Locating a High-Volume Memory Device Fabrication Facility: A Global Study

S. Refik Akgun, Rebecca A. Bayard, Ginger M. Davis, Margaret W. Draughon, James A. Michelini, and Marta A. Morales

Abstract— **This project assessed the costs associated with wafer fabrication and developed a heuristic for general semiconductor site decisions. After studying several theoretical models, the Capstone team used discrete choice analysis and the multinomial conditional logit model for its evaluation. Using empirical research, the team identified a set of variables that may have influenced past location decisions. These variables were used in the model to compare the likelihood of locating in one country among a set of possible choices. The team approximated the conditional logit model with the statistical package SAS/STAT and assessed the results for statistical merit and possible conclusions. The completed analytical framework identified capital subsidies, water availability, and labor quality as the influential variables based on statistical significance. This project is a joint effort between student teams from the Darden School of Business and the School of Engineering and Applied Science at the University of Virginia.**

I. Introduction

In his testimony to the US Subcommittee on Select Revenue Measures, Craig R. Barrett, Chairman of the Board for Intel Corporation: "The cost to build a new wafer fabrication facility today is $3 billion or more. Where, and when, to build a fabrication plant is the largest ongoing decision a semiconductor CEO must make." [1] Companies routinely make site decisions; however, each decision remains unique because of variable cost constraints and stakeholder requirements. Several factors can be attributed to the eventual location decision, but the impact of specific variables over time is difficult to generalize. Therefore, the team employed a Multinomial Conditional Logit Model to compile information about past decisions into one framework.

Manuscript received April 9, 2007.

James A. Michelini is with the University of Virginia, Charlottesville, VA 22903 USA (phone: 484-459-0567; e-mail: jam2ga@virginia.edu).

S. Refik Akgun is with the University of Virginia, Charlottesville, VA 22903 USA (e-mail: sra2a@virginia.edu).

Rebecca A. Bayard is with the University of Virginia, Charlottesville, VA 22903 USA (e-mail: rab5w@virginia.edu).

Ginger M. Davis is with the University of Virginia, Charlottesville, VA 22903 USA (e-mail: gingerdavis@virginia.edu).

Margaret W. Draughon is with the University of Virginia, Charlottesville, VA 22903 USA (e-mail mwd3w@virginia.edu).

Marta A. Morales is with the University of Virginia, Charlottesville, VA 22903 USA (e-mail: marta@virginia.edu).

This paper evaluates the effectiveness of applying an engineering framework to a location decision problem in the semiconductor manufacturing industry. Such a framework cannot be structured without knowledge of industry priorities and past decision drivers. The factors that affect location decisions can be categorized into four categories: operating costs, facility needs, regional incentives and technological considerations.

Items in a cost model include utilities, machinery, land and labor costs. The operating cost of a facility is often comparable across borders: the difference between land, utility and machinery costs are often negligible. Although manufacturers often locate abroad for more affordable labor, semiconductor industry leaders claim that the difference in wage is marginal due to the skilled nature of the workforce required. Most of these cost variables are location independent.

A wafer fabrication facility (fab) needs two basic resources to succeed in a particular location. An area must be able to sustain a new facility with abundant water and electricity resources. Besides these physical resource needs, the company also requires a quality labor supply to sustain the facility. Therefore, the local level of technical education greatly impacts the success of the facility. If an area meets these needs, industry executives look for opportunities with regional incentives.

The tax package offered by a particular country often determines where the facility will be located. The increasingly popular trend of locating in Asia is generally attributed to the proximity to Original Equipment Manufacturer (OEM) producers and the attractive tax holidays and benefits offered by Asian countries. However, several companies continue to develop large scale facilities within the United States, so internal expansion remains a competitive option.

The semiconductor industry is extremely competitive. To remain successful, manufacturers must continually improve productivity and product quality. Moore's Law states that the number of components within a chip must double every 18 months. The upgrade in wafer size from 200mm to 300mm will enable manufacturers to produce approximately 2.6 times more chips per wafer constructed [2]; this increase will improve manufacturing efficiency and enable cost-effective production of high technology wafers. As a result, the decision framework will accommodate the development of a new 300mm facility for a large, international semiconductor corporation.

There is a strong international demand for these high tech facilities, each country offers a unique blend of attributes. A formal quantitative framework allows the decision maker to evaluate their options while including many relevant factors. The engineering framework compared regional attributes and offered one perspective on which variables affected past location decisions.

