the functional forms differ from those in G-H and hence in general will imply different estimates of the political parameters. Thus, if the market structure for a sector is misspecified, the overall estimation of the political parameters is likely to be biased.

To account for the heterogeneity in protection structure across different market structures and the effect of this heterogeneity on the estimation of the political parameters, this paper proposes a general empirical specification that accommodates both perfectly and monopolistically competitive sectors. We base the empirical framework on G-H for perfectly competitive sectors and on Chang (2005) for monopolistically competitive sectors. In order to implement this heterogeneous market structure specification, however, we need to classify sectors into perfectly or monopolistically competitive sectors. We construct a market structure indicator variable to assign sectors into one of the two subsets. The import protection equations for the two subsets of sectors are then jointly estimated. Recall from the previous paragraph that the four responses of protection to import penetration – two for each market structure – are functions of the two political parameters and yet the political parameters are shared by all sectors in the same economy. Thus, the two protection

equations cannot be estimated separately; in fact, the response coefficients in the heterogeneous specification are explicitly related by two linear constraints.

We construct the market structure dummy based on the condition that firms in a monopolistically competitive sector necessarily face a demand elasticity larger than one, and the criterion that the degree of seller competition in a monopolistically competitive sector be lower than a specified threshold. We allow the threshold to vary across a wide range of observed values and estimate the heterogeneous model conditional on each resulting classification of sectors. We then search for the optimal specification of the market structure dummy that gives rise to the maximum likelihood for the heterogeneous model. The use of extraneous economic variables on *a priori* grounds to classify observations into different regimes follows from Fair and Jaffee (1972) in the literature of switching regressions. We generalize their approach by not requiring knowledge of the threshold value and let the data determine endogenously the optimal separation of the sample. This requires at most  $n$  searches (where  $n$  is the sample size), as compared to  $2^n$  searches (which is intractable) if no economic constraint is imposed on the construction of the market structure dummy.<sup>1</sup>

The heterogeneous model with the optimal classification of sectors is then compared to the homogeneous G-H model with perfect competition which is often adopted by previous cross-sectional G-H-type empirical studies. Each model is taken in turn as the null hypothesis and tested against the alternative model according to the non-nested *J-test* of Davidson and MacKinnon (1981). The heterogeneous model can be regarded as a joint hypothesis that the market structure dummy is correctly specified and that the model of Chang (2005) is correctly specified for the subset of monopolistically competitive sectors. The homogeneous perfect competition model, on the other hand, hypothesizes that the G-H model is correctly specified for all sectors. If a model as the null is rejected by the J-test, the data suggest that the alternative model provides some extra explanatory power (in a statistically significant way) over and above what the null specification does. On the other hand, if a model as the null is accepted, the test suggests that the alternative model does not significantly improve the statistical goodness-of-fit of the null. Both models can be rejected as the null hypothesis. In this case, the test suggests that neither model is complete; important

---

<sup>1</sup>Alternatively, one may contemplate the use of random switching regression models as developed by Quandt (1972), Kiefer (1978), Quandt and Ramsey (1978), Hartley (1978), and Phillips (1991) among others. However, as commented in Quandt (1972), applying random switching regression models does not allow individual observations to be identified with particular regimes but computes only the probability that one or the other regime was operative during the sample period. This does not seem particularly appealing to us, as we do not consider market structure to be a random event.

variables or structures are missing from the models. If both models are accepted, the test cannot differentiate between the two models. If one model is accepted, while the other is rejected, the test suggests that the first model provides a statistically-significant better fit to the data.

It is interesting to see that whether the proposed heterogeneous model will be supported by the data. First of all, the estimated coefficients of the import protection equation should satisfy the sign predictions of the theories of G-H and Chang (2005) respectively for the subset of perfectly and monopolistically competitive sectors. These sign predictions correspond to the two theories' predictions on the relationship between import protection and import penetration for organized and unorganized sectors. If the estimated coefficients of the heterogeneous model have the correct signs, and if the model is supported by the J-test against the homogeneous G-H model, three implications arise. First, this suggests that the model of Chang (2005) is borne out by the data for monopolistically competitive sectors, although the G-H model remains valid for the subset of perfectly competitive sectors. Second, this also suggests that the relationship between import protection and import penetration varies across sectors of different market structures and across sectors with or without political organization. Third, we may hence regard the political parameter estimates derived from the heterogeneous specification as being more representative of the economy under study than suggested by previous studies based on the homogeneous G-H model with perfect competition.

The rest of the paper is organized as follows. In Section 2, we review the basic elements of G-H and Chang (2005) and integrate the two models. In Section 3, we present the empirical model for the general specification allowing for heterogeneous market structures and explain the estimation methodology. The general specification is then compared to the homogeneous perfect competition specification based on non-nested hypothesis tests. The empirical results are given in Section 3.5. Section 4 concludes. Technical notes regarding the estimation methodologies are given in the appendices.

## 2. THE PROTECTION-FOR-SALE MODEL WITH HETEROGENEOUS MARKET STRUCTURE

In this section, we integrate the Protection-for-Sale models of Grossman and Helpman (1994) and Chang (2005). This construct allows the market structure to be heterogeneous across sectors – either perfectly or monopolistically competitive. Suppose that a country is populated by individuals with identical preference, but with potentially different endowments. The preference is given by:

$$U = C_0 + \sum_{i=1}^n U_i(C_i) \quad (1)$$

where  $C_0$  denotes consumption of good 0,  $C_i$  denotes consumption of good  $i$ ,  $i = 1, 2, \dots, n$ , and  $U_i$  is an increasing concave function. Good 0 serves as numeraire, with a world (and domestic) price equal to 1. Let  $P_i$  denote the domestic price of good  $i$ . The demand for good  $i$  implied by the preference in (1) is denoted  $D_i(P_i)$ , where  $D_i(\cdot)$  is the inverse of  $U'_i(\cdot)$ . The indirect utility of an individual with income  $E$  is given by  $V = E + \sum_{i=1}^n S_i(P_i)$ , where  $S_i(P_i) = U_i(D_i(P_i)) - P_i D_i(P_i)$  is the consumer surplus derived from good  $i$ . If sector  $i$  is monopolistically competitive,  $C_i$  represents the aggregate consumption of differentiated goods in sector  $i$ , with the aggregation following the Dixit-Stiglitz functional form (Dixit and Stiglitz, 1977):

$$C_i = \left( \sum_{k=1}^{m_i} c_{hi,k}^{\varrho_i} + \sum_{k=1}^{m_i^*} c_{fi,k}^{\varrho_i} \right)^{\frac{1}{\varrho_i}} \quad 0 < \varrho_i < 1 \quad (2)$$

where  $c_{hi,k}$  ( $c_{fi,k}$ ) is the consumption of home (foreign) variety  $k$  of good  $i$  and  $m_i$  ( $m_i^*$ ) is the number of varieties of good  $i$  produced at home (abroad). The corresponding aggregate price level  $P_i$  for the differentiated goods in sector  $i$  is:

$$P_i = \left( \sum_{k=1}^{m_i} p_{hi,k}^{1-\sigma_i} + \sum_{k=1}^{m_i^*} p_{fi,k}^{1-\sigma_i} \right)^{\frac{1}{1-\sigma_i}} \quad (3)$$

where  $p_{hi,k}$  ( $p_{fi,k}$ ) is the consumer price for home (foreign) variety  $k$  of good  $i$ , and  $\sigma_i = \frac{1}{1-\varrho_i} > 1$  is the elasticity of substitution among different varieties of good  $i$ .

Good 0 is taken to be a homogeneous good, produced one-to-one from labor, and traded freely and costlessly, so that the wage is equal to one at home and abroad. Production of the other goods

requires labor and a sector-specific input. The various specific inputs are available in inelastic supply  $\bar{K}_i$ ,  $i = 1, 2, \dots, n$ . If sector  $i$  is perfectly competitive, the production technology exhibits constant returns to scale. If sector  $i$  is monopolistically competitive, each variety of the differentiated goods is assumed to require a fixed amount of the sector-specific factor  $k_i$  in order to produce at all; after that, there is a constant unit labor requirement  $\theta_i$ . Thus, the number of varieties produced at home in sector  $i$  is  $m_i = \bar{K}_i/k_i$ . The technology abroad to produce the differentiated products is assumed to be the same as that at home. Thus, the number of varieties produced abroad in the same sector is  $m_i^* = \bar{K}_i^*/k_i$ , where  $\bar{K}_i^*$  is the amount of sector-specific factor  $i$  the foreign country is endowed with.

Let the domestic import policy  $\tau_i$  denote one plus the ad valorem import tariff rate and the domestic export policy  $s_i$  represent one plus the ad valorem export subsidy rate for sector  $i$ . Let  $\tau_i^*$  and  $s_i^*$  represent the corresponding foreign import and export policy for sector  $i$ , defined in the similar way.

Suppose sector  $i$  is perfectly competitive and the exogenous world price is  $P_i^*$ . Then, the domestic price is  $P_i = \tau_i P_i^*$  in an import-competing sector, and is  $P_i = s_i P_i^*$  in an export sector. The returns to specific factor  $i$  depend only on  $P_i$  and are denoted by  $\Pi_i(P_i)$ . It follows that the supply function of good  $i$  is  $Y_i(P_i) = \Pi_i'(P_i)$ .

