State Characteristics and the Location of Foreign Direct Investment within the United States: Minimum Chi-Square Conditional Logit Estimation

Authors	Vachira Arromdee, Cletus C. Coughlin, and Joseph V. Terza
Working Paper Number	1987-006B
Citable Link	https://doi.org/10.20955/wp.1987.006
Suggested Citation	Arromdee, V., Coughlin, C.C., Terza, J.V.; State Characteristics and the Location of Foreign Direct Investment within the United States: Minimum Chi-Square Conditional Logit Estimation, Federal Reserve Bank of St. Louis Working Paper 1987-006. URL https://doi.org/10.20955/wp.1987.006

Published In	Review of Economics and Statistics
Publisher Link	https://doi.org/10.2307/2109406

State Characteristics and the Location of Foreign Direct Investment within the United States: Minimum Chi-Square Conditional Logit Estimation

by

Cletus C. Coughlin
Research Department
Federal Reserve Bank of St. Louis
St. Louis, MO 63166

Joseph V. Terza
Department of Economics
The Pennsylvania State University
University Park, PA 16802

and

Vachira Arromdee
Department of Economics
University of Georgia
Athens, GA 30602

July 1989

A preliminary version of this paper was presented at the Federal Reserve Bank of St. Louis in March 1987. The authors are grateful to the participants for their constructive comments. Thanks are also due Eric Bond for his helpful comments.

State Characteristics and the Location of Foreign Direct Investment within the United States: Minimum Chi-Square Conditional Logit Estimation

I. Introduction

Despite the perception by many of the declining competitiveness of U.S. manufacturing across a broad range of sectors including the high technology sectors, foreign direct investment in U.S. manufacturing has grown at rapid rates. For example, the stock of foreign direct investment in U.S. manufacturing increased from $11.4 billion in 1975 to $91.0 billion in 1987.[1] This anomaly raises numerous questions concerning the determinants and effects of this investment; however, literature reviews by Ó hUallacháin (1986) and Arpan, Flowers, and Ricks (1981) indicate that the majority of these questions remain unanswered.

A subset of questions revolves around the location of foreign direct investment in the U.S. Carlton (1983) concluded that economists know very little about the determinants of new business location. A similar and more emphatic conclusion is appropriate with respect to the determinants of the location of manufacturing foreign direct investment in the U.S. Studies by Heller and Heller (1974), Wilkins (1979), Suzman et al. (1979), and Williams and Brinker (1985) examined a specific state (region). These studies tended to focus upon quantifying the size and scope of foreign direct investment and identifying possible reasons for the investment. To date, only Little (1978), Luger and Shetty (1985), and Glickman and Woodward (1987) have attempted to analyze empirically the determinants of the location of foreign direct investment throughout the United States.[2]

The present analysis focuses upon the pattern of foreign direct investment in manufacturing for a specific year, 1981.[3] This topic is especially interesting in light of the importance of foreign direct investment for economic development at the state level, the virtual absence of previous empirical work on the location determinants of foreign direct investment across all states, and the limited knowledge on new business location. Following Carlton (1983), Bartik (1985), and Luger and Shetty (1985) we develop a Conditional Logit Model (CLM) of the foreign firm's U.S. investment location decision.

In addition to conventional Maximum Likelihood Estimates (MLE) of the CLM, we report results obtained using our newly developed Minimum Chi-Square Estimator (MCE). We developed the MCE for two reasons related to the fact that the basic unit of observation for estimation is the state. First, the sample is small and MLE bias in small samples may be substantial (see Amemiya 1980). Secondly, 20% of the states had no foreign direct investment in the sample year 1981. Using simulated data we show that as the number of states having no investments increases the performance of MLE diminishes relative to MCE in terms of mean square error. This indicates that small sample MLE biases are exacerbated when the data on the dependent variable are sparse.

In the next section the CLM of foreign direct investment is detailed and the MCE is introduced. In Section III the MCE results are presented and analyzed. The potential small sample bias of MLE is evaluated via a simulation experiment in Section IV. The MCE and MLE results are also compared in that section. A final discussion highlights the basic findings concerning the determinants of foreign direct investment location in the U.S.

II. An Econometric Model of the Spatial Distribution of Foreign Direct Investment

In this section, we model the state-level determinants of the frequency distribution of foreign direct investments across the 50 states. For this analysis we take as the population foreign manufacturing firms who have decided to invest in the U.S.[4] Table 1 describes the direct investments across states of a sample of such firms for 1981. The total number of manufacturing foreign direct investments in the U.S. for the year 1981 was 274. California was the leading recipient with 37, while ten states had zero investments.

We assume that a foreign firm will choose to invest in a particular state if and only if doing so will maximize profit. Formally, the j^{th} state is chosen by the i^{th} firm if and only if

$$\Pi^*_{ij} = \max\{\Pi^*_{im};\ m = 1,\ \ldots,\ 50\} \tag{1}$$

where Π^*_{ij} denotes the profit of the i^{th} firm given that it locates in the j^{th} state ($j = 1, \ldots, 50$). Following Carlton (1983) we assume that

$$\Pi_{ij} = c + X_j\beta + \varepsilon_{ij} \tag{2}$$

where $\Pi_{ij} = \ln \Pi^*_{ij}/\theta$, c is the constant term, $X_j = [\ln X^*_{j1}, \ldots, \ln X^*_{jK}]$, $X^*_j = [X^*_{j1}, \ldots, X^*_{jK}]$ is a vector of observable characteristics for the j^{th} state, β is a vector of unknown coefficients to be estimated, ε_{ij} is the random term denoting the unobservable (by the researcher) unique profit advantages to the i^{th} firm from locating in the j^{th} state, and θ is the exponent of the random term in the untransformed version of the profit function.[5]

Assuming that the ε_{ij}'s are independent log-Weibull distributed McFadden (1974) shows that

$$P_j = \exp\{X_j\beta\}/\sum_{k=1}^{50} \exp\{X_k\beta\} \tag{3}$$

where P_j denotes the population relative frequency of locating in state j. The MLE of β is obtained by maximizing the likelihood function

$$L(\beta) = \prod_{j=1}^{50} P_j \tag{4}$$

There is a considerable literature on the problem of small sample bias in maximum likelihood logistic regression and discrimination models (see Anderson and Richardson 1979, Amemiya 1980, Davis 1985, Jennings 1986, Haldane 1955, McLachlan 1980, Schaefer 1983, and Walter 1985). Amemiya (1980) shows for the binary case that in small samples MLE performs poorly relative to Minimum Chi-square Estimation (MCE) in terms of mean square error. Even more troublesome than bias is the potential nonexistence of an optimum of (4) when the sample size is small (see Albert and Anderson 1984). Both of these problems are exacerbated when some of the response cells are empty. Given these potential problems with MLE and because our sample is both small and sparse, we have developed an alternative estimator that is of the simple minimum chi-square type. Our MCE virtually eliminates the existence problem and is designed to reduce bias relative the MLE.