II. Rationale and Scope

The purpose of the project is to understand the decision making process of facility locations in the semiconductor industry. The engineering team used a quantitative model to provide an analysis of past industry locations. This insight contributed to a larger project involving a team from the Darden School at the University of Virginia. The combined team client, Tiha Von Gzchzy, a former vice president for The Boston Consulting Group asked for an innovative platform to guide semiconductor location decisions in the future. The team's quantitative model demonstrated an in-depth knowledge of the industry that helped validate the combined team's perspective on future priorities and possible competitive advantages.

The scope of the engineering project includes all memory device semiconductor facilities, as defined by an industrial inventory of global semiconductors [3]. Therefore, the scope of the quantitative model excludes the possibility of greenfield investment. The model details historic priorities but does not offer a recommendation for future location decisions. The scope of our results, in fact, may be more effectively used by governments looking to attract semiconductor investment in their countries.

III. The Conditional Logit Model

Our capstone team used a conditional logit model to perform discrete choice analysis for the semiconductor industry. Cheng and Stough [4] propose discrete choice analysis for an empirical study of Japanese location selection in China. The researchers found that discrete choice models "reveal each individual choice maker's preferences, some of which may be lost in the aggregate methodologies" [4].

Additionally, Head, Ries, and Swenson [5] used this model in *Attracting foreign manufacturing: Investment promotion and agglomeration* to analyze the locations of Japanese firms within the United States. A conditional logit model relies upon the attributes of the available choices, not the characteristics of the decision maker to perform its statistical regression [6]. In the context of our project, the explanatory variables describe potential site locations while the response variable is the location choice of past memory fabs. The fundamental assumption for a conditional logit model is that a decision maker opts to maximize their utility among the available choices [6]. For a fabrication facility location, we decided to assume that maximized utility for a semiconductor firm was equivalent to maximized profit.

The discrete-choice model accommodated the set of possible locations and statistically compared the relative attractiveness of these alternatives. Reference [4] notes that this model assumes that a rational investor i would choose a region j to construct its new facility only if this region assures profit maximization. This model assumes the forecasted profit of each region to be a function of observable characteristics of each region. The forecasted profit, π, can be calculated with

$$\pi_{ij} = \chi_{ij}\beta_i + \varepsilon_{ij}\,. \tag{1}$$

Where χ is the vector of observable location characteristics of region j, β is the vector of estimated coefficients, and ε is the error term referring to unobserved characteristics of each alternative. The error term, ε, is assumed to be independently and identically distributed (IID). As a result, a region j is selected by an investor i, if and only if it maximizes their profits according to

$$\pi_{ij} > \pi_{is}\,, \text{ for } j \neq s. \tag{2}$$

The probability of investor i choosing a region j out of s alternatives is expressed

$$\Pr ob(j) = \frac{\exp(X_{ij}\beta_i)}{\sum_{s=1}^{S}\exp(X_{is}\beta)}\,. \tag{3}$$

In (3), β is estimated with maximum likelihood method (MLE). The reader must note that the suffix i for different investors does not apply to our case since we only have one investor.

In addition, we assume independence from irrelevant alternatives (IIA) property. The IIA property dictates that the odds-ratio between two alternatives does not change by the inclusion (or exclusion) of any other alternative [7]. In the context of facility location, this requirement means that the decision maker will select the same location regardless of the other available locations. If this property is not met by our model, the parameter estimates and results likelihoods would be meaningless [8]. To meet the IIA property, all error terms must be IID.

IV. Alternative Generation

The Conditional Logit Model accommodates a set of, in our case, locations from which a semiconductor company could choose. We used an industrial inventory of global memory chip facilities to generate this list of alternatives [3]. Due to the problem statement, we originally only looked for locations that historically supported high capacity manufacturing: 30,000 monthly wafer starts or more. This gave us the following countries: China, South Korea, Japan, Singapore and the United States.

As we continued with our modeling, we realized that the conditional logit model could only accommodate n-1 degrees of freedom. To accommodate more variables we extended our analysis to include all countries with memory device fabrication facilities. Table I shows the full decision set considered in the quantitative model.