Suppose sector  $i$  is monopolistically competitive instead. Assume that there are a large number of varieties (home and foreign combined) available to the consumer in sector  $i$ . Given the preference specified in (2), each variety's producer faces an approximately constant elasticity of demand,  $\sigma_i$ . Thus, with profit maximization, the producer of each variety charges the same price:  $p_{hi,k} = p_{hi} = p_{fi}^* = \frac{\theta_i \sigma_i}{\sigma_i - 1}$ , where  $p_{fi}^*$  is the producer price for each foreign variety of good  $i$  in the foreign market. Since the producer price of the home variety of good  $i$  is the same as that of the foreign variety, the difference in their consumer prices at the home market reflects trade interventions:  $p_{fi} = \frac{\tau_i}{s_i^*} p_{hi}$ . Given this, the aggregate price index for differentiated good  $i$  in (3) can be simplified as:

$$P_i = p_{hi} (m_i + m_i^* (\frac{\tau_i}{s_i^*})^{1-\sigma_i})^{\frac{1}{1-\sigma_i}}. \quad (4)$$

The utility function  $U_i$  in (1) for monopolistically competitive sectors is assumed to take the functional form:  $U_i = E_i \ln C_i$ . This amounts to assuming that an individual allocates a fixed amount of

expenditure  $E_i$  on good  $i$ . The rest of the world is assumed to share the same preference structure, but with a possibly different allocation of expenditure on various goods,  $E_i^*$ . Given this, we can derive the demand for a representative home and foreign variety of good  $i$  as:

$$\begin{aligned} c_{hi} &= \frac{E_i}{p_{hi}} \frac{1}{m_i + m_i^* \left(\frac{\tau_i}{s_i^*}\right)^{1-\sigma_i}} \\ c_{fi} &= \frac{E_i}{p_{fi}} \frac{\left(\frac{\tau_i}{s_i^*}\right)^{1-\sigma_i}}{m_i + m_i^* \left(\frac{\tau_i}{s_i^*}\right)^{1-\sigma_i}}. \end{aligned} \quad (5)$$

Similarly, a foreign individual will consume a representative home and foreign variety of good  $i$  according to:

$$\begin{aligned} c_{hi}^* &= \frac{E_i^*}{p_{hi}^*} \frac{\left(\frac{\tau_i^*}{s_i}\right)^{1-\sigma_i}}{m_i \left(\frac{\tau_i^*}{s_i}\right)^{1-\sigma_i} + m_i^*} \\ c_{fi}^* &= \frac{E_i^*}{p_{fi}^*} \frac{1}{m_i \left(\frac{\tau_i^*}{s_i}\right)^{1-\sigma_i} + m_i^*} \end{aligned} \quad (6)$$

where  $p_{hi}^* = \frac{\tau_i^*}{s_i} p_{fi}^*$  is the consumer price of a representative home variety of good  $i$  in the foreign market. Given the domestic and foreign demand for its product, a representative home producer of differentiated good  $i$  will produce at the scale of  $(Nc_{hi} + N^*c_{hi}^*)$ , where  $N$  ( $N^*$ ) is the total home (foreign) population. Thus, the returns to the specific factor used in differentiated sector  $i$  are:  $\Pi_i(\tau_i, s_i) = m_i(p_{hi} - \theta_i)(Nc_{hi} + N^*c_{hi}^*)$ .

The net tariff revenue from sector  $i$ , expressed on a per capita basis, is given by:

$$\begin{aligned} R_i &= (1 - I_i^m)(P_i - P_i^*)[C_i - \frac{1}{N}X_i] \\ &\quad + I_i^m[m_i^*(\tau_i - 1)\frac{p_{hi}}{s_i^*} c_{fi} - \frac{N^*}{N}m_i(1 - \frac{1}{s_i})p_{hi} c_{hi}^*] \end{aligned} \quad (7)$$

where  $X_i = Y_i(P_i)$  is the domestic aggregate output of good  $i$  in a perfectly competitive sector, and  $I_i^m$  is an indicator variable, which equals one if sector  $i$  is monopolistically competitive and zero otherwise. It is assumed that the government redistributes the revenue  $R = \sum_{i=1}^n R_i$  evenly to each individual.

Assume that each individual owns a unit of labor and at most one type of specific factor. Summing indirect utilities over all individuals, and noting that aggregate income is the sum of



labor income, returns to specific factors and tariff revenue, one obtains aggregate welfare:

$$W = N + \sum_{i=1}^n \Pi_i + N \sum_{i=1}^n (R_i + S_i). \quad (8)$$

We now describe the political structure. Suppose that in some subset of sectors  $L \subset \{1, 2, \dots, n\}$ , the specific-factor owners are able to form a lobby. Let  $\alpha_i$  denote the fraction of population that owns specific factor  $i$ . Summing indirect utilities over all individuals who belong to lobby  $i$ , we obtain lobby  $i$ 's aggregate well-being:

$$W_i = \alpha_i N + \Pi_i + \alpha_i N \sum_{l=1}^n (R_l + S_l). \quad (9)$$

Lobbies compete noncooperatively for the government's favor and propose contribution schedules,  $\mathbb{C}_i(\tau, s)$ , contingent on the trade-policy vector set by the government,  $(\tau, s)$ . Lobby  $i$ 's objective is to maximize the net aggregate well-being given by  $W_i - \mathbb{C}_i$ . Given the contribution schedules offered by the lobbies, the government in turn selects a trade-policy vector  $(\tau, s)$  to maximize its politically-motivated objective function, which is a combination of welfare and contributions:

$$G = \sum_{i \in L} \mathbb{C}_i + aW \quad (10)$$

where  $a \geq 0$  captures the weight of welfare in the government's objective relative to campaign contributions. As shown in G-H, if the contribution schedules offered by lobbies are truthful, the government's objective function in (10) is equivalent to:

$$\tilde{G} = \sum_{i \in L} W_i + aW. \quad (11)$$

To find the equilibrium trade policy, one can rewrite  $\tilde{G}$  as:

$$\tilde{G} = (a + \alpha_L)N + \sum_{i=1}^n (a + I_i)\Pi_i + (a + \alpha_L)N \sum_{i=1}^n (R_i + S_i). \quad (12)$$

where  $\alpha_L \equiv \sum_{i \in L} \alpha_i$  denotes the fraction of population that is represented by a lobby, and  $I_i$  is an indicator variable that equals one if sector  $i$  is organized and zero otherwise. We focus on the

endogenous import policy henceforth, and refer interested readers to Chang (2005) for details on the endogenous export policy. Let  $t_i \equiv \tau_i - 1$  denote the ad valorem import tariff rate. Taking the first-order derivative with respect to (12) yields the following results:

**PROPOSITION 1 (Endogenous Protection Structure)** *If the contribution schedules of the lobbies are truthful, the import policy that will emerge in the political equilibrium for a perfectly competitive sector satisfies*

$$\frac{t_i^o}{1 + t_i^o} = \frac{I_i - \alpha_L}{a + \alpha_L} \frac{z_i^o}{e_i^o} \quad (13)$$

where  $z_i^o = X_i^o/M_i^o$  is the equilibrium ratio of domestic output to imports, and  $e_i^o = -M_i^{o'}P_i^o/M_i^o$  is the elasticity of import demand.

On the other hand, for a monopolistically competitive sector, the import policy that will emerge in the political equilibrium satisfies

$$\frac{t_j^o}{1 + t_j^o} = \frac{I_j + a}{a + \alpha_L} \frac{\frac{\sigma_j - 1}{\sigma_j}}{\sigma_j + \frac{1}{\tilde{z}_j^o}} \quad (14)$$

where  $\tilde{z}_j^o = p_{hj}m_jc_{hj}^o / p_{fj}m_j^*c_{fj}^o$  is the equilibrium ratio of domestic output supplied to the domestic market relative to imports, and  $\sigma_j$  is the constant elasticity of substitution between domestic output and imports.

Several observations on Proposition 1 can be made. First, if the government is not politically motivated but maximizes aggregate welfare ( $a \rightarrow \infty$ ), the endogenous trade policy reduces to the welfare-maximizing trade policy, which is free trade for a perfectly competitive sector and a positive import tariff  $\frac{(\sigma_j - 1)/\sigma_j}{\sigma_j + 1/\tilde{z}_j^o}$  for a monopolistically competitive sector. Second, in both perfectly and monopolistically competitive sectors, the endogenous protection level is relatively higher than the benchmark welfare-maximizing level for organized sectors and lower for unorganized sectors. In particular, in perfectly competitive sectors, organized sectors receive positive import tariff protection while unorganized sectors face negative protection. In monopolistically competitive sectors, however, both organized and unorganized sectors receive positive import tariff protection. Third, in a perfectly competitive sector, a higher degree of import penetration ( $1/z_i$ ) corresponds to a lower level of import protection if the sector is organized, but to a higher level of import protection if the sector is unorganized. In a monopolistically competitive sector, however, a higher degree

of import penetration ( $1/\tilde{z}_i$ ) always corresponds to a lower level of import protection, regardless of whether or not the sector is organized. As a final remark, the political parameters ( $\alpha_L$  and  $a$ ) have comparable effects on the endogenous protection level under both market structures. As more people are politically represented (a larger  $\alpha_L$ ), the general protection level decreases, and as the government becomes more concerned with aggregate welfare (a larger  $a$ ), the endogenous protection level approaches the benchmark welfare-maximizing level.

### 3. THE ECONOMETRIC MODEL

If heterogeneity in market structure is ignored and all sectors are taken to be perfectly competitive, equation (13) as derived originally by Grossman and Helpman (1994) applies to all sectors. This theoretical specification has been the basis of empirical studies by Goldberg and Maggi (1999), Gawande and Bandyopadhyay (2000), and Eicher and Osang (2002). For example, Goldberg and Maggi (1999) form the following structural equation from (13):

$$\begin{aligned}\frac{t_i}{1+t_i}e_i &= \frac{I_i - \alpha_L}{a + \alpha_L}z_i + \epsilon_{p,i} \\ &= \gamma_p z_i + \delta_p I_i z_i + \epsilon_{p,i}\end{aligned}\tag{15}$$

where  $\gamma_p = -\alpha_L/(a + \alpha_L)$  and  $\delta_p = 1/(a + \alpha_L)$ . We call this specification the homogeneous perfect competition (PC) model. Theory (13) implies that the coefficients of the PC model should satisfy the following signs: (i)  $-1 < \gamma_p < 0$ , (ii)  $\delta_p > 0$ , (iii)  $\gamma_p + \delta_p > 0$ , for the nontrivial case where  $a > 0$  and  $1 > \alpha_L > 0$ .

