CEDE

DOCUMENTO CEDE 2003-10
ISSN 1657-7191 (Edición Electrónica)
ABRIL DE 2003

A DYNAMIC MODEL OF VERTICAL INTEGRATION FOR THE AMERICAN PULP AND PAPER INDUSTRY[*]

MARCELA MELÉNDEZ[**]

Abstract

The focus of this research is to learn about the factors that influence the decision of a manufacturing firm to vertically integrate into the production of its input. The American paper industry has a feature that makes it particularly suitable for this purpose: over the years paper mills of apparently similar characteristics have made different decisions with regards to their integration status. This work draws on the insight that there must be some unobserved mill characteristic that drives the decision process for a mill. Mills' choices of whether to exit the industry, and with regards to their integration status when they choose to stay in operation, depend on their productivity. This generates selection and simultaneity biases in a reduced form estimation. In order to deal with these issues, I propose a dynamic model in the spirit of Olley and Pakes (1996). This approach not only takes care of the estimation biases, but also allows me to learn about the unobserved characteristics of the firms in my data, and to use them to determine which firms vertically integrate and which firms do not. In addition, the model I propose allows me to learn about how vertical integration affects productivity and mill's entry and exit decisions.

JEL Classification: L22, C14, C61, C63

Key Words: Vertical Integration, Statistical Simulation Methods – Monte Carlo Methods, Dynamic Analysis, Computational Techniques

---

[*] This is the complete text of my Yale University Ph.D. thesis, presented in the format of a CEDE Working Paper. Advisors: Chris Timmins, Steve Berry and Pat Bayer.

[**] Assistant Professor, Department of Economics, Universidad de los Andes, Calle 18A No. 1-37 Este, Bogotá, Colombia. Email: mmelende@uniandes.edu.co

Chapter 1

Introduction and Summary

The primary focus of this research is to learn about the factors that influence the decision of a manufacturing firm to vertically integrate into input production. In an attempt to shed some light on the empirically observable regularities and motives that shape a manufacturing firm's decision to internalize the production of its input, I propose a dynamic model á la Olley and Pakes (1996) that accounts for unobserved firm-specific characteristics and for entry and exit. This approach takes care of the estimation biases from selection and simultaneity that arise in a reduced form model framework, and allows me to learn about the unobserved characteristics of the firms in my data, and use them to determine who vertically integrates and who does not. Furthermore, the model I propose allows me to learn about how vertical integration affects productivity and firm’s entry and exit decisions.

I use the U.S. pulp and paper industry in my research. I have collected data on individual mills for the years 1975, 1980, 1985, 1990 and 1995. The result is a panel dataset that allows me to track entry and exit, as well as integration status. Two features in particular make this industry suitable for my purposes. First, over the period 1975 to 1995 less than half of the operating paper mills in the U.S. were integrated backwards into pulp production. These mills accounted, however, for about 80% of the country's paper capacity. It seems a perfect setting to try to sort out the different reasons for a mill to vertically integrate. Second, the industry displays high rates of entry and exit over the period. While there were on average 550 active paper mills each year, only 361 of the 584 mills operating in 1975 survived to 1995. This behavior of the mills in the industry raises a reasonable concern for selection, and thus provides also an interesting setting to learn about the role of unobserved productivity and about the most convenient methodological approaches to deal with it in empirical estimation.

This dissertation is structured as follows. Chapter 2 presents my data and gives an overview of the American pulp and paper industry for the period 1975-1995. Chapter 3 presents a reduced form model of vertical integration for the paper industry and raises concerns about the possible estimation biases that may arise in a model that ignores the role of unobserved firm-specific productivity in a firm’s decision process. Chapter 4 presents a structural model of firm behavior that captures the choices facing a paper mill each period, with regards to both its operation status (in or out) and its vertical integration status (integrated or disintegrated). Chapter 5 presents a simulation exercise to show how the structural model from the previous chapter allows me to deal with the endogeneity and selection biases possibly present in a reduced form estimation framework. Chapter 6 describes in detail the steps required for the estimation of the structural model proposed, and presents some baseline results. Chapter 7 considers the possible directions in which this research may be extended. Chapter 8 concludes.

Chapter 2

Data and Industry Background

2.1 The papermaking process

Paper is basically comprised of a mat of fibers. Since its invention, many different fibers have been used for its manufacture. These include the bast fibers of flax and mulberry, the stalks of bamboo and other grasses, various leaf fibers, cottonseed hair, wool, asbestos, and the woody fibers of trees. Cotton and linen rags, straw, and wastepaper have also been used. Today, however, wood is the primary raw material from which most paper pulp is made.