We begin the discussion of our MCE conditional logit model by noting that (3) implies that

$$\log(P_j/P_1) = Z_j\beta \tag{5}$$

where $Z_j = X_j - X_1$. Since we cannot observe P_j, the population relative frequency, we make equation (5) operational by replacing P_j with p_j, the observed relative frequency of investment for the jth state. Following the approach taken by Haldane (1955), Schaefer (1983), and Walter (1985) in correcting for small sample bias in the binary logistic regression model, we show in the appendix that

$$\log\left(\frac{p_j + (1/2n)}{p_1 + (1/2n)}\right) = \log\left(\frac{P_j}{P_1}\right) + e_j = Z_j\beta + e_j \qquad (6)$$

$$(j = 1, \ldots, 49)$$

where n denotes the total number of observed investments in the sample, and e_j is the stochastic term such that $E[e_j] = 0$ except for terms of order smaller than n^{-1} in probability. We also show in the appendix that the covariance matrix of $e = [e_1, \ldots, e_{49}]'$, is $\Omega \equiv D\,\Sigma\,D'$, where Σ is the (50x50) matrix whose typical element is

$$\sigma_{ij} = \begin{cases} P_j(1-P_j) & \text{if } i = j \\ -P_iP_j & \text{if } i \neq j \end{cases}$$

and

D is the (49x50) matrix whose ij^{th} element is

$$d_{ij} = \begin{cases} \frac{1}{P_j} & \text{if } i = j \\ 0 & \text{if } i \neq j \\ -\frac{1}{P_1} & \text{if } j = 1. \end{cases}$$

Rewriting (6) in matrix notation we have

$$y = Z\beta + e \qquad (7)$$

where y is the 49x1 vector whose j^{th} element is $\log\left(\frac{p_j + (1/2n)}{p_1 + (1/2n)}\right)$, and Z is the 49xK matrix whose j^{th} row is Z_j. Note that the addition of the bias correction factor 1/2n to each of the observed state investment proportions ensures that the dependent variable of (6) will be defined even for states with zero investment frequencies. An analogous specification for the binary linear logit model is suggested by Maddala (1983, p. 30) and applied by Voos and Mishel (1986). To exploit efficiency gains we apply the following minimum chi-square estimator to (7)

$$\tilde{\beta} = (Z'\hat{\Omega}^{-1}Z)^{-1}Z'\hat{\Omega}^{-1}y. \tag{8}$$

Note that $\hat{\Omega}$ is obtained by substituting $\hat{P}_j$ for P_j in Ω where $\hat{P}_j = \exp\{X_j\hat{\beta}\}/\sum^{50}\exp\{X_k\hat{\beta}\}$ and $\hat{\beta}$ is the OLS estimator of β. In the appendix $\hat{\beta}$ is shown to be consistent and asymptotically normal with covariance matrix $(Z'\Omega^{-1}Z)^{-1}$. Furthermore, $(Z'\hat{\Omega}^{-1}Z)^{-1}$ is a consistent estimator for $(Z'\Omega^{-1}Z)^{-1}$.

The probability of selecting a specific state for a foreign direct investment transaction depends on the levels of its characteristics that affect profits relative to the levels of these characteristics in other states. A discussion of potential determinants follows the identification of one other relevant factor. The probability of a manufacturing foreign direct investment transaction in a state depends on the number of potential sites for locating the investment. Similar to Bartik (1985), state land area excluding federal land is used as a proxy for the number of potential sites. Bartik found that the land elasticity of new branch plants was approximately one. In other words, a state with twice as many potential sites as another state had a probability of selection twice as large, ceteris paribus. We reexamine this "dartboard theory" of industrial location with respect to foreign direct investment.

The factors affecting profits are obviously those factors affecting the revenues and costs of the investors. On the revenue side, state per capita income is a measure of market demand in a state and is expected to be related positively to foreign direct investment. The plausibility of this argument is reinforced by Glickman and Woodward's (1987) evidence that foreign manufacturers are serving regional markets. On the other hand, state per capita income would likely be an unimportant consideration for a manufacturing facility that serves a national market. Consequently, if state per capita income is a statistically significant determinant, its expected impact on foreign direct investment is positive.

A second variable that might be a proxy for market demand is manufacturing density. States with higher densities of manufacturing activity could attract more foreign direct investment because the foreign investors might be serving existing manufacturers. Manufacturing density, however, could also proxy for agglomeration economies. Irrespective of the interpretation, manufacturing density is expected to be related positively to foreign direct investment. In fact, using a different measure for agglomeration economies, Luger and Shetty (1985) found evidence that foreign plant start-ups in three industries were related positively to agglomeration economies.

With respect to the factors affecting costs, labor market considerations come to mind immediately. Wage rates, the labor-management environment, and the availability of labor are potentially important characteristics of labor markets. Higher wage rates are expected to deter foreign direct investment. Bartik (1985) found that higher wages were a negative, statistically significant determinant of the probability of locating a new branch plant in a state, and Luger and Shetty (1985) found a similar result with respect to foreign direct investment. In addition, Little (1978) found that state wage differentials were relatively more important for foreign than domestic investors. On the other hand, Glickman and Woodward (1987) did not find that wage differentials affected employment growth of foreign-owned firms in a state.

A characteristic of state labor markets that is widely publicized, especially by states with low rates, is the extent of unionized labor. For those states stressing the paucity of union activity, the selling point is managerial freedom to pursue profit maximization unencumbered by the restrictions imposed by union contracts. Bartik (1985) has generated empirical evidence that supports this position. Not only was unionization a deterrent, but the impact was large.

Related to the actual degree of unionization is the legislation governing the rights of employers and employees. Right-to-work legislation is one example. States with this legislation publicize its existence with the expectation that investment will be attracted; however, Glickman and Woodward (1987) found that right-to-work legislation was related negatively to the employment growth of foreign-owned firms in a state.

A final characteristic of labor markets is the unemployment rate. To the extent that the unemployment rate reflects a pool of potential workers, higher unemployment across states will likely be related positively to foreign direct investment. One complicating factor, however, revolves around the effects of unemployment insurance. Higher unemployment insurance benefits tend to increase unemployment rates by increasing the average duration of unemployment and encouraging the unemployed to seek higher paying employment (Ehrenberg and Oaxaca, 1976). The amount that a firm must pay in unemployment insurance premiums is linked to the frequency with which their employees claim benefits, but upper limits on these premiums are well below the actuarial rates for some firms. Consequently, firms with high labor turnover are subsidized by less volatile companies (Feldstein, 1976). This would deter foreign firms with low labor turnover from investing in a state because they would be required to subsidize the unemployed workers who were released by other firms. (On the other hand, the high turnover foreign firm would be encouraged to enter.) The preceding discussion suggests that unemployment rates might not be a good measure of labor availability across states and that factors underlying the higher rates could deter foreign direct investment. Thus, the likely empirical association between unemployment rates and foreign direct investment is uncertain.

The cost of energy is a frequently mentioned factor in governmental attempts to attract investment. Carlton (1983) found that energy costs, especially the price of electricity, affected the location of firms manufacturing fabricated plastic products, communication transmitting equipment, and electronic components. Higher energy costs are expected to relate negatively to foreign direct investment.

Another frequently mentioned consideration in attempts to attract foreign direct investment is the existence of a highly developed transportation network. Consequently, more highway mileage, more railroad mileage, and more public airport facilities, adjusted for state size, are expected to be related positively to foreign direct investment. Empirical support for the importance of highways in the location of new branch plants has been presented by Bartik (1985).

The existence of a highly developed transportation infrastructure depends on public funding, which leads to the possible effects of taxes on business location decisions. Newman and Sullivan (1988) concluded, contrary to

Wasylenko's (1981) earlier conclusion that taxes have very little effect on interregional business location decisions, that this is an unsettled question. The continued use of tax and fiscal inducements suggests that policymakers believe that taxes can affect business location, while recent econometric evidence is mixed. For example, Luger and Shetty (1985) found that higher tax rates were related negatively to foreign direct investment in one industry, but also found a positive relationship in another industry and no significant relationship in a third industry.[6]

A major problem in attempting to assess the impact of taxes on business location decisions is measuring state tax burdens. Two common methods of comparing state tax burdens are state and local taxes per capita and state and local taxes as a percentage of personal income. There are numerous problems with these measures (Kieschnick, 1983). Included among the problems are identifying the incidence of a tax, the possibility that the taxes are financing the provision of goods and services valued by business, and the possibility of the use of tax incentives. Thus, even though higher taxes are expected to be a deterrent, the preceding problems might prevent the anticipated empirical finding.