TABLE I
POTENTIAL LOCATIONS

Country	Existing High-Volume Facilities	Total Memory Facilities
United States	5	29
Japan	9	69
China	9	23
South Korea	10	27
Singapore	2	7
Russia	1	1
Ireland	0	1
United Kingdom	2	2
Germany	1	9
France	0	3
Italy	0	3

V. DATASET GENERATION

Art et al. [9] identified two hundred cost categories in his analysis of wafer factories. They conclude, however, that only twenty-three of them drive sixty per cent of the cost function. In our case, we had to account for fewer than twenty-three variables because some of those identified by Art et al. [9] (e.g., quality of wafer) were beyond the scope of this project.

Identified costs variables were then defined as location-dependent or as independent variables. Cost variables independent of location were treated as constant across the board. For example, if raw materials cost the same everywhere, this variable would be location independent and would be excluded from the location model because it would add no new information.

To identify the driving factors behind location selection, however, we had to identify location-dependent variables, which will be used as the explanatory variables of the statistical model.

A. Labor Quality

The semiconductor industry depends on the technical skills and expertise of its employers. Assuming that wages are a reflection of skills, higher wagers then represent higher productivity levels. Sometimes the technical education acquired at universities is not sufficient and companies must finance their own training. If a location does not support technical knowledge, the company would incur extra expenses trying to self-finance them and thus, reducing the attractiveness of the location. The team believed then, that the number of researchers in R&D per million people would be a good indicator for labor quality. Other variables were considered, such as high school graduates, but this did not reflect quality or technicality. Thus, a higher number of researchers in an area will reflect technical training and capability for a fab, increasing profits.

B. Water Availability

A fab requires about 2.5 million gallons of water per day [10] and a large amount of electricity. Thus, a site with scarce water resources will not be as attractive as a region with abundant water. The high demand of water for production purposes might increase costs if resources are not readily available. Thus, we decided to use total industrial freshwater withdrawals to indicate water availability. The higher the withdrawal the higher is the willingness to provide necessary water to paying industries. All the locations that were chosen are developed areas with electricity availability. Thus, electricity was not included in the model because it was considered location independent.

C. Business Climate

A good business climate indicates a professional atmosphere and may be a sign that a particular location is better for a company. The team chose the percentage of high technology exports within total manufacturing exports as an indicator of the quality of the economy. Under the assumption that technically-driven economies might welcome other technical industries, we decided that the exports were an indicator of the technological state of the economy. Additionally, we considered the highest marginal corporate tax rate. We assumed that the size of the facility would put the company in the highest bracket, and that lower rates would be preferred.

D. Infrastructure

A semiconductor fab is usually part of a complex supply and distribution network in which transportation is crucial. The team chose the ton-km of air cargo flown out of the state airports to represent infrastructure quality. Though it is not critical, it will avoid future transportation costs. We reasoned that for an airport to be built there needs to be a sufficient transportation infrastructure to access the airport, roads and shipping networks. The more an airport can ship the bigger its capacity is.

E. Tax Incentives

Tax incentives are very important because they can have a decisive impact on site selection. Since the cost of a new fab is approximately $ 3 billion, the company can expect to pay millions of dollars in taxes. To decrease these costs, companies will look for incentives from the government. Governments offer tax incentives to attract businesses to their areas because they create employment opportunities and economic stimulation. The usual form of tax incentives is either tax holidays or credits. Since this information is private, the team chose capital subsidies to represent government incentives. These were predicted using historical data from the and past aid packages [11], [12].

F. Agglomeration

According to [13] in *The Elusive Concept of Localization Economies,* similar and related industries cluster to create a localization economy. In such an economy, complementary industries thrive off each other's presence [13]. These complementary industries create a knowledge base that facilitates specialization. Increased clustering leads to increased specialization within an economy, which presumably leads to a more efficient and profitable economy. A wafer fabrication facility benefits from these economic clusters through specialized regional University research programs, existing supply and distribution channels for finished memory chips, and accessible power and water supplies. The team represented agglomeration with the number of high-volume facilities in each country.

G. Regional Indicators

In Cheng and Stough [4], regional indicator variables were used to help control the unobservable characteristics of each location. The inclusion of such variables helps the analysis meet the aforementioned IIA assumption by limiting the correlation between error terms. Therefore, these variables will be included throughout our analysis, regardless of statistical significance. We added the three regional indicators to the potential explanatory variables. Using Russia as a base case, there was an indicator for the United States, members of the European economy, and members of the Asian economy.