Before wood can be made into paper, paperboard, and other products, it must be reduced to its basic components to form pulp. Wood is made up primarily of cellulose fibers bound together with lignin, a glue-like binder, plus sugars, gums, resins and mineral salts in lesser quantities. The objective of pulp manufacturing is to separate the wood into fibers and other components, remove the undesirable components, and provide a means for treating the fiber to produce a suitable pulp for the paper mill. Depending on the type of wood used, and the requirements of the end product, the conversion of wood into wood fibers is carried out by one (or a combination) of three general types of processes. The simplest pulping method is generally referred to as *mechanical pulping*. It differs from other pulping methods in that the reduction of pulp to fibers is essentially a physical operation in which the fibers are actually pulled away from each other by the application of some type of mechanical force. In *chemical pulping* processes, the chips are cooked in a chemical solution under pressure until the fibers are separated by dissolution of their lignin binder and fall apart with very little or no mechanical action. There are several chemical-pulping processes using one of several chemical solutions, depending on the type of wood used and the type of paper desired. The most known are the soda, sulfite, and sulfate processes. Finally, *semi-chemical pulping* processes combine mechanical and chemical methods. These processes were developed particularly for the pulping of hardwoods and have many variations.

Pulp is made into paper in a paper mill. Paper mill layouts are generally divided into stock preparation, paper machine, and finishing operations. They vary somewhat based on the pulps used and grades of paper produced.

Most pulp cannot be used for papermaking as it comes directly from the pulp mill. To obtain the final desired qualities, pulp having different characteristics may be blended in, and dyes and special additives are also included to achieve the specific color and physical properties of the sheet. These operations are usually referred to as stock proportioning or stock blending. To impart mechanical strength to the final sheet, the pulp is then refined in a variety of machines, typical of which are the refiners, jordans, and beaters. Basically, this operation consists of passing the pulp repeatedly between sharp moving bars that cut and abrade the fibers. Before going to the paper machine the pulp is screened and, in some cases, cleaned by passing through centrifugal-type cleaners.

Paper machine operations can be subdivided into wet and dry end. The most common major component of the wet end portion of the paper machine on which the paper is formed is a foudrinier, consisting mainly of a continuous fine screen, called a wire, on which the pulp suspension is spread. Most of the water drains at the top of the wire to form a mat of fibers. A series of vacuum suction boxes draw more water from the mat, and the wet paper leaves the foudrinier machine at a consistency of about 20% fiber and additives, 80% water. The wet paper is sent to the presses where it is supported by endless woven or synthetic loops called felts, and passed between heavy press rolls that press out as much water as possible. The rest of the water is then evaporated on steam-heated rolls.

Most dried papers go through one or more additional finishing processes, one of which is calendering. This process consists of ironing the paper between heavy polished steel rollers to give it a much smoother surface. Some paper is wound in large rolls as it comes from the calenders, to be later rewound and cut into smaller rolls or sheets as required by the user. And some papers are produced specifically for further processing by converter-plants that make consumer products.

The different paper products may be classified using six broad paper categories: (1) newsprint and other papers produced from groundwood pulp; (2) book and writing papers; (3) wrapping papers; (4) paperboard; (5) sanitary papers; and (6) industrial papers.

2.2 The U.S. pulp and paper industry, 1975-1995

Data was collected for five cross-sections at five-year intervals, starting in1975, from the Lockwood-Post's Directory of the Pulp, Paper and Allied Trades.[1] For all the active mills in the U.S. for which pulp and paper production capacity was reported,[2] information on owner, location, capacities, and paper grades produced each year was recorded. Mills were coded according to their owner and location in order to be able to track them over time. Table 2.1 presents some summary statistics. It shows a growing industry that displays a tendency towards larger and fewer mills over the years.

The mill level data on pulp and paper production capacities were used to determine its degree of vertical integration. For each mill, the pulp to paper ratio was computed, and the mill was denoted as vertically integrated whenever this ratio was different from zero. In addition, the fact that roughly 30% of the firms in the industry own more than one mill suggested that a definition of vertical integration accounting only for production of both pulp and paper at the same location was too restrictive. Conversations with people in the

---

1 Formerly known as the "Post's Pulp and Paper Directory", it changed its name after a merger with the "Lockwood 's Directory of the Paper and Allied Trades" in 1987.