There are also special tax issues associated with foreign direct investment. The regional issue that has generated the most controversy is the use of unitary taxation by a number of states, but empirical evidence on the foreign direct investment impacts of unitary taxation is limited to Glickman and Woodward (1987). They found that unitary taxation reduced the employment growth of foreign-owned firms. The controversy stems from the difficulties of dividing a multi-jurisdictional company's taxable income across jurisdictions (Tannenwald, 1984). If a state determines that a corporate taxpayer is a component of a larger enterprise, then a state with unitary taxation taxes a fraction of the enterprise's worldwide income rather than the income earned in the state. (The fraction is usually determined by the state's share of the enterprise's worldwide payroll, sales, and property.) Multinational corporations have raised numerous objections to worldwide unitary taxation with the exposure to double taxation and the burdens of additional accounting requirements being primary objections. On the other hand, state tax officials argue that worldwide unitary accounting is the only method that prevents multinational corporations from evading taxes by reallocating profits from high-tax areas to low-tax areas. In 1981 there were five states using unitary taxation characterized as total worldwide combination - Alaska, California, Idaho, North Dakota, and Oregon.

In addition to taxation, state government spending policies can affect the business location decisions of foreign investors. Numerous categories of state expenditures such as spending on education and highways could affect foreign direct investment. Luger and Shetty (1985) have presented suggestive evidence on this issue.[7] In the present study one expenditure targeted for influencing foreign direct investment is examined. State governments promote foreign direct investment primarily through the provision of information and investment-incentive packages (Kline, 1982). A measure of investment promotion is the level of state expenditures on this activity. These expenditures are designed to increase international awareness about the state by providing business-related information to potential investors. This information is generally of a comparative nature and includes items such as wage rates, unionization rates, energy costs, transportation facilities, taxes, and educational facilities.

In summary, foreign direct investment at the state level is expected to be affected by the number of potential sites, demand, agglomeration economies, labor market characteristics, energy costs, transportation facilities, taxes, and promotion expenditures. The definitions of all independent variables and their expected impacts on foreign direct investment are listed in Table 2.

III. Empirical Results

The key empirical results obtained by the minimum chi-square method are listed in Table 3. Given the large number of potential independent variables, the results of six regressions are presented. The adjusted R^2, which ranges from .65 to .71, indicates that the model performed well. Since the underlying profit function is log-linear, the coefficients of the explanatory variables can be roughly interpreted as elasticities.[8]

The variables that appear in all variants are discussed first. Land area, the proxy for the number of sites, is a positive, statistically significant determinant of foreign direct investment location. The range of coefficient estimates from .66 to .82 suggests that the "dartboard theory" of industrial location is not strictly applicable to foreign direct investment.[9] More foreign direct investment transactions occur in larger states, but the relationship is not equiproportionate.

State per capita income, the proxy for market demand, is a positive, statistically significant determinant of foreign direct investment location. The range of coefficient estimates from 7.5 to 8.7 indicates that foreign direct investment is very responsive to differences in per capita income across states.

The importance of demand is also suggested by the statistical significance of manufacturing density; however, as suggested earlier this variable could also capture agglomeration economies. Irrespective of exactly what this variable is capturing, the more dense the manufacturing activity, the more likely is foreign direct investment to occur.

The results show that higher wage rates deter foreign direct investment with the elasticity estimate ranging from -6.6 to -5.0. These wage elasticity estimates are substantially greater in absolute value than the -.9 estimate by Bartik (1985) of the wage elasticity of new branch plants and slightly greater than Luger and Shetty's (1985) estimates for the wage elasticity of foreign plant start-ups of -3.0 in drug manufacturing and -4.4 in motor vehicle production.

Another labor market characteristic that appears in all variants is the unemployment rate. The results show that the unemployment rate is a positive, statistically significant determinant of foreign direct investment. Thus the unemployment rate is a signal of the availability of labor that affects investors.

The final variable common to all variants is state government spending targeted to attract foreign direct investment. These expenditures have a positive, statistically significant effect on foreign direct investment in every variant except the first. Even though the efficiency of these expenditures cannot be assessed in this analysis, the expenditures are affecting foreign direct investment according to the desires of state officials.

Due to the large number of potential explanatory variables, the remaining variables appear in only selected regressions. The empirical results of incorporating a proxy for energy costs are mixed. While specifications were found that yielded a negative, statistically significant result, generally speaking energy costs were not statistically significant. Variant 2 is a representative case.

The results indicate that taxes deter foreign direct investment. Variants 2 and 6 show that state and local taxes per capita is a negative, statistically significant determinant of foreign direct investment. Results that are not reported reveal that the use of state and local taxes as a percentage of personal income generates a similar result. In addition to the deterrent effects of the preceding general measures of state taxation, specific tax measures also deter foreign direct investment. States using "total worldwide combination" unitary taxation tend to be the sites for less foreign direct investment. This result is highlighted in variants 1 and 3.

The transportation infrastructure of a state appears to affect foreign direct investment. Variants 4, 5, and 6 reveal that, after adjusting for state size, highway miles, railroad miles, and the number of public airports are positive determinants of foreign direct investment. The inclusion of the highway and railroad variables, however, makes state and local taxes per capita insignificant.

While none of the preceding results are surprising, the remaining results conflict with expectations. Variants 2-6 indicate that higher rates of unionization are related positively to foreign direct investment. There is no obvious explanation for this result; however, a few comments are in order. First, the inclusion of unionization causes the coefficient estimate and the t-ratio for the unemployment rate to be reduced. The simple correlation coefficient for these two variables is .47, so unionization might be capturing some of the effect of unemployment. Second, the finding does not imply that the labor associated with the foreign direct investment will more likely be unionized. Third, either the industrial composition or the source country composition of foreign direct investment in 1981 could account for this result. The consistency of this finding for other years and for specific industries and source countries should be examined further.

In conjunction with the surprising impact of unionization is the empirical finding, not reported, that states with right-to-work legislation tended to have fewer foreign direct investments than states without the legislation. In light of the high correlation between unionization and right-to-work legislation (-.67), this empirical finding simply corroborates the surprising unionization finding.

IV. Minimum Chi-Square Estimation vs. Maximum Likelihood Estimation

To explore the differences between Minimum Chi-Square Estimation (MCE) and Maximum Likelihood Estimation (MLE), we also estimated the six variants of the model by the maximum likelihood method. Following a comparison of the results in terms of coefficient estimates and t-statistics for individual variables, we describe a Monte Carlo experiment designed to illustrate the potential severity of the MLE bias and how that bias is affected as the number of states with zero investments increases.

A comparison of the MCE results in Table 3 with the MLE results in Table 4 produces numerous differences between the two methods. For land area, MLE generates higher coefficient estimates and makes the conclusion that the dartboard theory of industrial location does not hold less tenable. For example, the null hypothesis that the coefficient is equal to one can be rejected at the .10 significance level for variants 1-5, but cannot be rejected for variant 6. Recall that using MCE the null hypothesis was rejected at the .01 significance level for variants 1-5 and at the .10 level for variant 6.

