Habitual and Occasional Lobbyers in the US Steel Industry: An EM Algorithm Pooling Approach

by

Randall Morck*, Jungsywan Sepanski**, and Bernard Yeung***

First Draft: July 19, 1995

Latest Draft: Sept 29th, 1999

Economic Inquiry, Vol. 39, No. 3, July, 2001, 365-378.

JEL classification: F13

* Stephen J. Janislowsky Distinguished Chair in Finance, Faculty of Business, University of Alberta, Edmonton, Alberta, Canada. T6G 2R6, email: randall.morck@ualberta.ca ph. 780 492 5683, Fax: 780 492 3325

** Assistant Professor, Department of Mathematics, Central Michigan University, email: sepan1jh@cmich.edu Ph. 517 774 3508, Fax: 517 774 2414

*** Abraham Krasnoff Professor of International Business and Professor of Economic, Stern School of Business, New York University, 44 W 4th St. Rm. 7-65, New York New York, 10012, email: byeung@stern.nyu.edu, Ph. 212 998 0425 Fax: 212 995 4221

Bernard Yeung's research is partly supported by the Center for International Business Education at the University of Michigan Business School. The authors are grateful to Simon Evenett and Joanne Oxley for their helpful comments.

Abstract

Using U.S. steel firm data, we find that lobbying for import protection appears to be habit-forming. To identify heterogeneity in lobbying behavior among firms, we use an Expectation-Maximization algorithm to sort our firms into groups with different propensities to lobby and estimate the determinants of lobbying in each group. A two-pool model emerges. occasional lobbyers' lobbying depends on their market performance and habitual lobbyers' lobbying only depends on past lobbying. The latter tends to be larger steel firms whose business is more focused in steel. Our evidence is consistent with dynamic economies of scale in protection seeking breeding protection-dependent firms.

JEL classification: F13

I. Introduction

Rent-seeking activities plausibly have dynamic economies to scale. Past rent-seeking experience should reduce the cost of further rent seeking and increase its return. Thus rent-seekers may, over time, become more prone to further rent seeking and even become dependent on rent seeking. Given the large welfare losses theoretically associated with rent seeking (e.g. Bhagwati 1982, 1988 and Magee *et al*. 1989), a better understanding of actual rent seeking behavior is required. In particular, empirical verification of the self-sustaining nature of rent seeking is of fundamental importance.

Lobbying for protection from import competition is a form of rent seeking. This paper uses data on lobbying for protection by firms in the steel industry in the 1970s and 1980s to show that a habit-forming effect does exist. In a preliminary investigation, we pool all firm level data and find that past lobbying increases the current tendency to lobby. However, not all firms have the same propensity to lobby. To allow heterogeneity among firms' dependence on past lobbying, we apply an EM (expectation-maximization) algorithm approach (Dempster, Laird, and Rubin, 1977) to a lagged-dummy model (Heckman, 1982a; 1982b). This lets our firms sort themselves into groups according to the determinants of their lobbying activity.

Our statistical results suggest that an acceptable division of our data is a division into two groups: occasional lobbyers, whose lobbying depends on the firm's market performance; and habitual lobbyers, whose lobbying is essentially unrelated to the firm's business situation and depends mainly on past lobbying. Firms that never lobby for protection end up in the first group. Greater firm size and greater focus in steel production are associated with increased lobbying in both groups, but the influence is stronger for occasional lobbyers. Past sales growth and spending on modern equipment are associated with curtailed lobbying by occasional rent-seekers, but are unrelated to lobbying by habitual rent-seekers. Changes in cash flow have no influence in either. Generally, habitual lobbyers are larger firms whose business is more concentrated on steel. They then are naturally more inclined than occasional lobbyers to initiate lobbying. However, their lobbying appears to have become a habit devoiced from their market performance. These results are consistent with the presence of economies of scale in rent-seeking and with rent seeking being habit-forming (e.g. Magee *et al.* (1989).

In the next section, we describe the intensive lobbying for protection in the U.S. steel industry in the 1970s and 1980s. We justify our contentions that lobbying for protection by steel firms in the sample period may be habit-forming, and that different groups of firms should have dissimilar propensities to lobby for protection. In section three, we explain the EM algorithm approach. We explain the data in the fourth section and report our results in section five. Section six concludes.

II. Habitual and Occasional Lobbying in the US Steel Industry

In the U.S., domestic firms under import competition pressure often complain to the government about "unfair" foreign practice. These complaints usually allege either unfair foreign government subsidies or dumping. The U.S. government then investigates the veracity of these claims and decides whether or not material injury has occurred. Sometimes, complainants invoke the "escape clause" (Section 201 of the 1974 Trade Act), which allows temporary protection if imports are causing material injury to a U.S. industry. Foreign firms must actively participate in these investigations to try to prevent biased readings of the data and the subsequent erection of trade barriers. The process is commonly regarded as prejudiced and as a form of administered trade protection (Finger, Hall, and Nelson, 1982, p. 452-466) that coerces foreign firms to curtail their penetration of the US market (e.g. see Hartigan, Perry, and Kamma, 1986 [p. 610-617] and Staiger and Wolak, 1994).

The American steel industry confronted intensifying import competition during the 1970s and early 1980s. Deardorff and Stern (1988) report the U.S. trade deficit in steel almost tripled, from $2 billion to $5.9 billion between 1973 and 1983. Crandall

(1987, p. 275) documents steel imports to the U.S. increasing from an annual average of 16.49 million tons in the 1973-1979 period to an annual average of 20.06 million tons in the 1980-1986 period, a 22% rise. In contrast, domestic steel output dropped 23% from an annual average of 93.83 million tons between 1973 and 1979 to an annual average of 72.16 million tons between 1980 and 1986. This heightened import competition, along with a declining demand, led to combined losses of $9.5 billion in the 1983-1986 period for the seven major U.S. integrated producers studied by DeAngelo and DeAngelo (1991, p. 4).

In response, some steel companies aggressively sought trade protection. After 1979, when authority over trade complaints was transferred from the Treasury Department to the International Trade Administration (I.T.A.) of the Department of Commerce[1], their lobbying intensified and focused on the 1974 Trade Act. In the early 1980's, more than 60% of all petitions for protection submitted to the U.S. government were filed by steel companies. According to Deardorff and Stern (1988, table 2.5), steel companies filed 75% of all “countervailing” complaints and 59% of all “antidumping” complaints in the period from 1980 to 1984.[2]

Faced with this onslaught of lobbying, the U.S. government implemented a series of protectionist policies. “Trigger price mechanisms” were established in 1977 and again

---

[1] The latter was widely perceived as more sympathetic to protectionist arguments. Table 2.5 in Deardorff and Stern (1988) shows that the number of investigations related to trade complaints jumped from one on two per year in the late 1970's to 8 in 1980, 8 in 1981, and 159 in 1982.