2 Capacity reported was for 605 out of 755 mills in 1975, 620 out of 738 in 1980, 589 out of 671 in 1985, 536 out of 629 in 1990, and 539 out of 638 in 1995.

trade revealed that a firm is likely to use the pulp produced at one location for the production of paper at another as long as the two locations are not too far apart. This lead to the conclusion that, for mills belonging to multi-plant firms, the distance between the individual locations needed to be incorporated into the definition of vertical integration. According to industry sources, the largest distance between a paper and a pulp mill that allows for cost-efficient transportation of the pulp to the paper mill lies between 350 to 400 miles. With this consideration in mind, the distance between integrated and specialized paper mills, as well as the distance between pulp mills and specialized paper mills owned by the same firm, was measured. Whenever such distance fell within the cost efficient range, production capacities were considered in order to determine whether the paper mill should be considered integrated.

2.2.1 Integration Patterns

Tables 2.2 through 2.6 show different aspects of vertical integration in the paper industry. From table 2.2 we see that while most of the pulp-producing mills are vertically integrated into paper production, this is not true for paper mills. In 1980 mills reporting in-site production of both pulp and paper accounted for 47% of all paper producing mills. The proportion is lower for all other years in the sample. Interestingly, however, roughly 80% of total paper capacity in the U.S. each year belongs to mills reporting capacities for both pulp and paper production.

Table 2.2 also reports the number of integrated mills that obtains when paper mills using pulp from a close-by mill belonging to the same owner are accounted for. We see that paper mills with in-site pulp production represent between an 86% and a 92% of all integrated mills over the years.

Table 2.3 looks at the proportion of vertically integrated paper mills in the U.S. by region. The distribution of integrated mills appears rather unequal when looked at in this manner, with a significantly higher proportion of integrated mills in the southern and western regions, where most of the forestland in the U.S. is found, and a significantly lower proportion in the Northeast.

Table 2.4 presents the distribution of integrated paper mills according to their pulp to paper capacities ratio. Two issues are noteworthy here. First, even though throughout the sample period, integrated firms with pulp to paper ratios greater than or equal to 1 are almost always in majority[3], over the years the proportion of firms in this range fell from 61% of all integrated mills in 1975 to 45% in 1995. At the same time, the proportion of firms in the two lower ranges was rising, particularly in the middle range of ratios between 0.5 and 1. This shows an increasing tendency of the vertically integrated firms to be self-sufficient, but not participate as sellers in the market-pulp market. Second, both the mean and the standard deviation of the pulp to paper ratios fell over the years.

[3] The exception is year 1990 for which the 0.5 to 1 range displays the higher proportion.

However, while the mean fell by only 8% from 1975 to 1995, the standard deviation did so by 35%. Very low and very high levels of vertical integration, as measured by this ratio, must have been found to be inefficient over the years.

Table 2.5 compares integrated versus nonintegrated paper mills over the sample years. We see that mills that produce newsprint paper have been mostly integrated. In contrast very few of the firms that produce industrial paper grades are vertically integrated. These two findings agree with the predictions of the transaction-costs model: it is the more standardized paper grades that should be produced by vertically integrated productive units. We also find that the proportion of integrated mills belonging to multi-plant firms is higher than the proportion of specialized paper mills, and that integrated mills are more likely to produce more than one paper type than nonintegrated mills.

Table 2.6 presents the distribution of paper capacity by relative size of mill capacity, distinguishing integrated and nonintegrated mills. We see that the vertically integrated paper mills are generally larger than the nonintegrated. This table is also intended to permit the exercise of applying the "survivor technique" developed by Stigler (1968) to find the optimal plant size of both integrated and nonintegrated mills[4] Market share has risen for the vertically integrated mills with production capacities above 0.5% and below 0.75% of total paper capacity. It has also risen for those with capacities above 1% and below 1.5% of total capacity. For the nonintegrated paper mills there is no size category with an increasing market share over the years. Most nonintegrated mills, however, have paper capacities below 0.1% of total capacity. This size category holds a basically steady market share over the sample period. Also, table 2.6 shows declining market shares for the nonintegrated mills with sizes smaller than 0.5% of total paper capacity. There appears to be a production scale at which vertical integration becomes efficient for a paper mill. The efficient mill size for vertical integration falls in the range above 0.5% of total paper capacity in the industry.