For manufacturing density, the coefficient estimates using MLE are higher, but no conclusions concerning statistical significance are altered. A similar statement concerning statistical significance is appropriate for state per capita income; however, MLE generates relatively lower coefficient estimates that MCE.

For the labor market characteristics, the coefficient estimates using MLE are lower for the wage (in absolute value), unemployment, and unionization rates than when using MCE. For each variable, the t-statistic using MLE is also substantially lower than using MCE. For the unemployment rate, the lower t-statistic means that using a significance level of .01 precludes a rejection of the hypothesis that the coefficient is equal to zero in 5 of the 6 variants.

A comparison of the public policy variables across the two methods yields various results. Generally speaking, per capita state and local taxes are not statistically significant using MLE, while the variable is statistically significant using MCE. For both methods, unitary taxation is statistically significant, but MLE produces absolutely higher coefficient estimates. The two methods yield virtually identical results for promotional expenditures to attract foreign direct investment.

For the remaining variables, energy costs and the three variables measuring transportation infrastructure, the estimation methods yield no differential results. Thus conclusions based on hypothesis tests are identical across estimation methods.

To further explore the differences between the estimation methods, we conducted a Monte Carlo experiment and applied both MCE and MLE to the simulated data. For the experimental design matrix we chose the actual sampled design matrix for variant 4 of the model and set the coefficients equal to those found in the fourth row of Table 3. We chose variant 4 as the model to simulate because it yielded the best fit with regard to adjusted R^2. We varied the sample size from 25 to 300 in increments of 25. The simulation results are reported in Table 5. The first column of Table 5 gives the sample size and the second shows the number of states out of 50 (on average over 200 replications) that had no investments. The remaining ten columns of the upper half of the table contain the ratio of mean square error for MCE to that of MLE for each of the respective coefficients of the model. Values less that 1.0, therefore, indicate the dominance of MCE. In the lower half of the table similar mean square error ratios are reported for the estimated variances of the coefficients.

The simulated sample size most nearly comparable to our actual data in terms of the number of zeros generated is somewhere between 200 and 225. At these sample sizes MCE and MLE appear to perform equally well. Note, however, that as the sample size decreases and therefore the number of states with no investment increases the bias of MLE relative to MCE increases in almost every case, indicating that the sparser is the data the more dominant MCE becomes. This basic result holds for both coefficient and variance estimation. Finally, we note that at a sample size of 25 the MLE broke down, indicating possible existence problems at very small sample sizes.

IV. Conclusion

The present research is an initial exploration into a topic that is of increasing importance to the U.S. in general and is of direct relevance to the economic development efforts of individual states. In view of the limited knowledge concerning business location decisions, the paucity of previous studies concerning the spatial distribution of foreign direct investment is not surprising. A model of the foreign firm's investment location decision based on profit maximization was developed. The model yielded a linear version of McFadden's (1974) conditional logit model. To anticipate the problem of small sample bias of maximum likelihood estimation and the consequences of a high frequency of zero investments, a minimum chi-square estimator was developed. A comparison of the results using the two methods revealed the desirability of minimum chi-square estimation in the present case.

Specific results were provided and discussed previously, so only certain summary results using the minimum chi-square estimator are highlighted. First, the number of potential sites is a key determinant, but the "dartboard theory" of industrial location does not appear to apply. Second, per capita income, a proxy for market demand, affects foreign direct investment. Third, manufacturing density is related positively to foreign direct investment. Fourth, characteristics of the labor market affected the distribution of foreign direct investment. Higher wages deterred foreign direct investment, while higher unemployment rates tended to attract foreign direct investment. Surprisingly, higher unionization rates were not associated with reduced foreign direct investment, but rather with increased foreign direct investment. Fifth, there is some evidence that higher transportation infrastructures tend to attract relatively more foreign direct investment. Sixth, contrary to the bulk of the previous literature dealing with the effect of taxes on business location decisions, taxes affect the location of foreign direct investment. There is strong evidence that unitary taxation deters foreign direct investment. More general measures of the overall level of taxation indicated a negative relationship between taxation and foreign direct investment. The role of government is not limited to taxation.

Expenditures to attract foreign direct investment are related positively to foreign direct investment.

In addition to further investigation of the impact of unionization numerous topics remain for in future research. While the current results suggest that state government taxation and spending affect foreign direct investment, this topic deserves additional scrutiny especially with respect to fiscal incentives. Another avenue of research is to investigate the possibility that the location determinants vary across both countries and industries. Via disaggregation it should be possible to ascertain if the location determinants of foreign direct investment from different source countries differ. It is also reasonable to anticipate that the location determinants of foreign direct investment differ by industry as well.

FOOTNOTES

1. The figure for 1975 was taken from *Selected Data on Foreign Direct Investment in the United States, 1950-79* and is based on 1974 benchmark data (see U.S. Department of Commerce, 1984). The figure for 1986 was taken from an article in the *Survey of Current Business*, "Foreign Direct Investment in the United States: Detail for Position and Balance of Payments Flows, 1987," and is based on 1980 benchmark data (see U.S. Department of Commerce, 1988). The estimates are sensitive to the benchmarks, but for present purposes the figures are not so sensitive as to raise doubts about the rapid increase in foreign direct investment in the United States.

2. Other research dealing with foreign direct investment throughout the United States is limited. Using various measures, McConnell (1980), Little (1983), O'hUallacháin (1985), and Arpan and Ricks (1986) have identified the changes over time in the pattern of foreign direct investment. Little (1986) found that the employment directly associated with foreign direct investment mitigated the transitional and cyclical stresses experienced by states between 1979 and 1983.

3. A more recent year than 1981 could have been chosen. This year was chosen for two reasons. The data on promotion expenditures to attract foreign direct investment is not as complete for more recent years and unitary taxation has become less frequent in more recent years.

4. The Department of Commerce (1982) defines foreign direct investment as the direct or indirect ownership by a foreign entity of 10% or more of the voting securities of an incorporated business enterprise or an equivalent interest in an unincorporated business enterprise. A foreign direct investment transaction in manufacturing could involve an acquisition, a merger, an equity increase, a joint venture, a new plant, or a plant extension.

5. See Carlton (1983), p. 441.

6. Evidence of conflicting results is plentiful. For example, Carlton (1983, p. 441) concluded that taxes and state incentive programs did not have major effects on the location of new branch plants across standard metropolitan statistical areas; however, Bartik (1985) found that state taxes deterred the location of new branch plants at the state level.

7. Luger and Shetty (1985) used an "efforts index" to summarize the various state programs to encourage foreign direct investment. This index included state programs such as subsidized job training, financing for industrial development, land and building subsidies, research and development assistance, and unemployment compensation. Foreign direct investment was found to be related positively to this efforts index.

8. The logit specification implies that

$$\partial \log P_j / \partial \log X_k = \beta_k (1 - P_j)$$

and since the average estimated value of P_j is .02

$\beta_k \simeq \partial \log P_j / \partial \log X_k$.

9. The null hypothesis that the coefficient is equal to one is rejected at the .01 significance level for variants 1-5 and at the .10 significance level for variant 6.

REFERENCES

Albert A., and J. A. Anderson. "On the Existence of Maximum Likelihood Estimates in Logistic Regression Models," *Biometrika* 71 (1984), 1-10.

Amemiya, Takeshi. "The n^{-2}-Order Mean Squared Errors of the Maximum Likelihood and the Minimum Logit Chi-square Estimator," *Annals of Statistics* 8 (1980), 488-505.