[2] Following the filing of an anti-dumping case, the Department of Commerce (or, prior to 1979, the Treasury Department) was to decide within 20 days if the case merited investigation. If the preliminary ruling was affirmative, the International Trade Commission was to decide within 45 days if there was material injury. If the ITC found injury, the Department of Commerce had 110 days to complete its investigation for dumping complaints and 40 days for unfair subsidy complaints. If the Commerce Department found dumping or unfair subsidies, importers of the product were required to post a bond equal to an estimate of the value of the unfair subsidies or dumping margin. The Department of Commerce was to conduct on-site verifications within 75 days. If these verifications showed unfair trade practices, the International Trade Commission was to arrive at an injury determination within 45 days. See Eichengreen and van der Yen, 1984 (p. 72) for further details.

in 1980, and “voluntary” export restraints followed in 1982 and 1984. Trigger prices were floor prices for various steel imports. Imports at prices below these floors were *prima facie* considered to be dumping. The 1977 trigger prices were based on Japanese steel mills' production cost converted to U.S. dollars at historical dollar yen exchange rates. These trigger prices became ineffective barriers as US inflation surged. The 1980 trigger prices were intentionally set low to avoid rankling US allies, and were widely viewed as ineffective from the outset. To derail further US protectionist measures, the EC agreed to “voluntary” export restraints in 1982. Protected by these restraints, non-EC steel firms rapidly penetrated the U.S. market. The Reagan administration promised comprehensive multiple bilateral “voluntary” export restraints in 1984, as U.S. steel firms filed a flurry of petitions for protection against imports from countries like Poland and Argentina. The EC export restraints were strengthened, and similar agreements were quickly reached between the US and all major steel exporters. The effects of these protection measures are examined in Crandall (1987) and Lenway, Morck, and Yeung (1996). Their evidence suggests that trade protection of the U.S. steel industry was the fruit of rent seeking and benefited managers and steel workers with tenure-seniority, but did little to improve the competitiveness of American steel firms.

It is possible that political lobbying may have dynamic economies of scale for several reasons. First, lobbying requires large up-front investments in intangible assets like political connections, professional lobbyers, lawyers, and knowledge of political and legal procedures and channels. Once these investments are made, the marginal cost of further lobbying is comparatively low. Second, firms may learn by doing; practicing lobbying now leads to more effective lobbying in the future. Finally, f lobbying requires high start-up costs, any long gap in lobbying may force firms to pay the high start-up costs again. Hence, firms that have invested heavily in lobbying will lobby continuously.[3]

[3] We are grateful to William Neilson for this suggestion.

Finance theory says firms evaluate investments according to their expected returns. As a firm acquires rent seeking experience, further investments in rent-seeking offer increasingly attractive returns compared to investments in productive assets. In short, lobbying can induce further lobbying, and firms can become "habitual" lobbyers, who essentially supply lobbying inelastically. (See, e.g., Murphy *et al.* 1993)

Protection can generate negative externalities. Lenway, Morck, and Yeung (1996) find that protection seeking US steel firms appeared to politically engineer protection that benefited their stakeholders; but that harmed those more profitable and innovative US steel firms that did not explicitly seek protection.[4] (See also Crandall, 1987 which studies the Steel industry, and for general results, see Magee et al. 1989 and Murphy et al, 1993).

Understandably, therefore, attitudes towards protection may vary across firms in the same industry. While many steel firms actively and repeatedly sought protection, many other firms did so rarely or not at all. Some might be free riding on other firms' lobbying. Some may be in the process of developing a lobbying habit. However, a small minority of steel firms clearly stated their objection to protection during various congressional hearings in the nineteen-eighties. For example, innovative and profitable steel firms like Nucor explicitly lobbied against trade barriers. Much previous research on lobbying uses industry level data, and so misses such intra-industry differences. Firm level studies, like Lenway, Morck, and Yeung (1996) point to such differences being potentially important.

[4] Using a cross-section of firm-level data, Lenway *et al.* (1996) find lobbying firms to be larger, less profitable and less invested in R&D than non-lobbyers. They show that comprehensive quota protection raised the value of only the lobbying firms, not the more competitive non-lobbying firms. Moreover, the impact of protection on non-lobbying firms' values is negatively related to their past spending on innovation. They also show that protection raises compensation to lobbying firms' CEO, and that this is unrelated to changes in profitability. Finally, although protection does not mitigate the loss of jobs in lobbying firms, it does increase the wages of workers in lobbying firms who manage to keep their jobs. For non-lobbying firms, protection significantly affects neither wages nor job losses.

In summary, there appear to be different types of firms with different tendencies to lobby. Occasional lobbyers lobby in response to negative changes in firm performance and to reductions in competitiveness. Habitual lobbyers undertake lobbying actions more inelastically, and their tendency to lobby depends on past lobbying rather than firm performance or characteristics.[5] Without any obvious indicators to differentiate habitual lobbyers from occasional lobbyers, we have a typical mixture model problem in the empirical investigation of the hypothesis[6]. We need to let our sample of US steel firms sort themselves statistically into groups based on the determinants of their lobbying. To do this, we use an EM (expectation-maximization) algorithm (Dempster, Laird, and Rubin, 1977) approach to a lagged-dummy model of the Heckman (1982 a, b) type. The next section provides technical details.

III. An EM Algorithm Approach

In this section, we present our model of lobbying behavior, explain why standard estimation techniques fail in this context, and the provide intuition for our estimation procedure, the expectation-maximization (EM) algorithm. The technical details of the variant of the EM algorithm we use are described in the appendix at the end of this paper.

5 It would be desirable to explore also the determinants of stance against trade protection. Unfortunately, we do not have enough data to do this.

6 One may conjecture that some variables like "firm size" or "sheer" lobbying frequency could serve as indices to separate the pool of steel firms into sub-pools. Applying switching regression techniques, one then can use the likelihood maximizing value of these chosen dimensions to partition our sample and apply a probit regression to each sub-sample. The more advanced mixture model techniques we employ here allow the data to sort themselves into sub-groups assuming that there is a latent variable along which sample partitioning makes sense. Our approach does not rely on our judgment about which dimensions to employ in partitioning our sample. Rather, as we shall show, our statistical results reveal the differences between sub-groups.

The Model of Lobbying Behavior

Our contention is that the determinants of lobbying activity may be different for different sorts of firms. In particular, we postulate that past lobbying experience might be the primary determinant of current lobbying activity for some firms, but that various current firm characteristics might be the primary determinants of lobbying in other firms. We therefore require a model of lobbying activity within which we can nest both possibilities.

To do this, we use a Heckman (1982 a, b) lagged-dummy model. We define the dummy variable

$$y_{it} = \begin{cases} 1 & \text{if firm } i \text{ lobbied in period } t \\ 0 & \text{otherwise} \end{cases}$$

We assume an unobservable underlying "lobbying profit" y_{it}^{*} for firm i in period t. This lobbying profit can depend on a vector of firm i's current characteristics, $\mathbf{x}_{it}$, and on the lagged value of the dummy variable $y_{i,t-}$, which is one if the firm lobbied during the previous period and zero if it did not. Firm i's profit from lobbying in period t is thus

$$y_{it}^{*} = b_k + a_k y_{i,t-1} + \mathbf{x}_{it}^{'} \cdot \mathbf{b}_k + e_{it} \tag{1}$$

where e_{it} is an *iid* normal random variable with mean zero and the unknown parameters b_k, a_k and $\mathbf{b}_k$ can take different values for each of the K different subsamples. The model thus allows firm i's lobbying profits at time t depends on the firm's current characteristics, $\mathbf{x}_{it}$, and on its past lobbying to different extents in different subsamples.

We cannot observe the lobbying profits, y_{it}^{*}, but only whether or not firm i actually lobbied in period t as recorded by the dummy variable y_{it}. We assume that a firm engages in lobbying activity if and only if its lobbying profit exceeds zero.[7] That is, firm i lobbies in period t if and only if $y_{it}^{*} > 0$. It follows that the probability firm i, belonging to subsample k, lobbies in period t is

$$
\begin{aligned}
P(y_{it} = 1) &= P(y_{it}^{*} > 0) \\
&= P(\varepsilon_{it} > -\beta_k - \alpha_k y_{i,t-1} - \mathbf{x}_{it}^{'} \cdot \mathbf{b}_k) \\
&= \Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}^{'} \cdot \mathbf{b}_k)
\end{aligned}
\tag{2}
$$

where P denotes probability, and Φ is the cumulative distribution function of the standard normal distribution. Equation (2) relates observed lobbying (rather than unobserved lobbying profit) to observed past lobbying and firm characteristics for firms in each subsample k.