2.2.2 Entry and Exit Patterns

Table 2.7 presents the rates of entry and exit in the U.S. paper industry over the sample years. The sample is divided into four periods, and an entrant is defined as any mill whose code is not in the database in the starting year of the period. Similarly, an exiting mill is defined as one whose code is not in the database in the last year of the period. Although the total number of paper mills fell from 584 in 1975 to 508 in 1995, this result obtains after substantial simultaneous entry and exit of mills, the ending balance being determined to a great extent by the high amount of exit observed during the 1985-90 period. Table 2.7 also shows that in each period, the share of exiting nonintegrated mills exceeded that of exiting integrated mills. It is also true, however, that for two out of the four periods in the

[4] This method finds the efficient mill size by assuming that competition of different sizes of mills sifts out the more efficient enterprises. Mills are classified by size category based on percentages of total industry production (or capacity) and the market share originating from each category is observed over time. Efficient mill size will be revealed by a rising market share and inefficient size(s) by declining market shares over the period.

data the number of nonintegrated paper mills entering the market is higher than the number of integrated mills.

Table 2.8 compares integrated entrants and integrated established mills each period, at the national level and by region. Integrated mills constitute a decreasing share of all entrants to the industry. Nevertheless, they account for a rather stable share of all established firms. At the regional level, the southern and western regions display higher proportions of integrated mills than the other regions, and this is true for both entrants and established firms.[5] During the 1990-95 period, however, the proportion of integrated entrants in the South is substantially below that of the earlier periods.

Table 2.9 focuses on the paper mills that survived the whole sample period without changing integration status. They amount in number to 290 mills. 65% of the vertically integrated mills that were active in 1975 survived to 1995, while only 40% of the nonintegrated paper mills did. So the survival rate was higher among the integrated mills each year, suggesting that integration was a profitable strategy contributing to enhance competitiveness.

Table 2.10 shows the mills that switched between integration and specialization each period. These constitute a minority of all mills. In general, a mill that starts out as vertically integrated ends the sample period as such, and the same holds for nonintegrated firms. The switching mills reported are paper mills (they switch from paper only to integration or from integration to paper only, but not into or out of paper production[6]) and several of them appear as switching from one status to the other more than once during the sample.

Finally, table 2.11 presents the Herfindahl indices for the regional pulp and paper markets. Regional differences are apparent for the pulp market, the South showing much higher concentration measures than the other regions. Concentration in the South, in addition, shows an increasing trend over time. The western region has the second highest concentration measure, although much less than the South, and displaying the opposite trend over the sample period. For the paper market we observe lower indices in general, with higher indices in the South than in the other three regions. It is worth noting that while there is considerable variation over time of the concentration measures for the pulp market, there is never as much for the paper market.

---

5 Only in the 1975-1980 period the proportion of integrated entrants in the Central region is higher than that of the West.

6 In my sample only one pulp mill integrated forward into paper production and one integrated mill disintegrated into a pulp only mill. Neither is the subject of my study.

Chapter 3

Reduced Form Model of Vertical Integration

3.1 Related Literature

3.1.1 Incentives for Vertical Integration

Central to the notion of vertical integration is the elimination of contractual exchanges and the substitution of internal exchanges within the boundaries of the firm. Perry (1989) classifies the incentives for vertical integration in three broad categories: technological economies, transactional economies, and market imperfections. Technological economies occur if fewer resources are used to produce the downstream output upon vertical integration with the upstream process. Transactional economies refer to the reduction in transaction costs when integration replaces external exchange. Finally, "market economies" stem from substituting integration for exchanges in imperfect markets: firms integrate to avoid costs attributed to monopoly or monopsony market power, price controls and rationing, externalities, imperfect or asymmetric information, and problems associated with uncertainty.

The first two categories of determinants view vertical integration as welfare increasing. As a result, transaction cost economics is primarily concerned with explaining and predicting patterns of vertical integration (Teece 1976, Levin 1991, Levy 1985, Spiller 1985). In contrast, vertical integration that arises in response to market imperfections may increase or decrease welfare. So public policy questions become of central interest. When firms make the decision to vertically integrate, it is likely that more than one of these categories apply. However, it is hard to identify the different types of incentives econometrically. Market imperfections usually translate into costs of exchange, and since there are often other possible sources to these costs, the econometrician is not always capable of identifying them separately.

3.1.2 The Transaction Costs Model

Like production, exchange is costly. Vertical integration is just one method of bilateral exchange. Contractual exchange is its primary alternative. Transaction cost economics examines the relative costs of contractual versus internal exchange. Its leading proponent is Williamson (1975,1985).