On the other hand, if all sectors are taken to be monopolistically competitive, equation (14) applies to all sectors. A structural equation based on (14) can be derived in parallel with (15) as:

$$\begin{aligned}\frac{t_i}{1+t_i}e_i &= \frac{I_i + a}{a + \alpha_L} \frac{e_i - 1}{e_i + \frac{1}{\tilde{z}_i}} + \epsilon_{m,i} \\ &= \gamma_m w_i + \delta_m I_i w_i + \epsilon_{m,i}\end{aligned}\tag{16}$$

where  $\gamma_m = a/(a + \alpha_L)$ ,  $\delta_m = 1/(a + \alpha_L)$ , and  $w_i = \frac{e_i - 1}{e_i + \frac{1}{\tilde{z}_i}}$ . Note that we have used  $e_i$  as a proxy for  $\sigma_i$  in (16), as  $\sigma_i$  being the elasticity of substitution among different varieties of good  $i$  is also the

elasticity of substitution between home and foreign varieties, which is approximately the elasticity of import demand,  $e_i$ . Theory (14) implies that the coefficients for monopolistically competitive sectors should observe the following signs: (i)  $1 > \gamma_m > 0$ , (ii)  $\delta_m > 0$ , (iii)  $\gamma_m + \delta_m > 1$ , similarly for the nontrivial case where  $a > 0$  and  $1 > \alpha_L > 0$ .

In this paper, we propose a generalized specification of endogenous protection structure based on Proposition 1, which accommodates both potential market structures:

$$\frac{t_i}{1+t_i}e_i = (1 - I_i^m)\{\gamma_p z_i + \delta_p I_i z_i + \epsilon_{p,i}\} + I_i^m\{\gamma_m w_i + \delta_m I_i w_i + \epsilon_{m,i}\}, \quad (17)$$

subject to the constraints:  $-\gamma_p + \gamma_m = 1$  and  $\delta_p = \delta_m$ . The constraints follow directly from the definitions of  $\gamma_p$ ,  $\delta_p$ ,  $\gamma_m$  and  $\delta_m$ , and the fact that all sectors, despite their market structure, share the same set of political parameters for an economy. We call this generalized specification the heterogeneous market structure (HC) model. Proposition 1 suggests that the coefficients of the HC model should display the following signs: (i)  $-1 < \gamma_p < 0$ , (ii)  $\delta_p > 0$ , (iii)  $\gamma_p + \delta_p > 0$ , (iv)  $1 > \gamma_m > 0$ , (v)  $\delta_m > 0$ , (vi)  $\gamma_m + \delta_m > 1$ . Given the constraints ( $-\gamma_p + \gamma_m = 1$  and  $\delta_p = \delta_m$ ) imposed in the estimation, it follows that conditions (i)–(iii) are equivalent to conditions (iv)–(vi).

### 3.1 Data and Measurement

The data set for the PC model was kindly provided by Eicher and Osang (2002) and was a reconstruction of the data set described in Goldberg and Maggi (1999). We construct two additional variables,  $I_i^m$  and  $\tilde{z}_i$ , which are required for the HC model. We discuss the measurement and endogeneity issues associated with estimation of the PC and HC models below and refer the reader to the above two sources for further explanations of the data set.

The sectors investigated are 106 United States manufacturing sectors at the 3-digit Standard Industrial Classification (SIC) level. The coverage ratios for nontariff barriers (NTB) in year 1983, instead of tariffs, are used to proxy for the noncooperative endogenous protection level predicted by the above theories. This is in view of the fact that tariff levels in reality are set cooperatively among nations in international trade negotiations. Since the coverage ratio can only take values between 0 and 1, the protection variable  $t_i$  is censored. We follow the benchmark mapping in Goldberg and Maggi (1999) between the latent protection level and the nontariff barrier; that is,

for coverage ratios less than 1, they reflect the equivalent tariff level. For latent protection levels higher than 100 percent, they are mapped to the coverage ratio 1.

The trade elasticity estimates from Shiells et al. (1986) are used to measure the import demand elasticity  $e_i$  for perfectly competitive sectors, and the elasticity of substitution  $\sigma_i$  for monopolistically competitive sectors. For perfectly competitive sectors, Goldberg and Maggi (1999) brought the import demand elasticity  $e_i$  to the left hand side of the structural equation as in (15), because the variable is endogenous by theory (13) and yet no suitable instrumental variables could be clearly identified. The elasticity of substitution  $\sigma_i$  for monopolistically competitive sectors is not endogenous by theory (14). To facilitate estimation and comparison, however, we also phrase the structural equation in (16) for monopolistically competitive sectors in a similar fashion such that the left-hand side variable is also a composite of the protection measure and the elasticity measure. The elasticity terms that remain on the right-hand side of the structural equation (16) for monopolistically competitive sectors can be taken as exogenous by theory (14).

The inverse import-penetration ratio  $z_i$  for perfectly competitive sectors in (15) is measured by the ratio of the value of shipments over imports as in Goldberg and Maggi (1999). The alternative inverse import-penetration ratio  $\tilde{z}_i$  for monopolistically competitive sectors in (16), however, requires slight modifications. It is measured as the ratio of the value of shipments to the domestic market over imports, that is, the value of shipments net of exports divided by imports. Because both imports and domestic productions depend on the protection level, these two inverse penetration ratios are endogenous. We follow Goldberg and Maggi (1999) in the selection of the explanatory variables for the inverse penetration ratio,  $z_i$ . These include physical capital, inventories, engineers/scientists, white-collar worker, skilled labor, semiskilled labor, cropland, pasture, forest, coal, petroleum, and minerals (which are the explanatory variables used for import penetration in Treffer, 1993), as well as seller concentration, buyer concentration, seller number of firms, buyer number of firms, scale, capital stock, unionization, geographic concentration, and tenure (which are a subset of the explanatory variables used for import protection in the same source). This collection of explanatory variables are also used for the composite inverse penetration ratio,  $w_i$ .

The political-organization dummy  $I_i$  is exogenous in theory, but is empirically measured based on sectoral political contributions. Goldberg and Maggi (1999) construct the dummy using the threshold level of \$1,000,000,000 for political action committee (PAC) contributions. A sector is

considered politically organized for trade-policy purposes, if the sector's PAC contribution is above the threshold. Since the political contribution level  $\mathbb{C}_i$  is endogenous in theory, constructing the political-organization dummy based on  $\mathbb{C}_i$  implies that  $I_i$  is potentially endogenous. Whether taking  $I_i$  as endogenous or exogenous does not appear to make much difference in Goldberg and Maggi (1999). There also arise significant methodological problems when  $I_i$  is taken as endogenous—we explain the technical details in Appendix A. We will hence treat  $I_i$  as exogenous.

The market structure dummy  $I_i^m$  specifies which of the two protection regimes in (17) applies to sector  $i$ , and hence is crucial to the estimation of the HC model. To construct the market structure dummy  $I_i^m$ , we begin with the observation that firms in a monopolistically competitive sector necessarily face a demand elasticity larger than one. To see this, note that for theory (14) to hold for monopolistically competitive sectors, the elasticity of substitution must be greater than one. Thus, sectors with trade elasticity measures less than or equal to one are necessarily classified as perfectly competitive. As a second condition, we require that the degree of seller competition in a monopolistically competitive sector be lower than a specified threshold. This criterion relies on the hypothesis that the degree of seller competition in monopolistically competitive sectors should be low enough such that individual firms maintain a reasonable degree of monopoly power. We use one of the exogenous variables, the seller number of firms, which is the number of competing companies in a sector divided by the sector's total sales, to measure the degree of seller competition. Let '*scomp*' denote the degree of seller competition and  $\bar{\kappa}$  the chosen threshold. Thus, the market structure dummy is defined as follows: a sector is classified as monopolistically competitive, if the elasticity measure is greater than one and if the sector faces a relatively low degree of competition ( $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq \bar{\kappa}$ ). We consider an extensive range of cutoff threshold  $\bar{\kappa}$  in the estimation of the HC model and let the data endogenously choose the optimal cutoff threshold. The search for the optimal classification of sectors based on the above criteria involves at most  $n$  searches conceptually, where  $n$  is the sample size. Specifically, we begin by setting  $\bar{\kappa}$  at the highest observed value of '*scomp*'; this is equivalent to imposing no restriction on the degree of competition but relies purely on the simplistic criterion of elasticity measure. We then go down the list of observed values of '*scomp*' in a descending order as potential cutoff thresholds  $\bar{\kappa}$  and estimate the HC model based on the resulting classification of sectors. We stop the search at the point where the number of sectors classified as monopolistically competitive reaches a preset lower

bound. This lower bound is set so that a tolerable degree of freedom is maintained for the subset of monopolistically competitive sectors in the estimation of the HC model.