Anderson, J. A., and S. C. Richardson. "Logistic Discrimination and Bias Correction in Maximum Likelihood Estimation," *Technometrics* 21 (1979), 71-78.

Arpan, Jeffrey S., Edward B. Flowers, and David A. Ricks. "Foreign Direct Investment in the United States: The State of Knowledge in Research," *Journal of International Business Studies* 12 (Spring/Summer 1981), 137-153.

Arpan, Jeffrey S. and David A. Ricks. "Foreign Direct Investment in the U.S., 1974-1984," *Journal of International Business Studies* 17 (Fall 1986), 149-153.

Bartik, Timothy J. "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States," *Journal of Business and Economic Statistics* 3 (January 1985), 14-22.

Berry, Willard M. and William A. Mussen. "Part I - A Report," in *Export Development and Foreign Investment: The Role of the States and Its Linkage to Federal Action*, Committee on International Trade and Foreign Relations, Washington, D.C.: National Governors' Association, 1980.

Carlton, Dennis W. "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables," *Review of Economics and Statistics* 65 (August 1983), 440-449.

Davis, Linda J. "Modification of the Empirical Logit to Reduce Bias in Simple Linear Logistic Regression, *Biometrika* 72 (1985), 199-202.

Ehrenberg, R. G. and R. L. Oaxaca. "Unemployment Insurance, Duration of Unemployment, and Subsequent Wage Gain," *American Economic Review* 66 (December 1976), 754-766.

Feldstein, Martin. "Temporary Layoffs in the Theory of Unemployment," *Journal of Political Economy* 84 (October 1976), 937-957.

Glickman, Norman J., and Douglas P. Woodward. *Regional Patterns of Manufacturing Foreign Direct Investment in the United States,* Special Report for U.S. Department of Commerce, Economic Development Administration, 1987.

Haldane, J. B. S. "The Estimation and Significance of the Logarithm of a Ratio of Frequencies," *Annals of Human Genetics* 20 (August 1955), 309-311.

Heller, H. Robert and Emily E. Heller. *Japanese Investment in the United States, With a Case Study of the Hawaiian Experience*, New York: Praeger Publishers, 1974.

Jennings, D. E. "Judging Inference Adequacy in Logistic Regression," *JASA* 81 (1986), 471-76.

Kieschnick, Michael. "Taxes and Growth: Business Incentives and Economic Development," in Michael Barker (ed.), *State Taxation Policy*, Durham, NC: Duke Press Policy Studies, 1983, 153-282.

Kline, John M. *State Government Influence on U.S. International Economic Policy*, Lexington, MA: Lexington Books, 1982.

Little, Jane Sneddon. "The Effects of Foreign Direct Investment on U.S. Employment during Recession and Structural Change," *New England Economic Review* (November/December 1986), 40-48.

Little, Jane Sneddon. "Foreign Investors' Locational Choices: An Update," *New England Economic Review* (January/February 1983), 28-31.

Little, Jane Sneddon. "Locational Decisions of Foreign Direct Investors in the United States," *New England Economic Review* (July/August 1978), 43-63.

Luger, Michael I., and Sudhir Shetty. "Determinants of Foreign Plant Start-Ups in the United States: Lessons for Policymakers in the Southeast," *Vanderbilt Journal of Transnational Law* 18 (Spring 1985), 223-45.

Maddala, G. S. *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge, MA: Cambridge University Press, 1983.

McConnell, James E. "Foreign Direct Investment in the United States," *Annals of the Association of American Geographers* 70 (June 1980), 259-270.

McFadden, Daniel. "Conditional Logit Analysis of Qualitative 'Choice Behavior," in Paul Zarembka (ed.), *Frontiers in Econometrics*, New York: Academic Press, 1974, 105-142.

McLachlan, G. J. "A Note on Bias Correction in Maximum Likelihood Estimation with Logistic Discrimination," *Technometrics* 22 (1980), 621-27.

Newman, Robert J., and Dennis H. Sullivan. "Econometric Analysis of Business Tax Impacts on Industrial Location: What Do We Know, and How Do We Know It?" *Journal of Urban Economics* 23 (March 1988), 215-34.

Ó hUallacháin, Breandán. "The Role of Foreign Direct Investment in the Development of Regional Industrial Systems: Current Knowledge and Suggestions for a Future American Research Agenda," *Regional Studies* 20 (April 1986), 151-162.

Ó hUallacháin, Breandán. "Spatial Pattern of Foreign Direct Investment in the United States," *Professional Geographer* 37 (May 1985), 154-162.

Schaefer, Robert L. "Bias Correction in Maximum Likelihood Logistic Regression," *Statistics in Medicine* 2 (1983), 71-78.

Serfling, Robert J. *Approximation Theorems of Mathematical Statistics*, Wiley, 1980

Suzman, Cedric L., and John M. Stopford, Peter W. Becker, Lawrence G. Franko, Mira Wilkins, and Jeffrey S. Arpan. *Foreign Direct Investment in the Southeast: West Germany, the United Kingdom, and Japan*, Atlanta: The Southern Center for International Studies, 1979.

Tannenwald, Robert. "The Pros and Cons of Worldwide Unitary Taxation," *New England Economic Review* (July/August 1984), 17-28.

U.S. Department of Commerce, Bureau of Economic Analysis. "Foreign Direct Investment in the United States: Detail for Position and Balance of Payments Flows, 1987," *Survey of Current Business* 68 (August 1988), 69-83.

U.S. Department of Commerce, Bureau of Economic Analysis, *Selected Data on Foreign Direct Investment in the United States, 1950-79*, Washington, DC: U.S. Government Printing Office, 1984.

U.S. Department of Commerce, International Trade Administration. *Foreign Direct Investment in the United States - 1981 Transactions*, Washington, DC: U.S. Government Printing Office, 1982.

Voos, Paula B., and Lawrence R. Mishel, "The Union Impact on Profits: Evidence from Industry Price-Cost Margin Data," *Journal of Labor Economics* 4 (January 1986), 105-133.

Walter, S. D. "Small Sample Estimation of Log Odds Ratios from Logistic Regression and Fourfold Tables," *Statistics in Medicine* 4 (1985), 437-44.

Wasylenko, Michael. "The Location of Firms: The Role of Taxes and Fiscal Incentives," in Roy Bahl (ed.), *Urban Government Finance*, Beverly Hills: Sage Publications, 1981, 155-190.

Wilkins, Mira. *Foreign Enterprise in Florida*, Miami: University Presses of Florida, 1979.

Williams, Steven D. and William J. Brinker. "A Survey of Foreign Firms Recently Locating in Tennessee," *Growth and Change* 16 (July 1985), 54-63.

Table 1

Manufacturing Foreign Direct Investment - 1981

State	Number	Rank	State	Number	Rank
Alabama	3	21-25	Montana	0	41-50
Alaska	1	32-40	Nebraska	0	41-50
Arizona	2	26-31	Nevada	2	26-31
Arkansas	2	26-31	New Hampshire	1	32-40
California	37	1	New Jersey	14	7
Colorado	3	21-25	New Mexico	1	32-40
Connecticut	12	8	New York	34	2
Delaware	4	17-20	North Carolina	4	17-20
Florida	10	9-10	North Dakota	0	41-50
Georgia	6	14-15	Ohio	17	3-4
Hawaii	0	41-50	Oklahoma	3	21-25
Idaho	0	41-50	Oregon	1	32-40
Illinois	10	9-10	Pennsylvania	16	5-6
Indiana	1	32-40	Rhode Island	3	21-25
Iowa	1	32-40	South Carolina	2	26-31
Kansas	0	41-50	South Dakota	0	41-50
Kentucky	4	17-20	Tennessee	6	14-15
Louisiana	1	32-40	Texas	17	3-4
Maine	2	26-31	Utah	0	41-50
Maryland	3	21-25	Vermont	1	32-40
Massachusetts	16	5-6	Virginia	8	11
Michigan	4	17-20	Washington	7	12-13
Minnesota	1	32-40	West Virginia	2	26-31
Mississippi	0	41-50	Wisconsin	7	12-13
Missouri	5	16	Wyoming	0	41-50

Source: U.S. Department of Commerce, International Trade Administration. Foreign Direct Investment in the United States - 1981 Transactions.