Bilateral monopoly between a buyer and a seller arises because gains from trade are enhanced by investments in assets specialized to their exchange. This is what Williamson called "asset specificity". Asset specificity may arise from investments in specific physical capital, specific human capital, site-specific capital, dedicated capital, or brand name capital. These transaction specific assets are the source of what Klein, Crawford and Alchian (1978) call "appropriable quasi-rents" - the difference between the value of the asset in its current use and its value in its next best use. In uncertain environments, it is usually prohibitively costly to write long-term contracts specifying all obligations under all contingencies. It is likely, then, that a situation arises in which one party engages in opportunistic behavior, and attempts to extract the quasi-rents of the other by threatening to dissolve the relationship unless price concessions are granted. Such opportunistic behavior involves costs of haggling, and may result in the failure to maximize joint profits. Provisions in the contracts that govern the exchange can lessen some of these problems. However, when asset specificity is substantial, contractual governance over opportunism can be very costly, and internal organization of the exchange through vertical integration may then be more efficient.

Empirical assessments of the transaction-cost explanation for vertical integration have found that a number of transaction-cost indicators, such as industry concentration (an indicator of the number of buyers and/or sellers competing in the market) and proxies for asset specificity, are positively associated with the extent of vertical integration in a given industry or firm. Spiller (1985) examines vertical mergers to test the theory. He estimates the gains from mergers using stock price information, and finds that they are negatively related to the distance between vertically related plants, a measure of transaction-specific assets. This provides some support for site specificity and vertical coordination as explanations of vertical mergers. Globerman and Schwindt (1986) evaluate the significance of transactional determinants in explaining the observed patterns of vertical integration in the Canadian forest products sector. They document and explain the integration patterns of the 30 largest forest-based enterprises using the transactions cost approach, and conclude that transactional considerations, particularly asset specificity, are robust empirical determinants of the governance structure. Hennart (1988) explores whether Williamson's theory of vertical integration holds for the upstream stages of the aluminum and tin industries. He concludes that the structure of these industries is broadly consistent with Williamson's predictions: the higher degree of vertical integration in aluminum can be explained by greater scale economies, higher barriers to entry, higher transportation costs, and greater asset specificity. Differences in vertical integration within the tin industry can be explained in terms of the same variables. Caves and Bradburd (1988) test the theoretical models of vertical integration on a cross-section of industries. They confirm the roles of contractual and transaction-cost factors such as small-numbers bargaining, lock-in effects, and the need for industries to share intangible assets. Finally, Ohanian (1993, 1994) tests a transaction-cost model of vertical integration of the U.S pulp and paper industry for the period 1900-1940. Using mill level data, she finds vertical integration of pulp and paper production to be positively associated with regional concentration, paper mill capacity, and production of standardized grades of paper.

3.2 The reduced-form model

In order to assess the extent to which transaction-costs explain the choice of a paper mill to integrate backward into pulp production, the following model is estimated:

$$VI_i = \alpha_0 + \alpha_1 SIZE_i + \alpha_2 NEWS_i + \alpha_3 SW_i + \alpha_4 CONC_i + \varepsilon_i$$

VI_i is a dummy variable denoting whether mill i is or is not vertically integrated. Recall that about 70% of the paper mills are themselves firms (i.e. are owned by a single-mill firm). Also recall that for the remaining 30% that belong to multi-plant firms, distance was incorporated into the definition of vertical integration in order to account for the possible integration that occurs between mills at different close-by locations when owned by the same firm. This implies that in this model focusing on the mill as the relevant decision-making unit, is equivalent to focusing on the firm. From now on both terms will be used interchangeably. As explained in Chapter 2, VI_i takes a value of 1 whenever the computed pulp to paper ratio is different from zero, and a value of 0 otherwise. $SIZE_i$ is the log of the paper capacity of firm i, measured in tons per day. It is a proxy for firm size, because data such as firm sales or assets are not available for the full sample of firms listed in the directory. $NEWS_i$ is a dummy variable indicating the production of newsprint, a standardized paper grade, that is usually produced on a larger scale than other specialized finer grades of paper which are produced in smaller batches (to order), and requires a less complex mix of pulps than the finer grades. It is a proxy for asset specificity, because newsprint production requires the specialization of the pulping assets to conform to the requirements of the papermaker. SW_i is a dummy variable that takes the value of 1 whenever the firm is located in the South or West regions, where most of the forestland in the U.S. is found. It is a proxy for site specificity. Finally $CONC_i$ is a concentration measure of the market in which the mill operates, and it is used as a proxy for the potential for a small-numbers bargaining problem. A circular market with a 100 miles radius was defined around each mill and two types of concentration measures were constructed. The first one is a Herfindahl-Hirshman Index for the market-pulp market, calculated from the market-pulp capacities that result after accounting for the pulp capacity that is absorbed by the paper mills belonging to the same owner and operating in close-by locations[7]. The pulp capacity share of each mill was excluded when building the index for the circular market in which it operates, so these concentration measures are mill-specific. The second measure used as a proxy for concentration is the pulp by mill ratio in each circular market.