### 3.2 The Full Econometric Model

Let SF stand for ‘structural form’. Given the above discussion, the full econometric model for the PC model we consider is, with  $\bar{c}_i = \frac{1}{2}e_i > 0$ ,

$$\begin{aligned}
y \text{ SF} &: y_i = \max\{0, \min(y_i^*, \bar{c}_i)\} \\
y^* \text{ SF} &: y_i^* = \frac{t_i^*}{1 + t_i^*} e_i = \gamma_p z_i + \delta_p I_i z_i + \epsilon_{p,i} \\
z_i &= \zeta_p' Z_i + u_{p,i} \\
\text{observed} &: y_i, z_i, I_i, Z_i, e_i, \bar{c}_i, i = 1, \dots, n, \text{ iid across } i.
\end{aligned}$$

where  $u_{p,i}$  and  $\epsilon_{p,i}$  are dependent, which makes  $z_i$  endogenous in  $y$  SF. The latent variable  $t_i^*$  indicates the true level of protection, which is censored at zero and one when measured by the NTB coverage ratio. It follows that the upper censoring point for  $y^*$  SF is  $\bar{c}_i$ , which varies across  $i$ . The vector  $Z_i$  consists of one and the explanatory variables for the inverse penetration ratio  $z_i$ . The HC model, on the other hand, takes the following form:

$$\begin{aligned}
y \text{ SF} &: y_i = \max\{0, \min(y_i^*, \bar{c}_i)\} \\
y^* \text{ SF} &: y_i^* = \frac{t_i^*}{1 + t_i^*} e_i = (1 - I_i^m) \{\gamma_p z_i + \delta_p I_i z_i + \epsilon_{p,i}\} + I_i^m \{\gamma_m w_i + \delta_m I_i w_i + \epsilon_{m,i}\} \\
z_i &= (1 - I_i^m) \{\zeta_{pp}' Z_i + u_{pp,i}\} + I_i^m \{\zeta_{pm}' Z_i + u_{pm,i}\} \\
w_i &= (1 - I_i^m) \{\zeta_{mp}' Z_i + u_{mp,i}\} + I_i^m \{\zeta_{mm}' Z_i + u_{mm,i}\} \\
\text{observed} &: y_i, z_i, w_i, I_i, I_i^m, Z_i, e_i, \bar{c}_i, i = 1, \dots, n, \text{ iid across } i \\
\text{s.t.} &: -\gamma_p + \gamma_m = 1 \text{ and } \delta_p = \delta_m.
\end{aligned}$$

where  $u_{pp,i}$  and  $\epsilon_{p,i}$ , as well as  $u_{mm,i}$  and  $\epsilon_{m,i}$ , are dependent, which makes  $z_i$  and  $w_i$  endogenous. In the above specification, we allow the endogenous variables ( $y_i^*$ ,  $z_i$ , and  $w_i$ ) to follow different regimes depending on the market structure, with  $y^*$  SF explicitly guided by Proposition 1. As discussed earlier, the SF coefficients of the two protection regimes must be related in the HC model by the constraints:  $-\gamma_p + \gamma_m = 1$  and  $\delta_p = \delta_m$ .

### 3.3 Estimation

In view of the *iid* assumption, the subscript  $i$  will be omitted often in the following. The goal is to estimate the SF parameters  $\gamma_p$  and  $\delta_p$  in the PC model, and  $\gamma_p$ ,  $\delta_p$ ,  $\gamma_m$ , and  $\delta_m$  in the HC model subject to the constraints:  $-\gamma_p + \gamma_m = 1$  and  $\delta_p = \delta_m$ . There are two econometric problems in pursuing the goal. One is the censoring in the  $y$  SF and the other is the endogeneity of  $z$  and  $w$ . The  $z$ - (and  $w$ -) equation regressor vector  $Z$  that includes unity is essentially an instrument vector for  $z$  (and  $w$ ), but due to the nonlinearity caused by the censoring, the usual instrumental variable estimator (IVE) for linear models is not applicable. As explained in Appendix A, we explored four methods. They differ in terms of whether the censored maximum likelihood estimator (CMLE) or the censored least absolute deviation estimator (CLAD) is used to estimate the  $y$  SF, and whether the minimum distance estimator (MDE) or the two-stage least square type estimator (2SLS) is used to account for the endogeneity of the  $z$ - (and  $w$ -) equation. The system maximum likelihood estimator might be yet another option. Given the relatively small sample size coupled with the relatively large number of parameters, we are led to adopt the 2SLS-CMLE combination after a preliminary data analysis, as it is less restrictive than the system MLE in terms of the requisite assumptions, and is more numerically stable than the MDE/CLAD.

Let  $\sigma_a$  and  $\rho_{a,b}$  stand for, respectively, the standard error of variable  $a$  and the correlation between variables  $a$  and  $b$ . For the PC model, substituting the  $z$ - equation into the  $y^*$  SF, we get the  $y^*$  SF-2SLS for the PC model and the conditional variance of the error term:

$$y^* \text{ SF-2SLS} \quad : \quad y^* = \gamma_p \zeta_p' Z + \delta_p I \zeta_p' Z + \{(\gamma_p + \delta_p I)u_p + \epsilon_p\}, \quad (18)$$

$$\sigma_{v_p}^2 \quad \equiv \quad V(\{\cdot\}|Z, I) = (\gamma_p + \delta_p I)^2 \sigma_{u_p}^2 + 2(\gamma_p + \delta_p I) \rho_{u_p, \epsilon_p} \sigma_{u_p} \sigma_{\epsilon_p} + \sigma_{\epsilon_p}^2. \quad (19)$$

The error term in  $\{\cdot\}$  is heteroscedastic depending on  $I$ , and consists of two errors. Let  $\hat{\zeta}_p$  and  $\hat{\sigma}_{u_p}$  denote the first-stage least square estimate (LSE) for  $\zeta_p$  and  $\sigma_{u_p}$  in the  $z$ - equation. In the second stage, the  $y$  SF-2SLS can be estimated applying the usual CMLE with  $\zeta_p$  and  $\sigma_{u_p}$  replaced by  $\hat{\zeta}_p$  and  $\hat{\sigma}_{u_p}$ , and with an adjustment owing to the heteroscedasticity. Define  $f(Z, I; \gamma_p, \delta_p, \zeta_p) \equiv \gamma_p \zeta_p' Z + \delta_p I \zeta_p' Z$ , which is the SF-2SLS regression function of the PC model in (18). The second-stage



log-likelihood function for the  $y$  SF-2SLS of the PC model is

$$\begin{aligned}
L_p = \sum_i \{ & 1[y_i = 0] \ln \Phi\left(\frac{-f(Z, I; \gamma_p, \delta_p, \hat{\zeta}_p)}{\sigma_{v_p, i}}\right) \\
& + 1[0 < y_i < \bar{c}_i] \ln \frac{\phi((y_i - f(Z, I; \gamma_p, \delta_p, \hat{\zeta}_p))/\sigma_{v_p, i})}{\sigma_{v_p, i}} \\
& + 1[y_i = \bar{c}_i] \ln \Phi\left(\frac{f(Z, I; \gamma_p, \delta_p, \hat{\zeta}_p) - \bar{c}_i}{\sigma_{v_p, i}}\right) \}, \tag{20}
\end{aligned}$$

which is to be maximized over  $\gamma_p$ ,  $\delta_p$ ,  $\sigma_{\epsilon_p}$ , and  $\rho_{u_p, \epsilon_p}$ ; note that  $\sigma_{v_p, i}$  is a function of  $I_i$  and the parameters, which are suppressed to simplify notations. Denote the resulting estimates  $(\dot{\gamma}_p, \dot{\delta}_p, \dot{\sigma}_{\epsilon_p}, \dot{\rho}_{u_p, \epsilon_p})$ .

Analogous steps are taken for the HC model. Substituting the  $z$ - and  $w$ - equation into the  $y^*$  SF, we get the  $y^*$  SF-2SLS for the HC model and the conditional variance of the error term:

$$\begin{aligned}
y^* \text{ SF-2SLS} \quad : \quad y^* = & (1 - I^m)(\gamma_p \zeta'_{pp} Z + \delta_p I \zeta'_{pp} Z) + I^m(\gamma_m \zeta'_{mm} Z + \delta_m I \zeta'_{mm} Z) \\
& + \{(1 - I^m)(\gamma_p + \delta_p I)u_{pp} + I^m(\gamma_m + \delta_m I)u_{mm} + (1 - I^m)\epsilon_p + I^m\epsilon_m\}, \tag{21}
\end{aligned}$$

$$\begin{aligned}
\sigma_{v_h}^2 \equiv V(\{\cdot\}|Z, I, I^m) = & (1 - I^m)[(\gamma_p + \delta_p I)^2 \sigma_{u_{pp}}^2 + 2(\gamma_p + \delta_p I)\rho_{u_{pp}, \epsilon_p} \sigma_{u_{pp}} \sigma_{\epsilon_p} + \sigma_{\epsilon_p}^2] \\
& + I^m[(\gamma_m + \delta_m I)^2 \sigma_{u_{mm}}^2 + 2(\gamma_m + \delta_m I)\rho_{u_{mm}, \epsilon_m} \sigma_{u_{mm}} \sigma_{\epsilon_m} + \sigma_{\epsilon_m}^2]. \tag{22}
\end{aligned}$$

The error term in  $\{\cdot\}$  is heteroscedastic depending on  $I$  and  $I^m$ , and consists of four errors. Define  $g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \zeta_{pp}, \zeta_{mm}) \equiv (1 - I^m)(\gamma_p \zeta'_{pp} Z + \delta_p I \zeta'_{pp} Z) + I^m(\gamma_m \zeta'_{mm} Z + \delta_m I \zeta'_{mm} Z)$ , which is the SF-2SLS regression function of the HC model in (21). Denote the first-stage LSE for  $\zeta_{pp}$ ,  $\sigma_{u_{pp}}$ ,  $\zeta_{mm}$ , and  $\sigma_{u_{mm}}$  in the  $z$ - and  $w$ - equations as  $\hat{\zeta}_{pp}$ ,  $\hat{\sigma}_{u_{pp}}$ ,  $\hat{\zeta}_{mm}$ , and  $\hat{\sigma}_{u_{mm}}$ . The resulting second-stage log-likelihood function for the  $y$  SF-2SLS of the HC model is

$$\begin{aligned}
L_h = \sum_i \{ & 1[y_i = 0] \ln \Phi\left(\frac{-g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \hat{\zeta}_{pp}, \hat{\zeta}_{mm})}{\sigma_{v_h, i}}\right) \\
& + 1[0 < y_i < \bar{c}_i] \ln \frac{\phi((y_i - g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \hat{\zeta}_{pp}, \hat{\zeta}_{mm}))/\sigma_{v_h, i})}{\sigma_{v_h, i}} \\
& + 1[y_i = \bar{c}_i] \ln \Phi\left(\frac{g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \hat{\zeta}_{pp}, \hat{\zeta}_{mm}) - \bar{c}_i}{\sigma_{v_h, i}}\right) \}, \tag{23}
\end{aligned}$$

which is to be maximized over  $\gamma_p$ ,  $\delta_p$ ,  $\sigma_{\epsilon_p}$ ,  $\gamma_m$ ,  $\delta_m$ ,  $\sigma_{\epsilon_m}$ ,  $\rho_{u_{pp}, \epsilon_p}$ , and  $\rho_{u_{mm}, \epsilon_m}$ , subject to the constraints:  $-\gamma_p + \gamma_m = 1$  and  $\delta_p = \delta_m$ . Denote the resulting estimates  $(\ddot{\gamma}_p, \ddot{\delta}_p, \ddot{\sigma}_{\epsilon_p}, \ddot{\gamma}_m, \ddot{\delta}_m, \ddot{\sigma}_{\epsilon_m}, \ddot{\rho}_{u_{pp}, \epsilon_p}, \ddot{\rho}_{u_{mm}, \epsilon_m})$ .