VI. Results

To approximate the conditional logit model within SAS, we utilized the proportional hazards procedure. This procedure modeled site selection through survival analysis where a location decision constituted an event. The dependent variable for such a procedure is the hazard rate, or likelihood of selecting a particular location. Therefore, we determined a parameter's influence according its hazard ratio. Any parameter with a hazard ratio greater than one positively impacts the likelihood of being selected [14].

A. Significant Variables

Regression analysis through the proportional hazards procedure determined the statistical significance of each potential explanatory variable when used as the sole predictor. Using a significance level of 0.05, High Technology Exports was deemed insignificant and excluded from further analysis.

Many of our variables were products of an advanced, technologically based economy. We therefore examined the correlations between significant variables for instances of strong correlation. Defining strong correlation to be greater than 0.7 or less than -0.7, both Agglomeration and the Airport Infrastructure were highly correlated with other potential explanatory variables. We then removed these two variables from the analysis. Mild correlation existed between the remainder of the variables, but these were still included for the first full regression. We then analyzed these remaining variables through the proportional hazards procedure to determine which variables significantly impacted location decisions. Table II shows the results of the regression analysis. Both Labor Cost and Corporate Tax Rate were excluded from further analysis because they were insignificant. Each of the regional indicator variables was kept to help meet the IIA assumption.

TABLE II
Initial Model Results

Variable	Parameter Estimate	Pr > Chi Squared
Research & Development	0.0007063	0.0063
Water Availability	0.01074	0.0036
Labor Cost	6.71949E-6	0.7883
Corporate Tax Rate	0.02105	0.5814
Capital Subsidy	0.00957	<.0001
Region 1	-0.98876	0.4332
Region 2	-0.95732	0.4806
Region 3	0.36540	0.7783

This result left us with the second model specification containing Research & Development, Water Availability, and Capital Subsidy as the significant influences on location decisions. Surprisingly, our model did not deem agglomeration to be a significant influence on past site locations. We suspect that it could be a lurking variable because of its strong correlation to Capital Subsidy and Water Availability. We ran another proportional hazards procedures in SAS to determine precise parameter estimates and hazard ratios of the remaining variables. These results are summarized in Table III. Not surprisingly, each quantifiable variable had a positive influence on the likelihood of location. This specification passed three separate test of model utility. The Likelihood Ratio, Score, and Wald tests all rejected the global null hypothesis that the parameters are equal to zero. The p-value for each test was less than 0.0001, a solid rejection. This result means that our model does reveal the preferences of decision makers and can accurately project the likelihood of choosing a particular site.

TABLE III
Final Model Results

Variable	Parameter Estimate	Hazard Ratio	Pr > Chi Squared
Research & Development	0.0008162	1.001	<.0001
Water Availability	0.01241	1.012	<.0001
Capital Subsidy	0.01041	1.010	<.0001
Region 1	-1.09855	0.333	0.3776
Region 2	-1.14116	0.319	0.3771
Region 3	0.57853	1.783	0.5850

B. Likelihood of Location

Once we determined the final model parameters, we used (3) to determine the likelihood of selecting each particular from the choice set. Table IV summarizes our findings. The predicted likelihoods are very similar to the expected likelihoods: the historic likelihoods match those determined by variable parameters. The similarity between expected and

TABLE IV
Likelihood of Location

Country	Expected Likelihood	Predicted Likelihood
United States	16.667%	16.489%
Japan	39.655%	38.557%
China	13.218%	13.112%
South Korea	15.517%	16.753%
Singapore	4.023%	4.294%
Russia	0.575%	0.568%
Ireland	0.575%	0.554%
United Kingdom	1.149%	1.395%
Germany	5.172%	5.728%
France	1.724%	1.364%
Italy	1.724%	1.186%

predicted likelihoods supports the findings of our analysis. In a model that predicts a future decision, the countries that offer the best labor quality, resource availability, and tax incentives should attract the most facilities. Therefore, a high-volume facility was most likely to be located in Japan, South Korea, the United States or China.