Firm size is expected to be positively associated with the decision to vertically integrate, because, all else equal, transaction-cost savings will be greater compared to those of a smaller firm. Temin (1988) observes that larger firms are more likely to assure a steady supply of inputs through integration because the costs of supply disruptions are greater for them. Larger firms are also more likely to integrate if economies of scale in the upstream process result in lower costs for their own production compared to small firms (Williamson 1985). Also, if the frequency of transactions rises with firm size, a greater frequency of transactions will increase the benefits to integrate and may justify the costs of reorganization (Williamson (1985)). The coefficients on the dummy variables are expected to be positive, since they proxy for different forms of asset specificity. Finally, concentration is also expected to be positively associated with the decision to vertically integrate. Generally, market concentration indicates the number of alternative suppliers (or buyers) a firm may turn to in the event of opportunistic behavior by another party. A highly concentrated market structure

[7] Market-pulp capacity is defined as the pulp capacity reported by mills producing only pulp, plus the difference between the pulp and paper capacities reported by those integrated mills with a pulp to paper ratio higher than 1.

has a greater potential for opportunistic behavior in market transactions and, hence, constitutes an incentive to vertically integrate.

The model is also estimated after pooling all the data years. A pooled regression and a random effects panel regression are run using the variables described above. The pulp production of the 100 miles radius circular market is considered as an alternative concentration measure. I also estimate a version of the panel including two new market characteristic variables. State-level income per capita is included as a measure of the size of the market in which the mill operates. I also include a measure of the proportion of ruralness of the state; I use the ratio of state-level non-metropolitan to metropolitan populations.

Finally, a version of the model is estimated pooling all the data years and taking differences. In this reduced form specification I include both current and lagged concentration measures, and time trends to allow for the possibility that industry-level shocks at certain moments in time affected the decision of the individual mills to become integrated or to disintegrate.

Empirical Results

The data set contains observations at five-year intervals between 1975 and 1995, for the U.S. paper mills that were active each year, and for which pulp and paper capacity was reported. Since there was entry and exit over the period, the number of mills is not constant across cross-sections.

The discrete dependent variable model

I begin by estimating the first model described in Section IV, including the same set of independent variables and a constant term for each cross-section, as well as for the pooled data. Estimation is done using alternative discrete dependent variable models to ensure that the results are not dependent on a possibly erroneous assumption about the underlying distribution. The linear probability model, the logit model, and the probit model produce similar answers. Estimation results for the probit model, as well as the corresponding partial derivatives, are presented in table 3.1.

The coefficient on size is always positive as expected and highly significant, implying that there is indeed a bigger incentive for larger firms to become vertically integrated. This strong result, however, may also be signaling the possible endogeneity of size. It may be that there is some unobserved firm characteristic causing mills to be both large and vertically integrated. This possibility that mill size is endogenously determined needs to be addressed.

The coefficient on the dummy variable indicating newsprint production, $NEWS_i$, is always positive and highly significant for all years but 1975, when it is still positive but insignificant. 1975 is also the year in the sample with proportionally fewer vertically integrated newsprint firms (see table 2.5). The coefficient on the dummy variable serving as a proxy for forestland, SW_i, is positive and significant for all cross-sections, with the exception of 1995, when the

Herfindahl-Hirshman concentration measure is used (model I in table 3.1). It would seem that investment near forestland area lost its specific asset condition for the later years in the sample. A possible explanation for this could be the growing tendency in the paper industry to substitute towards more environmentally friendly pulping processes that do not rely only on the availability of wood pulp. When the pulp-by-mill concentration measure is used (model II in table 3.1), this variable loses significance. It remains significant for the first two years in the sample, and the pattern of significance over time is not lost.

The surprising result from this exercise is that when the Herfindahl-Hirshman concentration measure is used, the coefficient is always negative and significant. This result is in clear contradiction to the one obtained by Ohanian (1993,1994) for the earlier years of the paper industry. A possible explanation may be in the fact that during the first four decades of the century the paper industry in the U.S was going through a relocation process from the Northern and Central regions towards the South and West. Investments close to the primary input sources were apparently a strong incentive for start-up firms to exploit the benefits of vertical integration. At the same time, the number of firms in the South and West was not yet as large as in the later years. So the concentration measures were higher in the regions where firms were finding it more efficient to integrate backward into pulp production. Once the industry settled, the differences in market concentration across regions became smaller (see the regional Herfindahl indices in table 2.11). The number of vertically integrated paper mills in the South and West remained, however, proportionally higher than in the North or Center of the U.S. This result could be signaling once again the presence of endogeneity in the model. If the higher the number of vertically integrated mills in the market in response to some unobserved firm characteristic, the higher the market concentration, then market concentration is endogenously determined. The correlation of the unobserved factors determining the vertical integration decision with those explaining the market’s concentration level would be causing my estimates to be biased and could be counteracting a true positive effect.