We explain in Appendix B the details of how to account for the effect of first-stage estimation errors on the second-stage asymptotic variance.

### 3.4 Hypothesis Testing

To assess the two competing models, we apply the *J-test* in Davidson and MacKinnon (1981) for non-nested hypotheses. Recall that  $f(Z, I; \gamma_p, \delta_p, \zeta_p)$  is the SF-2SLS regression function of the PC model in (18). Similarly, recall that  $g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \zeta_{pp}, \zeta_{mm})$  is the SF-2SLS regression function of the HC model in (21). Then, the J-test for the PC model as the null hypothesis proceeds as follows. First, create an artificially augmented regression function based on  $f$ :

$$\tilde{f} \equiv f(Z, I; \gamma_p, \delta_p, \zeta_p) + \mu_h \hat{g},$$

where  $\hat{g} = g(Z, I, I^m; \ddot{\gamma}_p, \ddot{\delta}_p, \ddot{\gamma}_m, \ddot{\delta}_m, \hat{\zeta}_{pp}, \hat{\zeta}_{mm})$  is the fitted value based on the parameter estimates from the HC model. Next, obtain the 2SLS-CMLE based on a modified likelihood function  $\tilde{L}_p$  as in (20) but with  $f(\cdot)$  replaced by the augmented regression function  $\tilde{f}$ . Test for  $\mu_h = 0$ . If  $\mu_h = 0$  is rejected, the PC model is rejected to the direction of the HC model.

The J-test for the HC model as the null hypothesis can be done analogously. Form an artificially augmented regression function based on  $g$ :

$$\tilde{g} \equiv g(Z, I, I^m; \gamma_p, \delta_p, \gamma_m, \delta_m, \zeta_{pp}, \zeta_{mm}) + \mu_p \hat{f},$$

where  $\hat{f} = f(Z, I; \dot{\gamma}_p, \dot{\delta}_p, \hat{\zeta}_p)$  is the fitted value based on the parameter estimates from the PC model. Obtain the 2SLS-CMLE based on a modified likelihood function  $\tilde{L}_h$  as in (23) but with  $g(\cdot)$  replaced by the augmented regression function  $\tilde{g}$ . Test for  $\mu_p = 0$ . If  $\mu_p = 0$  is rejected, the HC model is rejected to the direction of the PC model.

### 3.5 Result

In Table 1, we present the search process for the optimal classification of sectors for the HC model. As introduced in Section 3.1, the market structure dummy is defined as follows: a sector is classified as monopolistically competitive, if the elasticity measure is greater than one and if the sector faces a relatively low degree of seller competition ( $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq \bar{\kappa}$ ). The first column

of Table 1 lists the cutoff thresholds  $\bar{\kappa}$  that we explore and the second column their corresponding rankings among all observed values of seller competition. We begin by setting  $\bar{\kappa}$  at the highest observed value of seller competition ( $\bar{\kappa} = 1.3865$ ). This is equivalent to imposing no constraint on seller competition, and sectors are classified based on the magnitude of elasticity alone. In this case, there are 59 sectors (out of 106 sectors) with elasticity larger than one and classified as monopolistically competitive. The HC model is estimated based on the resulting classification of sectors, and the political parameter estimates and the likelihood value are reported in the fourth column. We then lower the cutoff threshold to the next largest observed value of seller competition and re-estimate the HC model. This is repeated for the rest of observed values of seller competition in a descending order. The number of sectors classified as monopolistically competitive weakly decreases as the cutoff threshold lowers. We stop the search at the point when only 35 sectors are classified as monopolistically competitive. This corresponds to  $\bar{\kappa} = 0.1873$ . The lower bound on the number of classified monopolistically competitive sectors is chosen based on the consideration that there are 22 parameters to be estimated for each market structure in the reduced form equations for  $z$  and  $w$  and that some reasonable degree of freedom has to be maintained. In addition, as the criteria become too stringent for a sector to be classified as monopolistically competitive, few monopolistically competitive sectors are left under the HC model. In this case, the HC model approaches the PC model, and it will not be such an informative exercise to compare the two. Table 1 indicates that as the cutoff threshold lowers, the likelihood of the HC model first rises (albeit not strictly monotonically) and then falls. The maximum likelihood of the HC model is found in the interior of the search domain at  $\bar{\kappa} = 0.2513$ . This corresponds to a specification of the market structure dummy according to:  $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq 0.2513$ .

In Table 2, we present the details of the estimation results for the case where the HC model is estimated based on the optimal specification of the market structure dummy. The PC model is independent of the market structure dummy and is estimated for all sectors. The summary statistics show that all sectors on average face an import demand elasticity of 2.4443 and a nontariff barrier coverage ratio of 0.1350. Among all sectors, less than half (0.4245) are politically organized for trade policy purposes. Based on the PC model, the resulting parameter estimates of  $\gamma_p$  and  $\delta_p$  are of correct signs in accordance with theory (13) and are significant. The estimate of  $\gamma_p + \delta_p$  is also of correct sign, but is not significant. According to these estimates, the weight of welfare in the

government's objective ( $\beta \equiv \frac{a}{1+a}$ ) is 0.9890, while the fraction of population represented by a lobby ( $\alpha_L$ ) is 0.7981. The results are broadly consistent with those of Goldberg and Maggi (1999).

For the HC model, based on the optimal specification of the market structure dummy:  $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq 0.2513$ , there are 41 monopolistically competitive sectors and 65 perfectly competitive sectors. Monopolistically competitive sectors show a theoretically-desired lower degree of seller competition (0.1040) than perfectly competitive sectors (0.3233). The summary statistics also indicate that on average, monopolistically competitive sectors face higher import demand elasticity, are more likely to be politically organized, and receive higher levels of import protection. Based on the HC model, the parameter estimates of  $\gamma_p$ ,  $\delta_p$ ,  $\gamma_m$ , and  $\delta_m$  are of correct signs consistent with Proposition 1. The estimate of  $\gamma_p + \delta_p$  (and correspondingly that of  $\gamma_m + \delta_m$ ) is greater than zero (greater than one) as predicted by Proposition 1. However, except for  $\gamma_m$ , the estimates are not statistically significant. This may be due to the fact that the degree of freedom is greatly reduced in the HC model with the sample broken into two regimes. Based on the above estimates, the weight of welfare in the government's objective ( $\beta \equiv \frac{a}{1+a}$ ) is 0.9989 and the fraction of population politically represented ( $\alpha_L$ ) is 0.5856.

The higher weight of welfare in the government's objective implied by the HC model, relative to the PC model, may be explained by the fact that the welfare-maximizing tariff for a monopolistically competitive sector is positive. Thus, even if a government were not politically motivated (with  $\beta = 1$ ), a monopolistically competitive sector would still receive a positive level of protection. Thus, for a given level of observed protection, it would imply a higher weight of political contribution (and a lower weight of welfare) placed by the government if the sector is classified as perfectly competitive than if it is classified as monopolistically competitive.

Given that both models' parameter estimates have correct signs, we evaluate their statistical goodness of fit based on the J-test discussed in Section 3.4. The results in Table 2 show that the PC model as the null is rejected at 1% significance level while the HC model with the optimal specification of the market structure dummy is accepted. Based on the J-test results, we conclude that the HC model given the optimal classification of sectors provides a statistically-significant better fit to the data than the PC model.

One might question that how robust the result is if the market structure dummy is not optimally specified for the HC model. As indicated in the last column of Table 1, the PC model is consistently

rejected regardless of how the market structure dummy is specified for the HC model. On the other hand, the HC model is accepted against the PC model provided that the specification of the market structure dummy is not too far off from the optimal specification. The HC model is accepted if the market structure dummy is specified as:  $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq \bar{\kappa}$ , for  $0.2513 \leq \bar{\kappa} \leq 0.3233$ . Overall, the above results cast doubt on the PC model as a complete model and suggest a direction of improvement toward the HC model with a properly specified market structure dummy.

The estimates of the HC model in Table 2 suggest a higher weight of welfare in the government's objective ( $\beta = 0.9989$ ) than indicated by previous studies based on G-H (0.986 in Goldberg and Maggi, 1999; 0.96 in Eicher and Osang, 2002). This is consistent with our earlier discussion that misspecifying a monopolistically competitive sector as perfectly competitive tends to assign the supposedly welfare-driven component of protection as politically motivated, and to bias upward (downward) the estimate of the government's weight on political contribution (aggregate welfare). The estimates of the degree of political representation ( $\alpha_L$ ) vary a lot in the literature, ranging from 0.98 in Gawande and Bandyopadhyay (2000), 0.88 in Goldberg and Maggi (1999), to 0.26 in Eicher and Osang (2002). Our estimate of  $\alpha_L$  at 0.5856 suggests an intermediate degree of political representation in the trade policy arena.