Table 2

Definitions and Expected Impacts of Independent Variables

LAND	-	natural logarithm of state land excluding federal land (+)
PINC	-	natural logarithm of state per capita income (+)
MANL	-	natural logarithm of state manufacturing employment per square mile of state land excluding federal land (+)
WAGE	-	natural logarithm of average state wage of production workers in manufacturing (-)
UNION	-	natural logarithm of percentage of unionized employees (-)
RW	-	dummy variable equal to one if a state has right-to-work legislation and zero otherwise (+)
UNEM	-	natural logarithm of state unemployment rate (?)
EN	-	natural logarithm of energy costs per dollar of value-added in manufacturing (-)
HIWAY	-	natural logarithm of highway miles per square mile of state land area (+)
RR	-	natural logarithm of railroad miles per square mile of state land area (+)
AIR	-	natural logarithm of the number of public airport facilities per square mile of state land area (+)
PTAX	-	natural logarithm of state and local taxes per capita (-)
TAXSPI	-	natural logarithm of state and local taxes as a percentage of state personal income (-)
TUNIT	-	dummy variable equal to one if a state has "total worldwide combination" unitary taxation and zero otherwise (-)
PROM	-	natural logarithm of state expenditures to attract foreign direct investment (+)

Table 3

Minimum Chi-square Estimates

VARIANT #	LAND	PINC	WAGE	UNEM	PROM	MANL	PTAX	EN	TUNIT	HIWAY	RR	AIR	UNION	R^2 ADJ
1	0.744[a]	8.511[a]	-5.040[a]	1.924[a]	0.060	0.397[a]		0.169	-.0996[a]					.65
	(9.01)	(9.77)	(-7.51)	(5.21)	(1.65)	(6.12)		(0.82)	(-4.03)					
2	0.656[a]	8.663[a]	-6.611[a]	1.570[a]	0.146[a]	0.152[a]	-0.691[b]						1.103[a]	.68
	(8.51)	(8.86)	(-9.41)	(3.66)	(4.06)	(2.74)	(-2.58)						(4.36)	
3	0.754[a]	7.537[a]	-5.658[a]	1.436[a]	0.076[b]	0.354[a]			-0.793[a]				0.711[a]	.69
	(8.96)	(8.55)	(-8.13)	(3.49)	(2.05)	(5.64)			(-3.16)				(2.84)	
4	0.738[a]	8.231[a]	-5.929[a]	1.673[a]	0.108[a]	.0116[c]	0.036			0.487[b]			0.733[b]	.70
	(8.74)	(8.40)	(-8.26)	(3.88)	(2.79)	(1.81)	(0.10)			(2.60)			(2.61)	
5	0.690[a]	7.681[a]	-5.821[a]	1.454[a]	0.097[b]	0.208[a]	0.094				0.354[a]		0.708[b]	.69
	(8.59)	(7.81)	(-8.19)	(3.42)	(2.44)	(3.30)	(0.26)				(2.75)		(2.55)	
6	0.822[a]	8.100[a]	-6.142[a]	1.239[a]	0.093[b]	0.169[a]	-0.501[c]					0.630[a]	1.083[a]	.71
	(8.59)	(8.12)	(-8.50)	(2.79)	(2.29)	(3.05)	(-1.82)					(2.97)	(4.29)	

[a] statistically significant at the .01 level (two-sided)

[b] statistically significant at the .05 level (two-sided)

[c] statistically significant at the .10 level (two-sided)

Table 4

Maximum Likelihood Estimates

VARIANT #	LAND	PINC	WAGE	UNEM	PROM	MANL	PTAX	EN	TUNIT	HIWAY	RR	AIR	UNION	LIKE-LIHOOD RATIO STAT
1	0.846[a]	8.091[a]	-4.337[a]	1.611[a]	0.095[c]	0.519[a]		0.203	-0.549[b]					313.57
	(9.63)	(8.71)	(-5.23)	(3.27)	(1.94)	(8.36)		(0.90)	(-2.57)					
2	0.767[a]	7.056[a]	-4.710[a]	0.960[c]	0.139[a]	0.388[a]	-0.437						0.784[a]	317.58
	(9.51)	(6.88)	(-5.55)	(1.75)	(2.85)	(5.85)	(-1.57)						(3.19)	
3	0.846[a]	6.913[a]	-4.654[a]	1.037[c]	0.103[b]	0.473[a]			-0.429[c]				0.613[b]	318.98
	(9.50)	(7.45)	(-5.57)	(1.93)	(2.10)	(7.57)			(-1.96)				(2.49)	
4	0.843[a]	7.036[a]	-4.527[a]	1.235[b]	0.109[b]	0.319[a]	0.095			0.412[c]			0.521[c]	321.61
	(9.20)	(6.90)	(-5.41)	(2.25)	(2.19)	(4.08)	(0.24)			(1.95)			(1.85)	
5	0.823[a]	6.853[a]	-4.668[a]	1.125[b]	0.098[c]	0.376[a]	0.127				0.323[b]		0.514[c]	323.10
	(9.54)	(6.69)	(-5.50)	(2.07)	(1.92)	(5.24)	(0.330)				(2.32)		(1.88)	
6	0.946[a]	6.658[a]	-4.767[a]	0.892[c]	0.094[c]	0.378[a]	-0.292					0.632[b]	0.854[a]	324.74
	(8.78)	(6.44)	(-5.68)	(1.70)	(1.77)	(5.22)	(-0.99)					(2.63)	(3.34)	

[a]statistically significant at the .01 level (two-sided)

[b]statistically significant at the .05 level (two-sided)

[c]statistically significant at the .10 level (two-sided)

Table 5

Mean Square Error Ratios MCE/MLE

Coefficient Estimates

ZEROS	SAMPLE SIZE	LAND	PINC	WAGE	UNEM	PROM	MANL	EN	PTAX	UNION	AIR
7.52	300	1.21	0.99	1.74	1.79	1.08	1.09	1.81	0.93	1.03	1.32
8.38	275	1.47	0.89	2.24	1.89	1.13	0.80	1.76	0.85	1.14	1.34
9.38	250	1.18	1.01	1.50	1.47	0.98	0.88	1.57	0.99	1.01	1.17
9.89	225	1.33	0.91	1.67	1.33	1.06	0.76	1.62	0.89	0.98	1.00
11.22	200	1.20	0.93	1.75	1.26	0.96	0.79	1.68	0.80	1.14	0.96
12.69	175	1.20	0.86	1.80	1.26	0.82	0.71	1.64	0.82	0.99	0.88
13.96	150	1.10	0.94	1.46	1.36	0.94	0.74	1.57	0.71	0.93	0.87
16.53	125	1.18	0.67	1.02	0.98	0.83	0.53	1.22	0.57	0.85	1.00
18.36	100	1.21	1.28	1.16	1.09	0.82	0.43	1.24	0.66	0.73	0.81
21.85	75.00	1.49	1.25	1.02	0.80	0.59	0.39	0.90	0.47	0.63	1.03
26.60	50.00	1.37	1.30	1.20	0.48	0.33	0.24	0.71	0.34	0.57	0.61