We assume that our sample of firms can be partitioned into K subsamples, with the parameters $q_k \equiv [b_k, a_k, \mathbf{b}_k]$ different for each subsample $k = 1, \ldots K$. We do not know which firms actually belong to which subsample, so we assume the true partition to be described by the multinomial random variable $\mathbf{z}_i = (z_{i1}, \ldots z_{iK})$, such that

$$
z_{ik} = \begin{cases} 1 & \text{if firm } i \text{ belongs to subsample } k \\ 0 & \text{otherwise} \end{cases}
$$

We assume the elements of $\mathbf{z}_i$ to be *iid* (independently and identically distributed) and to have a multinomial distribution with probabilities $\mathbf{q} = (q_1, \ldots q_K)$.

[7] Actually, all that is necessary if for firm i to lobby in period t if and only if its lobbying profit is greater than some fixed threshold. A nonzero threshold is simply absorbed into the intercept term β_k. See Maddala (1983).

The Estimation Problem

Our statistical procedure is complicated by the fact that we must use the same firm-level data for two purposes:

1). To partition the sample of firms according to the way their lobbying is related to past lobbying and other firm characteristics.

2). To estimate how past lobbying and firm characteristics determine current lobbying within each subsample.

If we knew to which subsample each firm belonged, estimating how past lobbying and firm characteristics determine current lobbying would be a straightforward application of logit or probit regression analysis. Alternatively, if we knew the true relationship linking current lobbying to firm characteristics and past lobbying for each theoretical subsample, dividing the firms into subsamples would also be trivial.

The problem is that we know neither. Moreover, the relationships linking current lobbying to past lobbying and other firm characteristics that we estimate depend on how the sample is partitioned, and the partition we choose depends on the relationships to which we compare our firm-level data.

The Intuition Behind the EM Algorithm

Such estimation problems can be overcome using a recursive algorithm. The specific approach we employ is the expectation-maximization (EM) algorithm described by Dempster, Laird, and Rubin (1977), and illustrated in Figure 1. The intuition underlying this approach is easily described by going through the steps of the algorithm.

In an "initialization step", guesses of parameter values $q \equiv [q_1, \ldots q_K]$ and the multinomial probabilities $\mathbf{q}$ are chosen. The initial values may be chosen in any way, and the robustness of the technique can be tested by seeing whether different initial guesses lead to the same final estimates or not.

We then apply the “expectation step” (E). We use these parameter values to estimate a matrix of estimated probabilities that firm i‘s data was generated by the parameters $q_k \equiv [b_k, a_k, \mathbf{b}_k]$. We denote these estimated probabilities $p_{ik} = P(z_{ik} = 1)$, and use them to give a weight to each firm in each subsample. This gives us an estimated partition of our firms into subsamples.

We then apply the “maximization step” (M). Given the estimated partition from the previous E step, we calculate maximum likelihood estimates of the parameters q and **q**. Given these updated parameter estimates, we can execute the E step again and derive an updated partition.

A recursive procedure is used. The E step is repeated using the new parameter values from the M step, and reassigns weights to firms across the subsamples. Given the revised estimated partition from the E step, the M step is then repeated to generate new parameter estimates. The expectation and maximization steps are repeated until a convergence criterion is satisfied; for example, when the absolute distance between estimates from consecutive iteration is less than a certain value. If the algorithm converges to an optimal partition estimate and set of parameter estimates, we can then assess the economic plausibility of these results.

The whole procedure can be repeated for $K = 1, 2, 3 \ldots$ and the results for each can be compared statistically using Akaike's (1974) information criterion[8]

$$\text{AIC}(K) = -2\, ln\, L + 2\, N(K), \tag{3}$$

[8] The AIC is a comparative measure between models with different numbers of constructs. It relates the goodness-of-fit of the model to the number of coefficients required to achieve a level of fit. AIC is similar to the *adjusted* R^2 in multiple regression. A small AIC occurs when large likelihood values are achieved with fewer estimated coefficients. The basic objective is to diagnose whether model fit has been achieved by "overfitting" the data with too many coefficients. Since there is no statistical test available, AIC could be used in most instances to comparisons between models (See Akaike, 1987).

where $N(K)$ is the number of free parameters in the model and L is the likelihood function of the sample . The lower that value of AIC(K), the better the "fit" of the partition into K subsamples.

IV. Data and Variables

Our sample of steel firms consists of all companies listed in the *Standard and Poor's Corporate Register* between 1977 to 1988 under S.I.C. codes 3312 (steel works), 3315 (blast furnaces), 3316 (rolling mills), and 3317 (finishing mills).[9] These are an exhaustive list of S.I.C. codes for steel production. Firms not included on the *Compustat* tapes are dropped.[10] The resulting sample is a panel of 890 firm-year observations spanning 121 firms. Our sample includes a fairly complete cross section of the steel industry. It contains all the integrated steel companies and 14 of the 42 mini-mills in Barnett and Crandall (1985). The mini-mills we omitted are relatively small, with capacities less than 400,000 tons in 1985. Because of exit from the industry, the panel is not balanced across years.

We collect publicly available information on firms' lobbying for protection. This includes petitions for escape clause protection, petitions for the imposition of a countervailing duty or antidumping measures, and complaints about foreign government practices. We also include testimony in support of trade protection at congressional hearings. The names of firms undertaking the above activities were compiled from the *Federal Register* and the *CIS Congressional Abstract Index*.[11] Protection-seeking activity by a subsidiary was considered protection-seeking by the parent firm. The parent

[9] SIC code 3313 is 'electro-metallurgical products except steel'; SIC code 3314 is not assigned.

[10] These firms include those that do not file 10-K forms and firms that have gone out of business. Many very small firms are not included on the *Compustat* tapes.

[11] We scanned the database using the keywords 'steel', 'steel trade', and 'trade'. We examined each retrieved piece and retained those that fit our definition of protection seeking.

companies of subsidiaries were found by searching the *Standard and* Poor's *Corporate Register, Moody's Industrial Manuals, Capital Adjustments,* the *Value Line Investment Survey* and the *Directory of Corporate Affiliates.* We used this information to construct a dummy variable y_{it} equal to one if firm i lobbied for protection year t.

Our premise is that there may be habitual lobbyers and occasional lobbyers. Habitual lobbyers are influenced by the dynamic economies of scale in lobbying and thus are prone to petition for more trade protection. Occasional lobbyers have not reached such a point. They may become habitual lobbyers in the future, or they may choose to free ride on other firms' lobbying effort, or they may even choose to invest in other industries. The point is that the propensities to lobby for protection of firms in different groups are affected differently by the same firm characteristics.

Which firm characteristics are likely to be important? First, there are likely to be static economies of scale in lobbying. Big firms have more financial resources and so should be able to absorb the fixed up-front costs of lobbying better than small firms. Big firms also plausibly have lower marginal costs of lobbying because they have in-house legal staff. At the same time, steel firms with a greater volume of business should benefit more from protection. We therefore need a measure of firm size, namely:

Total Assets (*Size*) is used as a measure of firm size. This variable is adjusted for inflation and is in millions of 1983 dollars.