## 4. CONCLUSION

Previous cross-sectional empirical studies based on Grossman and Helpman (1994) have often adopted homogeneous (perfect competition) market structure, where all sectors are taken to be perfectly competitive and share the same protection structure according to G-H. In this paper, we propose a general empirical specification that accommodates both monopolistically and perfectly competitive sectors, with the protection structure for monopolistically competitive sectors guided by the recent theory of Chang (2005). The results of this paper cast doubt on the protection-for-sale model with homogeneous perfect-competition market structure and suggest a direction of improvement toward the proposed protection-for-sale model with heterogeneous market structure. This empirical finding implies that the Chang (2005) model is borne out by the data for monopolistically competitive sectors; meanwhile, the predictions of G-H remain valid for the subset of perfectly competitive sectors. The finding also suggests that the response of endogenous protection

to import penetration is heterogeneous: it depends on the market structure of a sector (whether the sector is perfectly or monopolistically competitive) and its state of political organization (whether or not the sector is politically organized). When all of these aspects of heterogeneity are taken into account, the resulting estimates imply that the government's weight on aggregate welfare is higher than suggested by previous literature (with similar estimation framework) and the government is close to being a welfare maximizer. On the other hand, the degree of political representation falls in the intermediate range; around half of the population is politically represented in the trade policy arena.

A caveat is warranted. The sample studied here, as well as in previous G-H-type empirical work, has focused mainly on manufacturing sectors where the general protection level is low, and has excluded the agriculture sector where heavy protection persists. With the agriculture sector included, the conclusions are likely to change toward finding a more special-interest driven government.

## APPENDIX A: ESTIMATION METHODOLOGY FOR TOBIT MODEL WITH ENDOGENOUS REGRESSORS

In this appendix, we describe the estimators we explored in our preliminary analysis. The notations here differ somewhat from those in the main text for the sake of simplification. This appendix will also serve as a review on the possible estimation methods employed in previous G-H-type empirical studies. Define an indicator function  $1[A] = 1$  if  $A$  holds and 0 otherwise. Let SF and RF stand for 'structural form' and 'reduced form', respectively. Suppose the equations under consideration are, with  $c > 0$ ,

$$\begin{aligned} y \text{ SF} &: y_i = \max\{0, \min(y_i^*, c)\}, \quad \text{where } y^* \text{ SF is } y_i^* = \beta_w w_i + \beta_{dw} d_i w_i + u_i, \\ w_i &= z_i' \eta_w + v_{wi}, \quad \text{where } v_{wi} \text{ and } u_i \text{ are dependent} \\ d_i &= 1[d_i^* > 0], \quad \text{where } d_i^* = z_i' \eta_d + v_{di}, \quad v_{di} \text{ and } u_i \text{ are dependent,} \end{aligned} \tag{Model 1}$$

observed :  $d_i, z_i, w_i, y_i, \quad i = 1, \dots, N, \quad iid \text{ across } i.$

The goal is to estimate the SF parameters  $\beta_w$  and  $\beta_{dw}$ . The upper censoring point  $c$  can be allowed to vary across  $i$  so long as  $c_i$  is observed and independent of  $y_i$ . A simpler version of Model 1 is

obtained if  $d$  is exogenous:

same as Model 1 but  $d$  is exogenous under independence between  $v_d$  and  $u$ . (Model 2)

In this case, the slope of  $w$  in  $y^* = (\beta_w + \beta_{dw}d)w + u$  shifts exogenously depending on  $d$ . Substituting the  $w$  equation into the  $y^*$  SF, we get the  $y^*$  RF:

$$y^* \text{ RF} : y^* = z'(\beta_w \eta_w) + dz'(\beta_{dw} \eta_w) + \{(\beta_w + \beta_{dw}d)v_w + u\}$$

where the error term in  $\{\cdot\}$  is heteroscedastic depending on  $d$  and consists of two errors.

Write the  $y^*$  RF simply as  $y_i^* = x_i' \alpha + \varepsilon_i$ , such that  $y_i = \max\{0, \min(x_i' \alpha + \varepsilon_i, c)\}$ , where  $x_i$  is the exogenous regressor vector,  $\alpha$  is the parameter vector, and  $\varepsilon_i$  is the error term. Under the independence of  $\varepsilon$  from  $x$  and the normality  $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ , the *Censored MLE (CMLE)* is obtained with the log-likelihood function

$$\sum_i \{1[y_i = 0] \ln \Phi\left(\frac{-x_i' \alpha}{\sigma_\varepsilon}\right) + 1[0 < y_i < c] \ln \left\{ \frac{\phi((y_i - x_i' \alpha)/\sigma_\varepsilon)}{\sigma_\varepsilon} \right\} + 1[y_i = c] \ln \Phi\left(\frac{x_i' \alpha - c}{\sigma_\varepsilon}\right)\}$$

which is maximized with respect to (wrt)  $\alpha$  and  $\sigma_\varepsilon$ . CMLE converges reasonably well.

Denoting the conditional median of  $\varepsilon|x$  as  $Med(\varepsilon|x)$ , a semiparametric estimator requiring only  $Med(\varepsilon|x) = 0$  while allowing an unknown form of heteroskedasticity is the *Censored Least Absolute Deviation estimator (CLAD)* of Powell (1984). For the double censoring case, CLAD minimizes wrt  $\alpha$

$$\sum_i |y_i - \max\{0, \min(x_i' \alpha, c)\}|.$$

The requisite assumption for CLAD is weaker than that for CMLE. Also, no ‘nuisance parameter’ such as  $\sigma_\varepsilon$  appears for CLAD.

Eicher and Osang (2002) apply *Minimum Distance Estimator (MDE)* as explained in Lee (1996a). But MDE requires a somewhat different model from above: instead of the  $d$  equation, suppose

$$dw \text{ equation} : d_i w_i = z_i' \eta_{dw} + v_{dwi}.$$

Then, substituting the above  $w$  equation and this  $dw$  equation into the above  $y^*$  SF yields

$$\begin{aligned}
y^* \text{ RF-MDE} & : y_i^* = \beta_w(z_i'\eta_w + v_{wi}) + \beta_{dw}(z_i'\eta_{dw} + v_{dwi}) + u_i \equiv z_i'\eta_y + v_y, \quad \text{where} \\
v_y & \equiv \beta_w v_{wi} + \beta_{dw} v_{dwi} + u_i \quad \text{and} \\
\text{MDE Restriction} & : \eta_y = \beta_w \eta_w + \beta_{dw} \eta_{dw}.
\end{aligned}$$

This  $y^*$  RF in turn leads to

$$y \text{ RF-MDE} : y_i = \max\{0, \min(z_i'\eta_y + v_y, c)\}.$$

The MDE proceeds in two steps. First, estimate  $\eta_y$  for the  $y$  RF-MDE with a censored model estimator, and  $\eta_w$  and  $\eta_{dw}$  with Least Squares Estimator (LSE); denote the estimators as  $h_y$ ,  $h_w$ , and  $h_{dw}$ , respectively. Second, do the LSE of  $h_y$  on  $h_w$  and  $h_{dw}$  to estimate the two scalars  $\beta_w$  and  $\beta_{dw}$ , which works owing to the MDE Restriction above.

The MDE procedure works well in practice. But the shortcoming is the linear model assumption for  $dw$ , which is not tenable in principle because  $dw$  is not a continuously distributed random variable due to  $P(d = 0) = P(dw = 0) > 0$ . Recognizing this problem, one may apply a little different MDE: use a censored model

$$d_i w_i = \max(0, z_i'\eta_{dw} + v_{dwi})$$

and estimate this with a CMLE. This should provide a better approximation for  $dw$ , but leads to a new problem that the MDE restriction—and consequently  $y$  RF-MDE—holds only on  $d = 1$ . Estimating the  $y$  RF-MDE using only the subsample  $d = 1$  causes the usual sample selection problem if  $d$  is endogenous. If  $d$  is exogenous, then the MDE procedure works so long as the estimation of the  $y$  RF-MDE is done with the  $d = 1$  subsample. That is, the MDE procedure is tenable under Model 2.

Lee (1996b) proposes a *Two-Stage-LSE (2SLS)* type approach for limited dependent variable models with endogenous regressors. The idea is to replace each endogenous regressor with its



conditional mean given the instruments. For this, rewrite the  $y^*$  SF as

$$\begin{aligned} y^* \text{ SF-2SLS-1} \quad : \quad y_i^* &= \beta_w E(w|z_i) + \beta_{dw} E(dw|z_i) \\ &+ u_i + \beta_w \{w_i - E(w|z_i)\} + \beta_{dw} \{d_i w_i - E(dw|z_i)\} \end{aligned}$$

where the last three terms constitute the new (composite) error term. The conditional means  $E(w|z)$  and  $E(dw|z)$  can be estimated nonparametrically. In practice, the LSE for  $w_i = z_i' \eta_w + v_{wi}$  and  $d_i w_i = z_i' \eta_{dw} + v_{dwi}$  will do. The latter (practical) approach has the same problem as the above MDE, because the linear model for the product  $dw$  is not tenable in principle.

The 2SLS approach can be similarly applied to Model 2. Rewrite the  $y^*$  SF as

$$\begin{aligned} y^* \text{ SF-2SLS-2} \quad : \quad y_i^* &= \beta_w E(w|z_i) + \beta_{dw} d_i E(w|z_i) \\ &+ u_i + \beta_w \{w_i - E(w|z_i)\} + \beta_{dw} d_i \{w_i - E(w|z_i)\} \end{aligned}$$

where the last three terms constitute the new (composite) error term. The conditional mean  $E(w|z)$  can be estimated using the LSE for  $w_i = z_i' \eta_w + v_{wi}$ . Hence the problem mentioned above wrt  $y^*$  RF-MDE and  $y^*$  SF-2SLS-1 does not arise in this case. The  $y^*$  SF-2SLS leads straightforwardly to the  $y$  SF-2SLS, to which CMLE or CLAD can be applied.