Variance Estimates

7.52	300	1.62	0.75	19.35	106.90	1.60	0.33	12.85	1.03	0.36	12.71
8.38	275	0.49	1.39	22.62	122.80	1.64	0.65	161.50	1.44	0.23	0.95
9.38	250	2.69	0.53	39.16	31.36	1.20	0.89	142.50	1.13	2.13	11.90
9.89	225	0.39	0.79	15.02	29.83	0.12	0.47	29.33	1.10	0.64	1.15
11.22	200	0.73	0.31	67.10	11.33	0.39	0.42	68.22	0.35	4.01	0.52
12.69	175	0.58	1.63	18.24	10.47	0.08	0.38	181.70	1.40	0.76	0.41
13.96	150	0.38	0.91	26.84	11.39	0.19	0.36	41.50	2.75	0.40	0.34
16.53	125	0.46	0.34	3.21	5.19	0.14	0.80	9.22	1.09	0.15	0.96
18.36	100	0.38	2.74	1.76	5.00	0.06	0.91	23.37	2.02	1.74	0.32
21.85	75.00	2.11	3.02	0.15	0.15	0.55	1.04	0.60	2.42	3.34	0.95
26.60	50.00	0.49	0.97	0.35	0.54	0.17	0.82	0.39	1.48	0.73	0.94

Appendix

Asymptotic Properties of the Minimum Chi-square Estimator

In this appendix two important theorems will be stated and proven. The first motivates our choice of 1/2n as the bias correction factor. As will be shown this value ensures that the expectation of error vector e in the estimating equation (7) will converge to zero at a rate faster than n^{-1}. Note that this means that the expected value of the left-hand side of (6) converges to $\log(P_j/P_1)$ faster than if it were merely asymptotically unbiased. The second theorem establishes the consistency and asymptotic normality of the minimum chi-square estimator (8).

THEOREM 1: $E[e] = o(n^{-1})$.

Proof:

Let e^* be the 49x1 vector whose typical element is

$$e_j^* = \log\left(\frac{p_j + t^*}{p_1 + t^*}\right) - \log\left(\frac{P_j}{P_1}\right)$$

where $t^* = t/n$ and t is an arbitrary positive constant. Following Haldane (1955) we write, ν_j, the number of observations in state j as

$$\nu_j = nP_j + \eta_j$$

where n denotes the total number of observations in the sample, and η_j is the random error term. Note that $\nu = [\nu_1, \ldots, \nu_j]$ is multinomially distributed so $E[\eta_j] = 0$, $E[\eta_j^2] = nP_j(1 - P_j)$, and $E[\eta_j\eta_i] = -nP_jP_i$.

It can easily be shown that

$$e_j^* = \log\left(1 + \frac{t + \eta_j}{nP_j}\right) - \log\left(1 + \frac{t + \eta_1}{nP_1}\right).$$

We seek the value of t for which $E[e^*] = o(n^{-1})$.

Taylor expanding e_j^* yields

$$e_j^* = \frac{t + \eta_j}{nP_j} - \frac{(t + \eta_j)^2}{2n^2P_j^2} + \frac{(t + \eta_j)^3}{3n^3P_j^3} - \ldots..$$

$$- \frac{t + \eta_j}{nP_1} + \frac{(t + \eta_1)^2}{2n^2P_1^2} - \frac{(t + \eta_1)^3}{3n^3P_j^3} + \ldots..$$

$$= \frac{t}{nP_j} + \frac{\eta_j}{nP_j} - \frac{t^2}{2n^2P_j^2} - \frac{2t\eta_j}{2n^2P_j^2} - \frac{\eta_j^2}{2n^2P_j^2}$$

$$+ \frac{t^3}{3n^3P_j^3} + \frac{t^2\eta_j}{n^3P_j^3} + \frac{t\eta_j^2}{n^3P_j^3} + \frac{\eta_j^3}{3n^3P_j^3} - \ldots..$$

$$\ldots - \frac{t}{nP_1} - \frac{\eta_1}{nP_1} + \frac{t^2}{2nP_1^2} + \frac{2t\eta_1}{2n^2P_1^2} + \frac{\eta_1^2}{2n^2P_1^2}$$

$$- \frac{t^3}{3nP_1^3} - \frac{t^2\eta_1}{n^3P_1^3} - \frac{t\eta_1^2}{n^3P_1^3} - \frac{t^2\eta_1}{n^3P_1^3} - \frac{\eta_1^3}{3n^3P_1^3} + \ldots..$$

$$= \left[\frac{t}{nP_j} - \frac{t}{nP_1} - \frac{t^2}{2n^2P_j^2} + \frac{t^2}{2n^2P_1^2} + \frac{t^3}{3n^3P_j^3} - \frac{t^3}{3n^3P_1^3} + \ldots..\right]$$

$$+ \left[\frac{1}{nP_j} - \frac{t}{n^2P_j^2} + \frac{t^2}{n^3P_j^3} - \ldots..\right]\eta_j$$

$$- \left[\frac{1}{nP_1} - \frac{t}{n^2P_1^2} + \frac{t^2}{n^3P_1^3} - \cdots\cdots\right] \eta_1$$

$$- \left[\frac{1}{2n^2P_j^2} - \frac{t}{n^3P_j^3} + \cdots\right] \eta_j^2$$

$$+ \left[\frac{1}{2n^2P_1^2} - \frac{t}{n^3P_1^3} + \cdots\right] \eta_1^2 + \cdots\cdots;$$

Therefore, we can write

$$e_j^* = u_j + \phi_j$$

where

$$u_j = \frac{t}{nP_j} - \frac{t}{nP_1} + \frac{\eta_j}{nP_j} - \frac{\eta_1}{nP_1} - \frac{\eta_j^2}{2n^2P_j^2} + \frac{\eta_1^2}{2n^2P_1^2}$$

and $E[\phi_j]$ is $o(n^{-1})$. We must find the value of t for which $E[u_j] = 0$. Now

$$E[u_j] = \frac{t}{nP_j} - \frac{t}{nP_1} - \frac{nP_j(1 - P_j)}{2n^2P_j^2} + \frac{nP_1(1 - P_1)}{2n^2P_1^2}$$

$$= \frac{2t(nP_jP_1^2 - nP_j^2P_1) - (nP_jP_1^2 - nP_j^2P_1)}{2n^2P_j^2P_1^2}$$

so if $t = \frac{1}{2}$ then $E[u_j] = 0$ which implies that $E[e_j] = o(n^{-1})$.

Now we state and prove two useful lemmas.