Firms that are more specialized in steel should benefit more from protection than would more diversified firms. We thus need a measure of steel focus. For this, we use the dummy variable:

Concentration in steel production dummy (*Steel***)** is equal to one if a firm's primary line of business, as listed in the *Standard and* Poor's manual for that year is 3312, 3315, 3316, or 3317 - the four SIC codes for steel production. It is also set to one if all four steel SIC codes are included in the list of the firm's lines of business. Otherwise, the dummy is set to zero[12]. The dummy is used to capture a firm's concentration in steel production.

Innovative firms are more likely to have investment opportunities with returns higher than lobbying returns. We therefore need a measure of investment in innovation, namely:

Research and Development Spending (*RD/A*) is measured per dollar of total assets. If research and development spending is not reported, but all other financial data is available, we assume R&D spending to be nil. R&D spending is scaled by total assets to capture the intensity in investment in innovation.

Finally, firms should invest more in lobbying when the returns on their other investments are low and when their sales are declining. This situation also makes satisfying the "material injury" requirement in dumping and countervailing duty cases easier, so the likelihood of successful lobbying increases with poorer market performance. Such firms may also have invested less in physical assets in the past. To capture these characteristics, we use the following set of variables:

Accumulated Depreciation (*Depreciation*), taken from the firm's balance sheet and divided by the book value of its net plant and equipment, is used as a measure of the lack of past investment in productivity (i.e. accumulated depreciation / gross plant and equipment).

Sales Growth is defined as the firm's most recent sales figure minus its sales the previous year, all divided by the latter. This variable is constructed using deflated sales figures in 1983 dollars to correct for inflation.

---

[12] Another alternative is to use industry segment data to construct an index indicating a firm's involvement in steel industries. However, the alternative is not very attractive because segmented data are less reliable and adequate segmented data are unavailable for many firms in our sample.

Change in Returns on Assets (*CROA*) is the first difference of income before extraordinary items gross of depreciation and interest expenses per dollar of total assets. This variable measures the change in cash flow produced per dollar of corporate assets.

We postulate that, on a year-by-year basis, these firm characteristics and market performance variables should affect occasional lobbyers' behavior, but not that of habitual lobbyers. That is, habitual lobbyers have a more inelastic lobbying tendency than occasional lobbying firms.

Habitual lobbying firms have already invested and set up their lobbying apparatus. Dynamic economies of scale in lobbying imply that the marginal cost of lobbying for these firms should decrease over time, while their return on lobbying rises. Lobbying firms should eventually become inelastic participants in rent seeking. To capture the possibility that lobbying is habit forming, that past lobbying increases the current tendency to lobby, we include the lagged value of a our lobbying dummy as a final right hand side variable.

Past lobbying ($lobby_{-1}$) is the lagged value of a dummy indicating a firm's involvement in lobbying.

Our main focus is whether "$lobby_{-1}$" increases the tendency to lobby. If lobbying has dynamic economies of scale, past lobbying should increase the likelihood of current lobbying. In addition, we expect that the propensity to lobby increases with "*size*," "*steel*," and "*depreciation*," but decreases with "*R&D/A*," "*CROA*," and "*sales growth*." If lobbying is indeed habit forming, but not all firms are addicted to lobbying, we also expect the EM algorithm to produce more than one functional relationship between the set of independent variables and the probability to lobby. In particular, we expect that one functional form should describe habitual lobbying while other functional forms should describe occasional lobbying. For habitual lobbyers, past lobbying should be much more important to the extent that it may be the only significant determinant of current lobbying for that group. All the other variables should matter more for occasional lobbyers, while past lobbying ought to be much less important.

V. Results

Table 1 reports the correlation matrix of our data. The lobbying dummy is positively and significantly correlated with size, steel production focus, depreciation (a proxy for older physical assets), and with the dummy indicating past lobbying. Lobbying is negatively and significantly correlated with sales growth, but insignificantly negatively correlated with changes in the returns on assets. Contrary to our expectations, the lobbying dummy is positively correlated with R&D spending,[13] but the correlation is insignificant.

[Table 1 about here]

Overall, a quick scan of the data suggests that lobbyers are larger, more concentrated in steel production, and have invested less in modernizing their plant and equipment. Also, lobbyers suffer from declining sales and returns on assets. Past period lobbying also seems to increase the probability of current period lobbying.

We run the algorithm described in section III with all the independent variables included assuming 1, 2 and 3 pools in the data (i.e. K = 1, 2 and 3). The resulting estimates, standard deviation (STD) and p-value are reported in table 2.

[Table 2 about here]

[13] Many non-lobbying firms are smaller firms not reporting R&D spending. They also tend to exit the industry after 1984 (Lenway, Morck and Yeung (1996)). We obtain the correlation matrix based on firm-year entries. Larger lobbying firms which lasted for the whole sample period are than given a greater weight.

When there is only one pool (i.e. $K = 1$), the model is equivalent to an ordinary probit on pooled firm level panel data. In the one pool model, lobbyers tend to be larger firms that are more concentrated in steel production and that have declining market performance and low levels of investment in productivity improvements. The focal result, however, is that past lobbying significantly increases the likelihood of current lobbying, consistent with lobbying being habit-forming.

When we allow the data to form two pools (i.e. $K = 2$), our firms appear to cleanly sort themselves into habitual and occasional lobbyers. In the first pool, past lobbying positively and highly significantly affects the probability of current lobbying. *Size* and *Steel* also positively and significantly affect the probability of current lobbying. However, all other variables are insignificant. These are the characteristics of habitual lobbying: a firm finds the dynamic economies of scale in lobbying and becomes an inelastic supplier of lobbying.

In the second pool, lagged lobbying does not affect the probability of current lobbying. The coefficient for lagged lobbying is negative and insignificant. *Size*, *Steel* and *Depreciation* are positive and significant while *Sales Growth* is negative and significant, indicating that larger steel firms with older equipment become more inclined to seek protection when they experience poorer sales growth. Both *CROA* and *RD/A* are insignificant.

We hypothesize that the first pool contains habitual lobbyers while the second contains occasional lobbyers. All estimated coefficients (except that on past lobbying) for the pool of occasional lobbyers are greater in magnitude than those for the pool of habitual lobbyers, which is consistent with habitual lobbyers having a more inelastic tendency to lobby. Unfortunately, we are aware of no statistically rigorous way to test the differences between these sets of coefficient estimates.

The three-pool model distinguishes two groups, the first and third, with lobbying determinants similar to those of occasional and habitual lobbyers in the two-pool model.

A further group, the second pool contains firms whose lobbying depends on both firm characteristics and past lobbying. A plausible conjecture might be that these firms are in the process of becoming habitual lobbyers.

Unfortunately, the likelihood surface in the neighborhood of q = (1, 0, 0, ..) under the null hypothesis is discontinuous, so ordinary likelihood ratio tests cannot determine the most likely number of pools (K). However, we can use the Akaike's (1974) information criterion AIC(K)= $-2\ ln\ L + 2\ N(K)$, where $N(K)$ is the number of free parameters in the model, to roughly compare the "goodness of fit" of each alternative. The statistically desirable K has a low AIC.

Table 2 indicates that models with three pools or more (results for $K > 3$ are not shown) fit less well than the one pool and two pools models. However, the one and two pools models are virtually indistinguishable from each other.

The viability of the two pool hypothesis therefore must depend on whether the partition it implies is consistent with a plausible economic explanation of why habitual lobbyers should act differently than occasional lobbyers. To verify the hypothesis that the first pool contains habitual lobbyers while the second contains occasional lobbyers, we need to assign firms to either the first or second pool.