Considering our discussions so far, we can think of at least four methods to estimate  $\beta_d$  and  $\beta_{dw}$ :

|      | CMLE                             | CLAD                               |
|------|----------------------------------|------------------------------------|
| MDE  | $y$ RF estimated, normality      | $y$ RF estimated, zero median      |
| 2SLS | $y$ SF-2SLS estimated, normality | $y$ SF-2SLS estimated, zero median |

Other than the approaches listed in the above table, one may also attempt to apply the system MLE to Model 1 under  $(u, v_w, v_d) \sim N(0, \Omega)$ , where the covariance matrix  $\Omega$  has  $SD(u)$ ,  $SD(v_w)$ ,  $COR(u, v_w)$ ,  $COR(u, v_d)$ , and  $COR(v_w, v_d)$ ;  $SD(v_d)$  is not identified. But this MLE has two problems. First, with  $d$  endogenous, it seems difficult to obtain the likelihood function. Second, even if the likelihood function is found, estimating the correlation coefficients (and the standard deviations) is notoriously difficult in multivariate MLE's. In practice, often some correlations are assumed to be zero (e.g.,  $COR(v_w, v_d) = 0$ ), or an equi-correlation assumption is imposed (e.g.,

$COR(u, v_w) = COR(u, v_d)$ ). It is not clear how Goldberg and Maggi (1999) proceeded, as the likelihood function and the estimate for  $\Omega$  for their MLE were not shown. Alternatively, if Model 2 is adopted, the likelihood function becomes straightforward, with the size of the covariance matrix greatly reduced. Overall, in view of the econometrics problems discussed above in implementing Model 1, we choose to adopt Model 2 in the main text.

## APPENDIX B: ESTIMATION METHODOLOGY FOR THE STANDARD ERRORS OF THE 2SLS-CMLE

Let  $\alpha$  be the first stage parameter of dimension  $k_1 \times 1$  and  $a_N$  be the LSE. Let  $\beta$  be the likelihood parameter of dimension  $k_2 \times 1$  for the second stage, and  $b_N$  be the MLE. Denote the second stage score function as  $s(z_i, \alpha, \beta)$ ; omit  $z_i$  for simplification. Define

$$\nabla_{\alpha} s(\alpha, \beta) \equiv \frac{\partial s(\alpha, \beta)}{\partial \alpha'} \quad \text{and} \quad \nabla_{\beta} s(\alpha, \beta) \equiv \frac{\partial s(\alpha, \beta)}{\partial \beta'}.$$

By the definition of  $b_N$ , it holds that, using Taylor's expansion,

$$\begin{aligned} 0 &= \frac{1}{\sqrt{N}} \sum_i s(a_N, b_N) \\ \implies 0 &\simeq \frac{1}{\sqrt{N}} \sum_i s(a_N, \beta) + \left\{ \frac{1}{N} \sum_i \nabla_{\beta} s(\alpha, \beta) \right\} \sqrt{N}(b_N - \beta) \\ \implies \sqrt{N}(b_N - \beta) &= -\left\{ \frac{1}{N} \sum_i \nabla_{\beta} s(\alpha, \beta) \right\}^{-1} \frac{1}{\sqrt{N}} \sum_i s(a_N, \beta) \\ \implies \sqrt{N}(b_N - \beta) &= -\left\{ \frac{1}{N} \sum_i s(\alpha, \beta) s(\alpha, \beta)' \right\}^{-1} \frac{1}{\sqrt{N}} \sum_i s(a_N, \beta). \end{aligned}$$

To account for the first-stage error  $a_N - \alpha$ , define

$$H_N \equiv \frac{1}{N} \sum_i \nabla_{\beta} s(\alpha, \beta) = \frac{1}{N} \sum_i s(\alpha, \beta) s(\alpha, \beta)' \quad \text{and} \quad L_N \equiv \frac{1}{N} \sum_i \nabla_{\alpha} s(\alpha, \beta).$$

Apply Taylor's expansion to  $s(a_N, \beta)$  in the above expression for  $\sqrt{N}(b_N - \beta)$  to get

$$\sqrt{N}(b_N - \beta) \simeq -H_N^{-1} \cdot \left\{ \frac{1}{\sqrt{N}} \sum_i s(\alpha, \beta) + L_N \sqrt{N}(a_N - \alpha) \right\}.$$

With  $r_i$  denoting the first-stage LSE residual with  $Z_i$  as the regressor, observe

$$\begin{aligned}\sqrt{N}(a_N - \alpha) &= \frac{1}{\sqrt{N}} \sum_i \left( \frac{1}{N} \sum_i Z_i Z_i' \right)^{-1} Z_i r_i \\ &\equiv \frac{1}{\sqrt{N}} \sum_i \eta_i, \quad \text{where } \eta_i \equiv \left( \frac{1}{N} \sum_i Z_i Z_i' \right)^{-1} Z_i r_i.\end{aligned}$$

Hence

$$\begin{aligned}\sqrt{N}(b_N - \beta) &\simeq -H_N^{-1} \left\{ \frac{1}{\sqrt{N}} \sum_i s(\alpha, \beta) + L_N \frac{1}{\sqrt{N}} \sum_i \eta_i \right\} \\ &= -H_N^{-1} \frac{1}{\sqrt{N}} \sum_i \{s(\alpha, \beta) + L_N \eta_i\} \\ &= -H_N^{-1} \frac{1}{\sqrt{N}} \sum_i q_i, \quad \text{where } q_i \equiv s(\alpha, \beta) + L_N \eta_i.\end{aligned}$$

Therefore, with  $\rightsquigarrow$  denoting convergence in law,

$$\sqrt{N}(b_N - \beta) \rightsquigarrow N(0, H^{-1} E(qq') H^{-1}) \quad \text{where } H \equiv E\{s(\alpha, \beta)s(\alpha, \beta)'\}.$$

Consistent estimators for  $H$  and  $E(qq')$  are

$$\begin{aligned}\hat{H}_N &\equiv \frac{1}{N} \sum_i s(a_N, b_N) s(a_N, b_N)', \\ Q_N &\equiv \frac{1}{N} \sum_i \{s(a_N, b_N) + \hat{L}_N \eta_i\} \{s(a_N, b_N) + \hat{L}_N \eta_i\}' \quad \text{where } \hat{L}_N \equiv \frac{1}{N} \sum_i \nabla_\alpha s(a_N, b_N).\end{aligned}$$

In practice,  $s$  can be obtained by numerical derivatives, and  $\nabla_\alpha s$  can be obtained using numerical derivatives once more. If the first-stage LSE involves two equations, each with regressor  $Z_{ji}$ , residual  $r_{ji}$ ,  $\eta_{ji}$ , and the parameter  $\alpha_j$ ,  $j = 1, 2$ , then set  $\alpha \equiv (\alpha'_1, \alpha'_2)'$ ,  $a_N \equiv (a'_{1N}, a'_{2N})'$ ,  $\eta_i \equiv (\eta'_{1i}, \eta'_{2i})'$  and proceed as above.

## REFERENCES

Bernheim, B. D., Whinston, M. D., 1986. Menu auctions, resource allocation, and economic influence. *Quarterly Journal of Economics* 101 (1), 1–31.

- Chang, P.-L., 2005. Protection for sale under monopolistic competition. *Journal of International Economics* 66, 509–526.
- Davidson, R., MacKinnon, J. G., 1981. Several tests for model specification in the presence of alternative hypothesis. *Econometrica* 49 (3), 781–93.
- Dixit, A. K., Stiglitz, J. E., 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67 (3), 297–308.
- Eicher, T., Osang, T., 2002. Protection for sale: An empirical investigation: Comment. *American Economic Review* 92 (5), 1702–1710.
- Fair, R. C., Jaffee, D. M., 1972. Methods of estimation for markets in disequilibrium. *Econometrica* 40 (3), 497–514.
- Gawande, K., Bandyopadhyay, U., 2000. Is protection for sale? Evidence on the Grossman-Helpman theory of endogenous protection. *Review of Economics and Statistics* 82 (1), 139–152.
- Goldberg, P. K., Maggi, G., 1999. Protection for sale: An empirical investigation. *American Economic Review* 89 (5), 1135–1155.
- Grossman, G. M., Helpman, E., 1994. Protection for sale. *American Economic Review* 84 (4), 833–850.
- Hartley, M. J., 1978. Estimating mixtures of normal distributions and switching regressions: Comment. *Journal of the American Statistical Association* 73 (364), 738–741.
- Helpman, E., 1999. The structure of foreign trade. *Journal of Economic Perspectives* 13 (2), 121–144.
- Kiefer, N. M., 1978. Discrete parameter variation: Efficient estimation of a switching regression model. *Econometrica* 46 (2), 427–434.
- Lee, M.-J., 1996a. *Methods of moments and semiparametric econometrics for limited dependent variable models*. Springer Verlag, New York.
- Lee, M.-J., 1996b. Nonparametric two stage estimation of simultaneous equations with limited endogenous regressor. *Econometric Theory* 12, 305–330.

- Phillips, R. F., 1991. A constrained maximum-likelihood approach to estimating switching regressions. *Journal of Econometrics* 48, 241–262.
- Powell, J. L., 1984. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* 25, 303–325.
- Quandt, R. E., 1972. A new approach to estimating switching regressions. *Journal of the American Statistical Association* 67 (338), 306–310.
- Quandt, R. E., Ramsey, J. B., 1978. Estimating mixtures of normal distributions and switching regressions. *Journal of the American Statistical Association* 73 (364), 730–738.
- 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 (3), 497–519.
- Trefler, D., 1993. Trade liberalization and the theory of endogenous protection. *Journal of Political Economy* 101 (1), 138–160.