LEMMA 1: Let e* be the (49x1) vector defined in Theorem 1 and let e^o be the (49x1) vector whose typical element is

$$\log(\frac{p_j}{p_1}) - \log(\frac{P_j}{P_1}),$$

where $t^* = t/n$ and t is an arbitrary positive constant. Then

$$\sqrt{n}(e^* - e^o) = o_p(1).$$

Proof: The desired result will be proven if we can show that

$$\sqrt{n}\left[\log\left(\frac{p_j + t^*}{p_1 + t^*}\right) - \log(\frac{p_j}{p_1})\right] = o_p(1).$$

It is easy to show that the bracketed term on the left-hand side of the above expression can be rewritten as

$$\log\left(1+ \frac{2t^* + p_j}{p_j}\right) - \log\left(1 + \frac{2t^* + p_1}{p_1}\right).$$

Taylor expanding we obtain

$$\frac{2t^* + p_j}{p_j} - \left(\frac{2t^* + p_j}{p_j}\right)^2 + \left(\frac{2t^* + p_j}{p_j}\right)^3 - \ldots.$$

$$\ldots. - \frac{2t^* + p_j}{p_j} - \left(\frac{2t^* + p_j}{p_j}\right)^2 + \left(\frac{2t^* + p_j}{p_j}\right)^3 + \ldots.$$

$$= \frac{2t^*}{p_j} + 1 - \frac{2t^*}{p_1} - 1$$

$$- \frac{(2t^*)^2}{p_j^2} - \frac{4t^*}{p_j} - 1 + \frac{(2t^*)^2}{p_1^2} + \frac{4t^*}{p_1} + 1$$

$$+ \frac{(2t^*)^3}{p_j^3} + \frac{3(2t^*)^2}{p_j^2} + \frac{3(2t^*)}{p_j} + 1$$

$$- \frac{(2t^*)^3}{p_1^3} - \frac{3(2t^*)^2}{p_1^2} - \frac{3(2t^*)}{p_1} - 1 + \ldots.$$

Multiplying by $\sqrt{n}$ yields

$$\frac{(2t/\sqrt{n})}{p_j} - \frac{(2t/\sqrt{n})}{p_1} - \frac{4t^*(t/\sqrt{n})}{p_j^2} - \frac{4t/\sqrt{n}}{p_j} + \frac{4t^*(t/\sqrt{n})}{p_1^2} + \frac{4t/\sqrt{n}}{p_1}$$

$$+ \frac{8t^{*2}(t/\sqrt{n})}{p_j^3} + \frac{12t^*(t/\sqrt{n})}{p_j^2} + \frac{6(t/\sqrt{n})}{p_j}$$

$$- \frac{8t^{*2}(t/\sqrt{n})}{p_1^3} + \frac{12t^*(t/\sqrt{n})}{p_1^2} + \frac{6(t/\sqrt{n})}{p_1} + \ldots.$$

and the probability limit of this is equal to zero. Therefore $\sqrt{n}(e^* - e^o) = o_p(1)$.

LEMMA 2: Let e^o be defined as in Lemma 1. Then

$$\Omega^{-\frac{1}{2}} \sqrt{n}\ e^o \xrightarrow{d} N(0,I)$$

where

$$\Omega \equiv D\ \Sigma\ D',$$

Σ is the (50x50) matrix whose typical element is

$$\sigma_{ij} = \begin{cases} P_j(1-P_j) & \text{if } i = j \\ -P_iP_j & \text{if } i \neq j \end{cases}$$

and

D is the (49x50) matrix whose ij^{th} element is

$$
d_{ij} = \begin{cases} \frac{1}{P_j} & \text{if } i = j \\ 0 & \text{if } i \neq j \\ -\frac{1}{P_1} & \text{if } j = 1. \end{cases}
$$

Proof: First note that

$$\Sigma^{-\frac{1}{2}} \sqrt{n}\,(p - P) \xrightarrow{d} N(0, I)$$

where p and P denote the (1x50) vectors whose typical elements are p_j and P_j, respectively (see Serfling 1980, p. 109).

Then using Theorem A of Serfling (1980, p. 122) we have that

$$\Omega^{-\frac{1}{2}} \sqrt{n}\; e \xrightarrow{d} N(0, I)$$

where

$$\Omega \equiv D\, \Sigma\, D'$$

and the ij^{th} element of D is

$$
\left.\frac{\partial \log(p_i/p_1)}{\partial p_j}\right|_{P = p} = \begin{cases} \frac{1}{P_j} & \text{if } i = j \\ 0 & \text{if } i \neq j \\ -\frac{1}{P_1} & \text{if } j = 1. \end{cases}
$$

Now we can establish the asymptotic properties of the minimum chi-square estimator (8).

THEOREM 2: The minimum chi-square estimator (8) is asymptotically normal and consistent.

Proof:

We first establish that the GLS estimator

$$\tilde{\beta}_{GLS} = (Z' \; \Omega^{-1} \; Z)^{-1} \; Z'y$$

is asymptotically normal and consistent and then prove that

$$\text{plim } Z'(\hat{\Omega}^{-1} - \Omega^{-1})Z = 0$$

and

$$\text{plim } Z'(\hat{\Omega}^{-1} - \Omega^{-1}) \sqrt{n} \; e = 0$$

thus establishing the asymptotic equivalence of GLS and MCE.

$$\tilde{\beta}_{GLS} = \beta + (Z' \; \Omega^{-1} \; Z)^{-1} \; Z' \; \Omega^{-1} \; e$$

so

$$\sqrt{n}(\tilde{\beta}_{GLS} - \beta) = (Z' \; \Omega^{-1} \; Z)^{-1} \; Z' \; \Omega^{-1} \sqrt{n} \; e.$$

By Lemma 1, however, this is asymptotically equivalent to

$$(Z' \; \Omega^{-1} \; Z)^{-1} \; Z' \; \Omega^{-1} \sqrt{n} \; e^{o}$$

and by Lemma 2 $\sqrt{n} \; e^{o}$ converges in distribution to $N(0, \Omega)$ so

$$\sqrt{n} \; (\tilde{\beta}_{GLS} - \beta) \xrightarrow{d} N(0, \; (Z' \; \Omega^{-1} \; Z)^{-1}).$$

Now

$$\text{plim}(\tilde{\beta}_{GLS}) = \beta + (Z' \; \Omega^{-1} \; Z)^{-1} \; Z' \; \Omega^{-1} \; \text{plim}(e)$$

But plim e = 0 because

$$\text{plim}\left[\log\left\{\frac{p_j + (1/2n)}{p_1 + (1/2n)}\right\} - \log(\frac{P_j}{P_1})\right]$$

$$= \log\left\{\frac{\text{plim}(p_j) + \lim_{n\to\infty}(1/2n)}{\text{plim}(p_1) + \lim_{n\to\infty}(1/2n)}\right\} - \frac{P_j}{P_1}$$

$$= \frac{P_j}{P_1} - \frac{P_j}{P_1} = 0$$

Therefore

$$\text{plim}(\tilde{\beta}_{GLS}) = \beta.$$

We have thus established the consistency and asymptotic normality of $\tilde{\beta}_{GLS}$. Now we establish the asymptotic equivalence of $\tilde{\beta}$ and $\tilde{\beta}_{GLS}$. First note that

$$\text{plim}[Z'(\hat{\Omega}^{-1} - \Omega^{-1}) \sqrt{n}\ e] = Z'\ (\text{plim}(\hat{\Omega}^{-1}) - \Omega^{-1}) \cdot \text{plim}(\sqrt{n}\ e).$$

Obviously

$$\text{plim}(\hat{\Omega}^{-1}) = \Omega^{-1}$$

and by Lemmas 1 and 2 $\sqrt{n}e$ converges in distribution to $N(0, \Omega)$. Therefore, $\sqrt{ne} = 0_p(1)$ so

$$\text{plim}[Z'\ (\hat{\Omega}^{-1} - \Omega^{-1}) \sqrt{n}\ e] = 0.$$

It is also easy to show that

$$\text{plim}[Z'\ (\hat{\Omega}^{-1} - \Omega^{-1})\ Z] = 0.$$

We can therefore conclude that the MCE (8) is consistent and asymptotically normal with asymptotic covariance matrix $(Z'\ \Omega^{-1}\ Z)^{-1}$.