The probit estimates shown in Table 2 are based on estimates of the probabilities that each firm i belongs to each subsample k. As was described in section 3, we re-estimate the probability that firm i belongs in subsample k, $p_{ik} = P(z_{ik} = 1)$, each time we execute the E step of the EM algorithm. The values of p_{ik} used in the final M step before convergence, the step that generated the parameter estimates shown in Table 2, can be used to assign each firm to the subsample most likely to have generated its observed data. That is, we assign firm i to subsample k if and only if $p_{ik} > p_{ij}$ for all j where $1 \leq j \leq K$ and $j \neq k$

The comparisons between the two pools are tabulated in Table 3. Pool one contains 57% of our firms leaving 43% in pool two. Firms that never lobbied all end up in pool two, the occasional lobbyers' pool. The habitual lobbyers' mean and median lobbying frequency per firm are 34% and 25%, respectively, while that the occasional lobbyers' mean and median are 3% and 0%. Both the means and the medians are statistically significantly different. Habitual lobbyers are more "steel focused" than occasional lobbyers. The mean and median steel focus dummy for the former are 0.62 and 1, respectively, compared with 37% and 0 for the latter. While the difference in means is not statistically significant, the difference in medians is. The mean of the habitual lobbyers' size is 1.16 of the mean of the occasional lobbyers' size, but this difference is statistically insignificant. The habitual lobbyers' median size is 160% of the occasional lobbyers' median size, and the difference is statistically significant. The comparisons in terms of size and steel focus suggest that the occasional lobbyers' pool is composed of both smaller steel-focused firms and large firms for which steel is only one of several lines of business. In contrast, habitual lobbyers are typically large firms focused in steel.

Interestingly, the differences in the other firm characteristics between the two groups are utterly insignificant. The two pools of firms have similar depreciation, R&D spending, sales growth, and change in returns on assets. The lack of material differences in these firm characteristics highlights the contrast between the two groups' lobbying tendency: equally poor market performance prompts occasional lobbyers' to lobby but has no relation to habitual lobbyers' tendency to lobby, which depends only on past lobbying.

These findings are economically sensible, and so lend credence to the two-pool model. Habitual lobbyers, being larger and more focused in the steel industry, have good reasons to initiate lobbying. Once they have started, their lobbying becomes a habit, in the sense that past lobbying leads them to lobby again regardless of their market performance. For other firms, an intuitively sensible relationship between lobbying and changes in firm performance and past investment in productivity holds, and past lobbying does not predict present lobbying.

One plausible interpretation of our result is that to successfully win protection, a firm must lobby continually for several years. However, our data indicate that many firms continued lobbying well beyond 1984, the year comprehensive multiple bilateral voluntary export restraints agreements were set up. They kept on lobbying despite the existence of comprehensive protection that cleanly arrested the decline in their market performance.[14] This behavior observation is consistent with our premise that lobbying has dynamic economies of scale. Once a firm has invested heavily in lobbying, it is more prone to use lobbying intensively and continuously. [15]

VI. Conclusions

This paper examines political lobbying for trade protection by American steel firms. By pooling all firms' data, we show that, on average, past lobbying increases the likelihood of current lobbying. When we let our data sort themselves into groups according to the determinants of their lobbying, two groups emerge: occasional and habitual lobbyers. Occasional lobbyers' lobbying is closely tied to firm performance and strategy variables. Non-lobbyers group themselves with occasional lobbyers. Habitual lobbyers are larger and less diversified steel firms, and account for the lion's share of lobbying. Thus, one would expect that they are more likely to initiate lobbying. Their

[14] See Lenway *et al.* (1996).

[15] An analogous sorting procedure puts 46%, 31%, and 23% of the firms into the first, second and third pools respectively. The first and third pools have lobbying characteristics analogous to the occasional and habitual lobbyers on the two-pool model. The second pool firms' lobbying has intermediate features. If it contains firms in the process of becoming habitual lobbyers, their lobbying should intensify over time. This is observed. The fraction of firm-years in which lobbying occurs rises in pool two from 0.12 prior to 1983 to 0.21 after 1982. For pools one and three, lobbying intensity falls from 0.04 to 0.01 and from 0.17 to 0.13 respectively. Pool two firms show rising ROA and falling R&D, consistent with their becoming "hooked" rent-seekers. However, these grouping are highly uncertain because the AIC of the three-pool model indicates relatively very poor statistical fit.

lobbying becomes habitual, in the sense that current lobbying is explained by past lobbying but is relatively unrelated to their performance. Overall, our findings support the view that, because of the dynamic economies of scale in rent-seeking, lobbying for protection can become habitual.

References

Amemiya, T (1985). *Advanced Econometrics*, Harvard University Press, Cambridge, MA.

Akaike, H. (1987) "Factor Analysis and AIC." *Psychometrika*, 52, 317-332.

Akaike, "A New Look at Statistical Model Identification," *IEEE Transaction on Automatic Control*, 6, 1974, 716-723.

Barnett, Donald F. and Robert Crandall, *Up from the Ashes: the Rise of the Steel Minimill in the United States,* the Brookings Institution, Washington, DC, 1986.

Baumol, William J., "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy,* 98, 1990, 893-921.

Becker, Gary S., Michael Grossman, and Kevin M. Murphy, 1994, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, Vol. 84 No. 3, June, pp. 396-418.

Bhagwati, Jagdish N., "Directly Unproductive, Profit-seeking (DUP) Activities," *Journal of Political Economy,* Vol. 90 No. 5, October 1982, 1982, 988-1002.

Bhagwati, Jagdish, *Protectionism,* Cambridge, Mass: MIT Press, 1988.

Crandall, Robert W., "The Effects of U.S. Trade Protection for Autos and Steel," *Brookings Papers on Economic Activity*, 1. 1987, 272-288.

DeAngelo, Harry and Linda DeAngelo, "Union Negotiations and Corporate Policy: A Study of Labor Concessions in the Domestic Steel Industry during the 1980s," Journal of Financial Economics, Vol. 30, November, 1991, 2-43.

Deardorff, Alan, and Robert Stern, "Current Issues in Trade Policy," in *U.S. Trade Policies in a Changing World Economy*, Robert Stern (ed.), MIT Press, Cambridge: Mass. 1988, 15-68.

Dempster, A. P., N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via EM Algorithm," *Journal of Royal Statistical Society* B 39, 1977, 1-38.

Eichengreen, Batty, and Hans van der Ven, "U.S. Antidumping Policies: the Case of Steel," in *The Structure and Evolution of Recent U.S. Trade Policy,* Robert E. Baldwin and Anne O. Krueger (eds), N.B.E.R. conference Report, University of Chicago Press, 1984.

Finger. J. M., H. Keith Hall, and Douglas R. Nelson, "The Political Economy of Administered Protection," *American Economic Review*, Vol. 72 3, June, 1982, 452-466.

Hartigan, James C., Philip R. Perry, and Sreenivas Kamma, "The Value of Administered Protection: a Capital Market Approach," *Review of Economics and Statistics*, Vol. 68, 1986, 610-617.

Heckman, J., "Statistical Models for the Analysis of Discrete Panel Data," in C. Manski and D. McFadden (eds*), Structural Analysis of Discrete Data: With Econometric Applications*. Cambridge, Mass: MIT Press, 1982a.