The pulp-by-mill concentration measure yields a different result. The probit coefficient on this concentration proxy is slightly positive for all years in the sample and always strongly significant. The fact that an alternative concentration measure yields a result more in line with what is expected, contributes to strengthen the hypothesis that the negative result of the first model may be due to endogeneity of market concentration to vertical integration. Certainly this hypothesis needs to be further explored.

Panel data analysis

By pooling together all of the cross-sections I am able to build an unbalanced panel with 2593 observations. I begin by estimating the probit model in which the dependent variable is the vertical integration dummy, as a conventional panel with common slope parameters imposed across all firms. The same probit model is then estimated under the assumption of random coefficients. Estimation results are presented in table 3.2. The random-effects probit returns larger coefficients and slightly lower t-statistics, but in general results in both cases are similar. Consistent with the predictions of the transaction-costs model, the coefficients obtained for the firm size variable, the newsprint production dummy, and the forestland dummy are in both cases positive and strongly significant. As stated above, three alternative concentration measures are used, and included both in their current and lagged versions. The first two columns of table 3.2 show the results when the concentration measure employed is the Herfindahl-Hirshman Index. As before, the coefficient on the current period concentration index is negative and significant. However, interestingly, the coefficient on its lag is positive as expected, and strongly significant. This result says that mills respond with a lag to market concentration in terms of making decisions about their integration status. This is by no means a counter-intuitive result, since a decision to integrate requires investing in the purchase of pulp-processing machinery and equipment. This result obtains for all alternative proxies of market concentration used. Concentration measures labeled II and III refer respectively to pulp-per-mill and total pulp in the 100-miles circular market. These two proxies for concentration are likely to exhibit less endogeneity related to the decision of a mill to vertically integrate. The coefficients on the current period concentration are in these two cases small but positive, and strongly significant.

The fixed-effects versions of these models were estimated switching to a logit specification and are available at request. Estimation in this case presents some inconveniences. First, the two explanatory dummy variables are lost. The forestland dummy is by definition invariant within each group since firms are coded over time according, among other features, to their location. With a few exceptions, the newsprint production dummy does not vary within groups either. Because it is unlikely that firms will decide to switch in or out of newsprint production given the specialized machinery and the scale of operations that production of this paper type requires, a good part of this variation can be attributed to measurement error. It thus seems convenient to drop this variable from the fixed-effects panel estimation. Second, fixed-effects estimation takes into account only those mills whose vertical integration status changed over the period. All other mills are automatically dropped during the estimation process and that takes us from a sample of 643 mills (2593 observations) to one of 142 mills (612 observations). The estimation results from the fixed-effects version of this model need to be considered with caution. The explanatory variables included lose significance, probably for the reasons presented above, and the fixed effects estimation overall contributes no additional information to this research.

Table 3.3 presents the results of the panel estimation using change data. The purpose of this exercise is to further explore the relationship between integration and market concentration, by trying to explain in which way market concentration affects the decision of a paper mill to change integration status. For this purpose both a change in concentration and a lagged concentration measure are included as explanatory variables. The concentration measure used in this exercise was the Herfindahl-Hirshman index corrected, as described above, to exclude the mills' own market share in the 100-miles-radius circular market in which I assume they participate. The data was split according to the direction of the change of integration status and in both cases (switch to integration and switch to disintegration) the mills that did not change status were included. The coefficients on the concentration variables have the correct sign in all cases, with the exception of that on the change in concentration variable in the pooled panel of the switchers who integrated. The fixed effects estimation corrects the sign but renders this variable insignificant. However, the power of the lagged market concentration to explain why mills choose to integrate or disintegrate is confirmed by these results. The coefficients on the year dummies are found to be strongly significant, signaling that there must have been industry-specific shocks that also affected the firms´ decision processes.

Conclusions

Paper mills choose to vertically integrate into pulp production because internalizing the production of their input allows them to avoid the transaction costs associated with outsourcing. The empirical exercises presented above contribute elements that confirm the role of asset and site specificity as conditions that render integration convenient from a cost-side perspective. In addition my results point towards cost benefits, arising from economies of scale, that make a large paper producer more likely to integrate than a small one. However, the reduced-form specification of the model does not properly address the possible endogenous relation of mill size and vertical integration that has been suggested above. Similarly, although the panel analysis portion of this chapter contributes to clear the role of market concentration in explaining a paper mill's decision to vertically integrate, the possible endogeneity of this relationship also needs to be explored further.