Table 1: The HC model estimates conditional on the classification of sectors

| classification of sectors                           |                                    |                      | HC model estimates |            |            |          |          | J-test |      |
|---|------------------------------------|----------------------|--------------------|------------|------------|----------|----------|--------|------|
| cutoff<br>value<br>of<br>$scomp$ ( $\bar{\kappa}$ ) | cutoff<br>ranking<br>of<br>$scomp$ | no. of MC<br>sectors | a                  |            | $\alpha_L$ |          | $L_h$    | PC     | HC   |
| 1.3865  | 106                                | 59                   | 849.0493           | (740.7607) | 0.4545     | (0.4453) | -78.3968 | R**    | R**  |
| 1.3222  | 105                                | 59                   | 849.0493           | (740.7607) | 0.4545     | (0.4453) | -78.3968 | R**    | R**  |
| 1.2246  | 104                                | 59                   | 849.0493           | (740.7607) | 0.4545     | (0.4453) | -78.3968 | R**    | R**  |
| 1.2014  | 103                                | 58                   | 830.3585           | (669.8679) | 0.5210     | (0.4391) | -77.9522 | R**    | R**  |
| 0.9717  | 102                                | 57                   | 810.7296           | (616.9224) | 0.5327     | (0.4179) | -77.3769 | R**    | R**  |
| 0.9716  | 101                                | 56                   | 857.9891           | (657.7732) | 0.4926     | (0.4155) | -75.1774 | R**    | R**  |
| 0.6550  | 100                                | 56                   | 857.9891           | (657.7732) | 0.4926     | (0.4155) | -75.1774 | R**    | R**  |
| 0.6397  | 99                                 | 55                   | 858.8432           | (659.0478) | 0.4914     | (0.4155) | -74.9309 | R**    | R**  |
| 0.6289  | 98                                 | 54                   | 851.6470           | (673.4271) | 0.4464     | (0.4140) | -74.7138 | R**    | R**  |
| 0.5953  | 97                                 | 53                   | 909.7454           | (760.2341) | 0.4318     | (0.4343) | -72.7223 | R**    | R**  |
| 0.5896  | 96                                 | 52                   | 874.5466           | (695.5136) | 0.4379     | (0.4181) | -72.4156 | R**    | R**  |
| 0.5628  | 95                                 | 51                   | 661.9317           | (433.5991) | 0.4787     | (0.3519) | -73.5389 | R**    | R*** |
| 0.5429  | 94                                 | 51                   | 661.9317           | (433.5991) | 0.4787     | (0.3519) | -73.5389 | R**    | R*** |
| 0.4926  | 93                                 | 50                   | 696.3284           | (472.3664) | 0.4718     | (0.3636) | -71.3375 | R**    | R*** |
| 0.4868  | 92                                 | 50                   | 696.3284           | (472.3664) | 0.4718     | (0.3636) | -71.3375 | R**    | R*** |
| 0.4619  | 91                                 | 50                   | 696.3284           | (472.3664) | 0.4718     | (0.3636) | -71.3375 | R**    | R*** |
| 0.4380  | 90                                 | 49                   | 727.0918           | (510.5308) | 0.4718     | (0.3789) | -69.7894 | R**    | R**  |
| 0.4262  | 89                                 | 48                   | 763.1615           | (555.1879) | 0.4624     | (0.3892) | -67.4248 | R**    | R**  |
| 0.4220  | 88                                 | 47                   | 807.7733           | (629.8685) | 0.4926     | (0.4354) | -65.1257 | R**    | R**  |
| 0.3668  | 87                                 | 47                   | 807.7733           | (629.8685) | 0.4926     | (0.4354) | -65.1257 | R**    | R**  |
| 0.3549  | 86                                 | 46                   | 802.1698           | (630.5880) | 0.4672     | (0.4306) | -63.3556 | R**    | R*** |
| 0.3507  | 85                                 | 46                   | 802.1698           | (630.5880) | 0.4672     | (0.4306) | -63.3556 | R**    | R*** |
| 0.3233  | 84                                 | 45                   | 825.5517           | (670.0232) | 0.5107     | (0.4711) | -61.0028 | R***   | A    |
| 0.3229  | 83                                 | 44                   | 869.5347           | (594.5170) | 0.5537     | (0.4182) | -59.7069 | R***   | A    |
| 0.3115  | 82                                 | 43                   | 903.2278           | (637.4129) | 0.5463     | (0.4303) | -57.5475 | R**    | A    |
| 0.2973  | 81                                 | 43                   | 903.2278           | (637.4129) | 0.5463     | (0.4303) | -57.5475 | R**    | A    |
| 0.2920  | 80                                 | 43                   | 903.2278           | (637.4129) | 0.5463     | (0.4303) | -57.5475 | R**    | A    |
| 0.2623  | 79                                 | 42                   | 920.0701           | (664.4405) | 0.5341     | (0.4374) | -57.1308 | R***   | A    |
| 0.2594  | 78                                 | 42                   | 920.0701           | (664.4405) | 0.5341     | (0.4374) | -57.1308 | R***   | A    |
| 0.2528  | 77                                 | 42                   | 920.0701           | (664.4405) | 0.5341     | (0.4374) | -57.1308 | R***   | A    |
| 0.2513  | 76                                 | 41                   | 888.3396           | (593.6911) | 0.5856     | (0.4136) | -55.9352 | R***   | A    |
| 0.2382  | 75                                 | 40                   | 494.6710           | (237.9385) | 0.5536     | (0.2767) | -67.8295 | R***   | R*** |
| 0.2353  | 74                                 | 39                   | 443.6104           | (184.3986) | 0.6248     | (0.2547) | -66.7684 | R**    | R*** |
| 0.2335  | 73                                 | 39                   | 443.6104           | (184.3986) | 0.6248     | (0.2547) | -66.7684 | R**    | R*** |
| 0.2332  | 72                                 | 38                   | 435.6234           | (172.8254) | 0.6088     | (0.2406) | -64.9324 | R***   | R*** |
| 0.2125  | 71                                 | 38                   | 435.6234           | (172.8254) | 0.6088     | (0.2406) | -64.9324 | R***   | R*** |
| 0.2105  | 70                                 | 38                   | 435.6234           | (172.8254) | 0.6088     | (0.2406) | -64.9324 | R***   | R*** |
| 0.2029  | 69                                 | 38                   | 435.6234           | (172.8254) | 0.6088     | (0.2406) | -64.9324 | R***   | R*** |
| 0.2011  | 68                                 | 37                   | 432.2619           | (172.2964) | 0.6056     | (0.2401) | -63.8470 | R***   | R*** |
| 0.2010  | 67                                 | 36                   | 427.5867           | (169.9796) | 0.6374     | (0.2434) | -62.8964 | R***   | R*** |
| 0.1885  | 66                                 | 36                   | 427.5867           | (169.9796) | 0.6374     | (0.2434) | -62.8964 | R***   | R*** |
| 0.1873  | 65                                 | 35                   | 416.3748           | (165.2915) | 0.6419     | (0.2449) | -62.4645 | R***   | R*** |

Note:

1. The variable ‘seller competition ( $scomp$ )’ is the number of seller firms in an industry divided by total industry sales. A sector is classified as monopolistically competitive (MC) if  $e_i > 1$  and  $scomp_i \leq \bar{\kappa}$ .
2. Figures in parentheses are standard errors.
3. R\* indicates rejection at the 10% significance level, R\*\* indicates rejection at the 5% significance level, and R\*\*\* indicates rejection at the 1% significance level. ‘A’ indicates acceptance at least at the 10% significance level.

Table 2: classification of sectors based on elasticity and seller competition  
( $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq 0.2513$ )

|   | PC             |           | HC        |            |           |
|---|----------------|-----------|-----------|------------|-----------|
| $\gamma_p$  | -0.0088***     | (0.0032)  | -0.0007   | (0.0004)   |           |
| $\delta_p$  | 0.0110*        | (0.0059)  | 0.0011    | (0.0008)   |           |
| $\sigma_{\epsilon_p}$                               | 0.6693         | (0.1037)  | 0.1193    | (0.0148)   |           |
| $\rho_{u_p, \epsilon_p}(\rho_{u_{pp}, \epsilon_p})$ | -0.9149        | (0.6486)  | -1        |            |           |
| $\gamma_m$  |                |           | 0.9993*** | (0.0004)   |           |
| $\delta_m$  |                |           | 0.0011    | (0.0008)   |           |
| $\sigma_{\epsilon_m}$                               |                |           | 1.3661    | (0.0572)   |           |
| $\rho_{u_{mm}, \epsilon_m}$                         |                |           | -1        |            |           |
| $\gamma_p + \delta_p$                               | 0.0022         | (0.0045)  | 0.0005    | (0.0007)   |           |
| $\gamma_m + \delta_m - 1$                           |                |           | 0.0005    | (0.0007)   |           |
| $a$   | 90.1326*       | (48.1782) | 888.3396  | (593.6911) |           |
| $\alpha_L$  | 0.7981**       | (0.3271)  | 0.5856    | (0.4136)   |           |
| $L_p$ v.s. $L_h$                                    | -99.0480       |           | -55.9352  |            |           |
| J test:   |                |           |           |            |           |
| $\mu_h$ v.s. $\mu_p$                                | 0.7454***      | (0.2797)  | 0.4192    | (0.5138)   |           |
| $\tilde{L}_p$ v.s. $\tilde{L}_h$                    | -92.8236       |           | -53.4008  |            |           |
| Summary Statistics:                                 | no. of sectors | $\bar{e}$ | $scomp$   | $\bar{I}$  | $\bar{t}$ |
| $I_i^m = 1$   | 41             | 4.5843    | 0.1040    | 0.5122     | 0.1561    |
| $I_i^m = 0$   | 65             | 1.0944    | 0.3233    | 0.3692     | 0.1217    |
| Total   | 106            | 2.4443    | 0.2385    | 0.4245     | 0.1350    |

Note:

1.  $I_i^m = 1$ , if  $e_i > 1$  and  $scomp_i \leq 0.2513$ , and  $I_i^m = 0$  otherwise. The variable ‘seller competition ( $scomp$ )’ is the number of seller firms in an industry divided by total industry sales.
2. Figures in parentheses are standard errors.
3. A \* sign indicates significance at 10% level, a \*\* sign indicates significance at 5% level, and a \*\*\* sign indicates significance at 1% level.
4. For parameters,  $\sigma_{\epsilon_p}$  and  $\sigma_{\epsilon_m}$ , which are not tested against zero, we do not attach the \* sign.
5. The estimates of  $\rho_{u_{pp}, \epsilon_p}$  and  $\rho_{u_{mm}, \epsilon_m}$  in the HC model reach the lower bound  $-1$ . Note that  $\rho = -1$  does not make the likelihood function of the HC model degenerate as it would do in a bivariate normal density function.