Heckman, J., "The Incidental Parameter Problem and the Problem of Initial Conditions in Estimating a Discrete Stochastic Process and Some Monte Carlo Evidence on Their Practical Importance," in C. Manski and D. McFadden (eds*), Structural Analysis of Discrete Data: With Econometric Applications*. Cambridge, Mass: MIT Press, 1982b.

Krueger, Ann O. "The Political Economy of the Rent-Seeking Society," *American Economic Review*; Vol. 64 no 3 June 1974, 291-303.

Lenway, Stefanie, Randall Morck, and Bernard Yeung, "Rent-Seeking, Protectionism and Innovation in the American steel industry," *Economic Journal*, Vol. 106 No. 435, March, 1996, 410-421.

Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics. Cambridge University Press, New York.

Magee, Stephen P., William A. Brock, and Leslie Young. *Black Hole Tariffs and Endogenous Policy Theory Political Economy in General Equilibrium,* Cambridge University Press, 1989.

McLanchlan, G. J. and K. E. Basford, *Mixture Models: Inference and Application to Clustering*, New York: Marcel Dekker, 1988.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "Why is Rent-Seeking so Costly to Growth," *American Economic Review*, Vol. 83 2, May, 1993, 409-414.

Render, R. and H. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Reviews*, 26, 1984, 195-240.

Staiger, Robert W. and Frank A. Wolak. "Measuring Industry-Specific Protection: Anti-dumping in the United States," Brookings Papers on Economic Activity: Microeconomics. 1994, 51-103. "

Wedel, M and W. S. Desarbo, " A review of recent developments in latent class regression models," in *Advanced Methods of Marketing Research*, Ed,. R. Bagazzu, 1994, 352-388.

Wedel, M and W. S. Desarbo, "A mixture likelihood for generalized linear models," *Journal of Classification*, 12, 1995, 21-55.

Table 1: Correlation matrix

	Lobby	Size	Sales Growth	Change ROA	R&D/A	Steel	Depreciation
Size	0.246 (0.000)						
Sales Growth	-0.085 (0.017)	0.009 (0.805)					
Change ROA	-0.002 (0.946)	-0.033 (0.355)	0.198 (0.000)				
R&D/A	0.024 (0.498)	0.489 (0.000)	-0.024 (0.509)	-0.050 (0.164)			
Steel	0.252 (0.000)	-0.112 (0.002)	-0.007 (0.835)	0.032 (0.365)	-0.205 (0.000)		
Depreciation	0.388 (0.000)	0.604 (0.000)	-0.034 (0.349)	0.010 (0.775)	0.379 (0.000)	0.125 (0.000)	
$Lobby_{-1}$	0.005 (0.000)	-0.024 (0.000)	0.015 (0.035)	0.023 (0.078)	-0.001 (0.466)	0.008 (0.000)	-0.045 (0.000)

p-value in parentheses

Table 2: An E-M Algorithm pooling approach applied to lagged-dummy model: $y^*_{it} = a y_{i,t-1} + x_{it} b + e_{it}$. The lobbying dummy $y_{it} = 1$ if $y^*_{it} > 0$, $y_{it} = 0$ otherwise. All independent variables included.

	Const.	Size	Steel	Sales Growth	Change ROA	R&D /A	Depre-ciation	$Lobby_{-1}$
K=1								
AIC = 386.1684								
Estimate =	**-2.040**	**0.394**	**0.752**	**-0.152**	0.057	**-0.237**	**0.160**	**0.914**
STD =	0.140	0.113	0.171	0.071	0.069	0.139	0.081	0.185
p-value =	**0.000**	**0.001**	**0.000**	**0.031**	0.415	**0.087**	**0.048**	**0.000**
K=2								
AIC = 389.0822								
Pool 1: Habitual Lobbyers								
Proportion =	0.569							
Estimate =	**-1.876**	**0.514**	**0.893**	-0.128	-0.005	-0.284	0.073	**0.911**
STD =	0.164	0.155	0.212	0.080	0.086	0.175	0.093	0.210
p-value =	**0.000**	**0.001**	**0.000**	0.109	0.957	0.104	0.430	**0.000**
Pool 2: Occasional Lobbyers								
Proportion =	0.431							
Estimate =	**-4.216**	**0.721**	**1.869**	**-0.545**	0.138	-0.394	**0.775**	-0.535
STD =	0.898	0.327	0.819	0.230	0.142	0.342	0.238	0.501
p-value =	**0.000**	**0.028**	**0.023**	**0.018**	0.331	0.249	**0.001**	0.285
K=3								
AIC = 403.1564								
Pool 1: Inexperience Lobbyers								
Proportion =	0.460							
Estimate =	**-2.438**	**0.492**	**1.009**	-0.173	0.068	**-0.457**	**0.454**	-0.316
STD =	0.288	0.191	0.314	0.128	0.111	0.248	0.161	0.373
p-value =	**0.000**	**0.010**	**0.001**	0.175	0.539	**0.065**	**0.005**	0.398
Pool 2: Experience Lobbyers								
Proportion =	0.312							
Estimate =	**-2.191**	**1.055**	**1.486**	**-0.254**	-0.087	-0.293	-0.101	**0.674**
STD =	0.275	0.264	0.342	0.112	0.113	0.215	0.129	0.279
p-value =	**0.000**	**0.000**	**0.000**	**0.023**	0.443	0.172	0.432	**0.016**
Pool 3: Habitual Lobbyers								
Proportion =	0.229							
Estimate =	**-2.170**	-0.262	0.076	-0.089	0.162	-0.083	0.240	**2.754**
STD =	0.342	0.267	0.465	0.207	0.210	0.476	0.201	0.548
p-value =	**0.000**	0.327	0.870	0.666	0.441	0.862	0.233	**0.000**

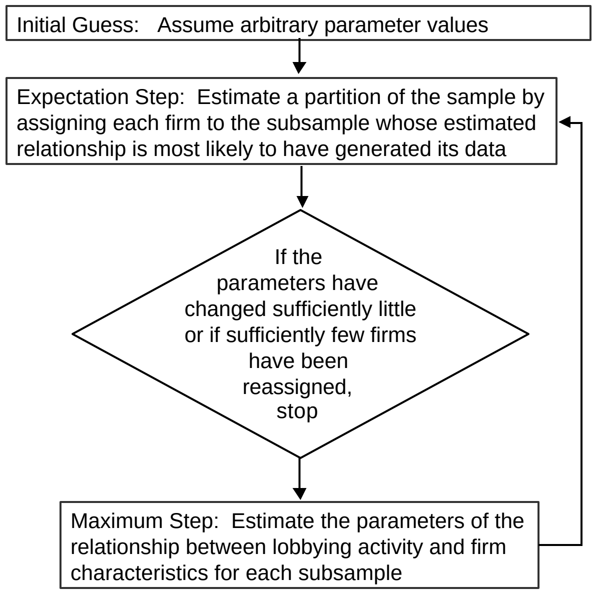
AIC = -2loglikelihood + 2(number of free parameters)

Table 3: The comparisons between the habitual lobbyers' and the occasional lobbyers' firm characteristics. A firm i is assigned to pool one if $LLF_{1,i} > LLF_{2,i}$ and to pool two otherwise. LLF is as defined in Eq. 1 and the parameters in LLF_1 and LLF_2 are taken, respectively, from the first and second sub-block in the K = 2 panel in Table 2.