VII. Conclusion

Discrete choice analysis was an appropriate method for seeking to understand the critical factors in historical site selections. This method gave us quantifiable data that was useful when proving our understanding. The biggest strength of such a modeling technique was its ability to discern the truly relevant preferences of decision makers. While the model did return likelihoods for locating a facility in a particular country, these should not be used to make a site selection for a high-volume facility. There are too many ambiguous factors associated with each particular location problem that can not be quantified. Potentially, this analysis might be useful for location decisions when considering less specific industries. A larger sample size and more general business requirements would alleviate concerns over ambiguous characteristics. We recommend discrete choice analysis for those seeking to understand the driving factors behind an industry. In particular, we recommend it for regional policymakers looking to attract semiconductor investment. By creating policy centered on the critical factors we discovered, a country will inevitably attract new wafer fabrication facilities.

Appendix

Table V shows the variables incorporated into the model and the sources associated with their values.

TABLE V
DATA SOURCES

Variable Name		Source
Research & Development	Number of researchers in R&D per million people	World Bank [15]
Water Availability	Total water withdrawals * % withdrawals for industry	World Bank [15]
Labor Cost	Labor costs per worker in manufacturing	World Bank [15]
Corporate Tax Rate	Highest marginal corporate tax rate	World Bank [15]
Airport Infrastructure	Air freight (million-ton-km)	World Bank [15]
Agglomeration	Number of existing high volume memory wafer fabs	Semizone [3]
High Technology Exports	High technology exports as a percent of manufacturing	World Bank [15]
Capital Subsidy	Value of capital tax credits in millions of US dollars	Trade Magazines, [11],[12] Chambers of Commerce

Acknowledgment

The authors wish to thank the team of five Darden Graduate School of Business students and their advisor, Tiha Von Ghyczy. Without their industrial expertise and insight into site selection, this project would never have been successful.

References

[1] United States Congressional Subcommittee on Select Revenue Measures. *Statement of Craig R. Barrett, Ph. D., Chairman of the Board, Intel Corporation.* Hearing 22 June, 2006. Washington: Government Printing Office, 2006

[2] United Microelectronics Corporation. (2005). *World Class Manufacturing.* Retrieved March 16, 2007, 2007, from http://www.umc.com/English/class_300/images/300mm(2003).gif.

[3] "SemiZone: Fabs." *Stanford Center for Professional Development.* http://www.semizone.com/fab/tech-list.tcl?id=5&st=1. 2007.

[4] Cheng, Shaoming; Stough, Roger. "Location Decisions of Japanese New Manufacturing Plants in China: A Discrete-Choice Analysis". Springer-Verlag 2006, March 14. pp: 1- 19.

[5] Head, K.C., J.C. Ries, and D.L. Swenson. 1999. Attracting Foreign Manufacturing: Investment Promotion and Agglomeration. *Regional Science and Urban Economics* 29: 197 – 218.

[6] Greene, W. H. (1997). Chapter 19: Models with discrete dependent variables. *Econometric Analysis* (5th ed., pp. 871) Prentice Hall.

[7] Dahlberg, Matz; Eklog, Matias. Relaxing the IIA Assumption in Locational Choice Models: A comparison between conditional logit, mixed logit, and multinomial probit modrels. Department of Economics, Uppsala University, 2002.

[8] Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, The Massachusetts Institute of Technology.

[9] Art, David, O'Halloran, Michael, Butler, Brian. Wafer Fab Contruction Cost Analysis & Cost Reduction Strategies: Applications of SEMATECH's Future Factory Analysis Methodology. IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 1994

[10] Peters, L. (1998, February 1). Ultrapure water: Rewards of recycling. [Electronic version]. *Semiconductor International*. Retrieved March 19, 2007, from Reed-Electronics database.

[11] R&D Credit Coalition. International R&D tax incentives. R&D Credit Coalition. Retrieved March 14, 2007. (2006)

[12] Moor, J. (2004). *The UK semiconductor industry: How can a trade association support a new strategy for UK competitiveness?* National Microelectronics Institute. Retrieved March 17, 2007, from NMI database.

[13] Malmberg, A., & Maskell, P. (2001). The elusive concept of localization economies– towards a knowledge-based theory of spatial clustering. Paper presented at the *AAG Annual Conference: 'Industrial Clusters' Revisited: Innovative Places Or Uncharted Spaces? Session,* New York.

[14] Machin, D., Cheung, Y. B., & Parmar, M. K. B. (2006). Cox's proportional hazards model. *Survival analysis: A practical approach* (2nd ed., pp. 121-154). Hoboken, NJ: Wiley.

[15] World Bank: Countries and Regions. *The World BankGroup.*http://go.worldbank.org/FFZ0CTE2V0. 2007.