Variable	**Habitual lobbyers' Group**		**Occasional lobbyers' Group**		**Prob-value: Δs in means**	**Prob-value: Δs in median**
	mean	**median**	**mean**	**median**		
Lobbying frequency	0.340	0.25	0.032	0	**0.016**	**0.000**
Steel focus dummy	0.619	1.00	0.372	0	0.604	**0.024**
Total assets	890.4	431.9	765.9	269.6	0.933	**0.061**
Depreciation in plant and equipment / Total Asset	0.338	0.2808	0.321	0.2683	0.940	0.974
R&D / total assets	0.0058	0	0.0084	0.001	0.868	0.689
Sales growth (log Δs)	0.016	0.004	0.026	0.021	0.877	0.217
Change in ROA	-0.001	0	-0.0001	-.001	0.977	0.833

We use the Wilcoxon statistics to test for the Δs in the median. Other tests lead to similar results.

Figure 1. The Intuition Behind the Expectation-maximization (EM) Algorithm



Appendix

In this appendix, we present the technical details of the expectation-maximization EM algorithm of Dempster, Laird and Rubin (1997) as employed in this paper. Let i = 1, … I, denote firms, t = 0, ... T_i, denote time periods, and k = 1, … K denote groups of firms. Let x_{it} be the column vector of firm characteristics for the i^{th} firm in the t^{th} period. Let the dummy variable y_{it} be one if firm i engaged in lobbying activity at time t.

The unobservable profit of firm i from lobbying in time t, y_{it}^*, is assumed to be defined by Heckman (1982 a, b) lagged-dummy model of the form

$$y_{it}^* = \beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k + \varepsilon_{it} \tag{A1}$$

where b_k, a_k and $\mathbf{b}_k$ are unknown parameters, which can be different for different subsamples of firms, and the errors e_{it} are *iid* with a standard normal distribution. The model assumes that firm i lobbies in period t if and only if its profit from doing so is above a threshold profit level, $\tilde{y}_{it}^*$. It follows that the probability firm i will lobby in period t fs

$$\begin{aligned} P(y_{it} = 1) &= P(y_{it}^* > 0) \\ &= P(\varepsilon_{it} > -\beta_k - \alpha_k y_{i,t-1} - \mathbf{x}_{it}' \cdot \mathbf{b}_k) \\ &= \Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k) \end{aligned} \tag{A2}$$

for firm i a member of subsample k. P denotes probability, and Φ is the cumulative distribution function of the standard normal distribution. Note that we can assume $\tilde{y}_{it}^*$ = 0 without loss of generality for any nonzero threshold will be absorbed into the intercept term b_k in (A2).

Define the vector of parameters $q_k = (b_k, a_k, \mathbf{b}_k)$ and recall the definitions of the vector $\mathbf{y}_{it} = (y_{i1}, \ldots y_{it})'$, and matrix $\mathbf{x}_{it} = (\mathbf{x}_{i1}, \ldots \mathbf{x}_{it})'$ for $t = 1, \ldots T_i$. The likelihood function contribution by the i^{th} firm, a member of subsample k, is then

$$
\begin{aligned}
f_i &= f\left(\mathbf{y}_{iT_i} \middle| \mathbf{x}_{iT_i}, y_{i0}, x_{i0}, q_k\right) \\
&= P\left(y_{i1} \middle| y_{i0}, x_{i0}, q_k\right) P\left(y_{i2} \middle| \mathbf{y}_{i1}, \mathbf{x}_{i1}, y_{i0}, x_{i0}, q_k\right) \cdot \ldots \cdot P\left(y_{iT_i} \middle| \mathbf{y}_{i,T_i-1}, \mathbf{x}_{i,T_i-1}, y_{i0}, x_{i0}, q_k\right) \qquad \text{(A3)} \\
&= \prod_{t=1}^{T_i} \left\{ \Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k)^{y_{it}} \left[1 - \Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k)\right]^{1-y_{it}} \right\}
\end{aligned}
$$

Note that the product is taken from period 1 to period T_i. Period 0 is excluded because firm i's lobbying decision in time t depends on its lobbying activities in the previous period, as measured by $y_{i,t-1}$, as well as on the firm's characteristics in period t, $\mathbf{x}_{it}$.

The log likelihood function for the subsample is thus

$$
L_k = \sum_{i=1}^{I} \sum_{t=1}^{T} \left\{ \begin{array}{c} y_{it} \ln\left[\Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k)\right] + \\ (1 - y_{it}) \ln\left[1 - \Phi(\beta_k + \alpha_k y_{i,t-1} + \mathbf{x}_{it}' \cdot \mathbf{b}_k)\right] \end{array} \right\} \qquad \text{(A4)}
$$

If we knew which firms belonged in which subsample, we could obtain maximum likelihood estimates of the parameters $q_k = (b_k, a_k, \mathbf{b}_k)$ in (A4) by running a probit regression procedure over each of the K subsamples.

We do not know which firms actually belong to which subsample, so we assume the actual partition to be defined by the multinomial random variable $\mathbf{z_i} = (z_{i1}, \ldots z_{iK})$ with

$$
\mathbf{z}_{ik} = \begin{cases} 1 & \text{if firm } i \text{ belongs to subsample } k \\ 0 & \text{otherwise} \end{cases}
$$

We assume the elements of $\mathbf{z}_i$ to be *iid* and to have a multinomial distribution with parameters K and $\mathbf{q} = (q_1, \ldots q_K)$. If we knew the true values of $\mathbf{q}$ and q º $(\mathrm{q}_1, \ldots \mathrm{q}_K)$, we could assign each firm to the subsample whose parameters are most likely to have generated its observed data, and so assign $\mathbf{z}_{ik}$ to zero or one for all firms.

Because we know the true values of neither the $\mathbf{z}_i$ nor ($\mathbf{q}$, q), we must estimate both. As was pointed out in the text, the difficulty is that we must assume values for the $\mathbf{z}_i$ to estimate ($\mathbf{q}$, q), and must assume values of ($\mathbf{q}$, q) to estimate the $\mathbf{z}_i$. The EM algorithm allows us to estimate the $\mathbf{z}_i$ and ($\mathbf{q}$, q) recursively. The expectation (E) step lets us estimate $P(\mathbf{z}_{ik} = 1)\ \forall\ i, k$ given known values of ($\mathbf{q}$, q). The maximization (M) step lets us estimate ($\mathbf{q}$, q) given known values of the $\mathbf{z}_i$. We begin by using the expectation (E) step to estimate the $\mathbf{z}_i$ assuming arbitrary initial guesses as to the value of ($\mathbf{q}$, q). These estimates of the $\mathbf{z}_i$ let us use the maximization (M) step to produce updated values of ($\mathbf{q}$, q). The idea is to continue updating $\mathbf{z}_i$ and ($\mathbf{q}$, q) until some convergence criterion is met.

The E step works as follows. Assume we know ($\mathbf{q}$, q). The joint likelihood function of the sample is

$$L\left(\mathbf{y}_{iT_i}, \mathbf{z}_i \middle| \mathbf{x}_{iT_i}, y_{i0}, x_{i0}, \mathrm{q}\right) = \prod_{k=1}^{K} \prod_{i=1}^{I} q_k f\left(\mathbf{y}_{i,t-1} \middle| \mathbf{x}_{i,t-1}, y_{i0}, x_{i0}, \mathrm{q}_k\right)^{z_{ik}} \qquad \text{(A5)}$$

for $f(\mathbf{y}_{iT} \mid \mathbf{x}_{iT}, y_{i0}, x_{i0}, \mathrm{q}_k)$ defined as in (A3). The joint log likelihood function of the sample is therefore

$$\ln L = \sum_{k=1}^{K} \sum_{i=1}^{I} z_{ik} \ln\left[f\left(\mathbf{y}_{i,T_i} \middle| \mathbf{x}_{i,T_i}, y_{i0}, x_{i0}, \mathrm{q}_k\right)\right] + \sum_{k=1}^{K} \sum_{i=1}^{I} z_{ik} \ln\left[q_k\right] \qquad \text{(A6)}$$

By Bayes' rule, the estimated probability that firm i belongs in subsample k is

$$p_{ik} = E\left[z_{ik} \middle| \mathbf{y}_{i,T_i}, \mathbf{x}_{i,T_i}\right] = \frac{q_k f\left(\mathbf{y}_{iT_i} \middle| \mathbf{x}_{iT_i}, y_{i0}, x_{i0}, \mathrm{q}_k\right)}{\sum_{k=1}^{K} q_k f\left(\mathbf{y}_{iT_i} \middle| \mathbf{x}_{iT_i}, y_{i0}, x_{i0}, \mathrm{q}_k\right)} \qquad \text{(A7)}$$

U sing (A7), we can assign weights to firm i in each subsample k. We can also assign firm i to the subsample k whose lobbying model is most likely to have given rise to that firm's observed data. The former procedure produces Table 2; the latter generates Table 3.

The M step works as follows. Assume we know the $\mathbf{z}_i$. Substituting p_{ik} in (A7) into the place of z_{ik} in (A6) yields

$$Q = \sum_{k=1}^{K}\sum_{i=1}^{I} p_{ik} \ln\left[f\left(\mathbf{y}_{i,T_i} \middle| \mathbf{x}_{i,T_i}, y_{i0}, x_{i0}, \mathrm{q}_k\right)\right] + \sum_{k=1}^{K}\sum_{i=1}^{I} p_{ik} \ln\left[q_k\right] \qquad \text{(A8)}$$

The M step chooses ($\mathbf{q}$, q) to maximizes Q subject to the constraint that $\sum_{k=1}^{K} q_k = 1$. We thus maximize

$$H = Q - l\ (\sum_{k=1}^{K} q_k\ \ -1), \qquad \text{(A9)}$$

where l is a Langrange multiplier. We want to maximize H with respect to the parameters q = (q$_1$, … q$_K$) and $\mathbf{q}$ = (q_1, … q_K).

Let $\Phi(q_k) = \Phi\left(b_k + a_k y_{i,t-1} + x_{it}' b_k\right)$ and $f(q_k) = f\left(b_{0k} + a_k y_{i,t-1} + x_{it}' b_k\right)$ for k = 1, … K where $f(\ .\ )$ is the density function of the standard normal distribution function $\Phi(\ .\ )$. Using this notation, the first order conditions are

$$\begin{aligned}
\frac{\partial H}{\partial b_{0k}} &= \sum_{i=1}^{I}\sum_{t=1}^{T_i} p_{ik} \frac{y_{it} - \Phi(q_k)}{\Phi(q_k)[1-\Phi(q_k)]} f(q_k) = 0 \\
\frac{\partial H}{\partial a_k} &= \sum_{i=1}^{I}\sum_{t=1}^{T_i} p_{ik} \frac{y_{it} - \Phi(q_k)}{\Phi(q_k)[1-\Phi(q_k)]} f(q_k) y_{i,t-1} = 0 \\
\frac{\partial H}{\partial b_k} &= \sum_{i=1}^{I}\sum_{t=1}^{T_i} p_{ik} \frac{y_{it} - \Phi(q_k)}{\Phi(q_k)[1-\Phi(q_k)]} f(q_k) x_{it} = 0 \\
\frac{\partial H}{\partial q_k} &= \sum_{i=1}^{I} \frac{p_{ik}}{q_k} - l = 0
\end{aligned} \qquad \text{(A10)}$$

for k = 1,… K-1 and

$$\frac{\partial H}{\partial l} = \sum_{k=1}^{K} q_i - 1 = 0 \qquad \text{(A11)}$$

Solving (A10) and (A11) algebraically gives us maximum likelihood estimates of q**.** Since $\sum_{k=1}^{K} p_{ik} = 1$ for all *i*, we can estimate the elements of **q** as

$$q_k = \frac{1}{I}\sum_{i=1}^{I} p_{ik} \tag{A12}$$

Solving (A10) and (A11) algebraically is theoretically possible in that the number of equations and number of unknown parameters are equal. However, in practice it is generally easier to use a numerical maximization procedure to estimate q. To do this, we define the vector

$$B'(q_k) = \left(\frac{\partial H}{\partial b_{0k}}, \frac{\partial H}{\P a_k}, \frac{\partial H}{\P b_k'}\right)$$

as the first derivative of Q with respect to q_k. The Fisher score matrix $C(q_k)$ is derived

by deleting terms with mean 0 in the matrix of second derivatives of Q with respect q_k.

$$C(q_k) = \begin{bmatrix} \sum_{i,t} p_{ik} \frac{f^2(q_k)}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) y_{i,t-1}}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) x_{it}'}{\Phi(q_k)\{1-\Phi(q_k)\}} \\ \sum_{i,t} p_{ik} \frac{f^2(q_k) y_{i,t-1}}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) y_{i,t-1}^2}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) y_{i,t-1} x_{it}'}{\Phi(q_k)\{1-\Phi(q_k)\}} \\ \sum_{i,t} p_{ik} \frac{f^2(q_k) x_{it}}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) y_{i,t-1} x_{it}}{\Phi(q_k)\{1-\Phi(q_k)\}}, & \sum_{i,t} p_{ik} \frac{f^2(q_k) x_{it} x_{it}'}{\Phi(q_k)\{1-\Phi(q_k)\}} \end{bmatrix}$$

The updated estimate $\hat{q}_k$ are then computed by using the Newton-Raphson algorithm

$$\hat{q}_k = \tilde{q}_k + \mathbf{C}^{-1}(\tilde{q}_k)\mathbf{B}(\tilde{q}_k) \tag{A13}$$

for $k = 1, \ldots K$ where $\tilde{q}_k$ is the parameters estimate from previous iteration.

Amemiya (1985) shows that the EM algorithm produces maximum likelihood estimates based on the likelihood function

$$\prod_{i=1}^{I} f\left(\mathbf{y}_{i,T_i}\middle|\mathbf{x}_{i,T_i}, y_{i0},\, x_{i0},\, \mathrm{q}\right) = \prod_{i=1}^{I}\sum_{k=1}^{k} q_k\, f\left(\mathbf{y}_{i,T_i}\middle|\mathbf{x}_{i,T_i}, y_{i0},\, x_{i0},\, ,\, z_{ik}\mathrm{q}_k\right), \tag{A14}$$

which is the likelihood function of a mixture model, as in Wedel and Desarbo (1994, 1995). It is well known that maximum likelihood estimates are consistent under regular conditions. Therefore, the EM estimates, which are exactly the maximum likelihood estimates, are consistent.

To summarize, our procedure is as follows:

Initial step: Choose an arbitrary initial guess as to the values of ($\mathbf{q}$, q).

E-step: Substitute the estimates of ($\mathbf{q}$, q) into equation (A7) to compute updated estimates of $\hat{p}_{ik}$.

M-Step: Obtain maximum likelihood estimates of q by maximizing (A9) using (A13) and $\mathbf{q}$ using (A12).

We repeat the E and M steps until the sum of the absolute distances between all the current and updated parameter estimates is less than 10^{-6}. For discussion of the convergence properties of the EM algorithm, see Amemiya (1985), Render and Walker (1984), Wedel and Desarbo (1994, 1995), and references therein.