In addition, the fact that we observe paper mills of similar sizes operating in close proximity, making different choices about integrating or not, makes one think that there must be some idiosyncratic mill specific characteristics driving these decisions, that is not adequately captured by the current specification of the model.

Chapter 4

Dynamic Model of Vertical Integration

4.1 Motivation

The primary focus of this research is to learn about the factors that influence the decision of a manufacturing firm to vertically integrate into the production of its input. The reduced form econometric approach presented in the previous section sheds some light on the empirically observable regularities and motives that shape a paper mill's integration decision. My results, however, appear to suffer from an endogeneity bias. The reduced form exercise also provides no answer to the fact that mills of similar characteristics make different decisions with regards to their integration status. The insight is that there must be some unobserved mill characteristic that drives the mill in its decision process.

In order to deal with these issues, I propose a dynamic model á la Olley and Pakes (1996) that accounts for unobserved plant-specific characteristics and for entry and exit. This approach takes care of the estimation biases that are a concern from previous work, and allows me to learn about the unobserved characteristics of the firms in my data, and use them to determine who vertically integrates and who does not. In addition, the model I propose allows me to learn about how vertical integration affects productivity and mill’s entry and exit decisions.

Two features make the U.S. paper industry particularly suitable for my purposes. First, over the period 1975 to 1995 less than half of the operating paper mills in the U.S. were integrated backwards into pulp production. These mills accounted, however, for about 80% of the country's paper capacity each year. It is, thus, a perfect setting to try to sort out the different motives for a firm to vertically integrate. Second, the industry displays very high rates of entry and exit over the period. While there were on average 550 active paper mills each year, only 290 of the 584 mills operating in 1975 survived to 1995.

4.2 Related Literature

4.2.1 Dynamic Models of Firm Behavior

Models that explicitly incorporate the dynamics consistent with some firms thriving while others lag and exit include Jovanovic (1982), Hopenhayn (1992), and Hopenhayn and Rogerson (1993).

Jovanovic (1982) proposes a passive learning model of industry evolution based on firm heterogeneity and self-selection. Each firm has some true underlying production cost, c, which is a draw from a normal distribution with mean $\bar{c}$ and variance σ_c^2. The firm knows the cost distribution, but not its own cost parameter. Each period the firm's unit cost of production fluctuates randomly around the mean. From observing it over time, the firm learns about its underlying cost and is able to estimate it consistently as the time average of its observations. Each period the firm decides whether to exit or stay in operation based upon its current cost information. The evolution of the economy is then driven by the learning and selection decisions of these optimizing agents. Productivity can be viewed here as the dual of costs. From an analytical point of view, the notion that a firm's decision is based on its entire history of productivity draws would hardly be empirically tractable. The length of the dependence period would need to be restricted to that of the observed data.

Hopenhayn (1992) proposes a model in the same line of passive learning, in which firms are subject to a random productivity shock every period. This productivity shock follows a first-order Markov process that is independent across firms. The distribution of future productivity is assumed to be stochastically increasing in this period's productivity. Surviving firms pay a fixed cost each period, then observe their productivity shock, and decide on a level of output for that period. Entrants pay an entry fee, and then draw from a common underlying distribution of productivity shocks, and choose output. Exiting firms make no profits, and pay no costs. This framework allows Hopenhayn to derive equilibrium conditions that imply predictions about the productivity of entrants, incumbents, and exiting firms.

Hopenhayn and Rogerson (1993) propose a variant of this model and use it to evaluate the aggregate implications of government policies that make it costly for firms to adjust their labor. Using a value function that explicitly includes an adjustment cost for labor, they develop an equilibrium model of the reallocation process of labor across firms. They prove the existence of an equilibrium that has entry, exit, and the growth and decline of firms over time.

Ericson and Pakes (1995) introduce the concept of active learning. They develop a dynamic model of a small, imperfectly competitive industry with a stochastic process of accumulation for the state variable, in which an active exploration process drives the firms' learning about their efficiency. A variation of this framework may be applied to investigate how the firm's learning process about its efficiency determines its decision to be vertically integrated or not, and how this learning process itself is affected by the firm's choice of integration status.

4.2.2 Estimation of Productivity

Since productivity is not directly observable, to examine it requires one to first of all find a way of measuring it. If one is going to estimate total factor productivity, the simplest way to do this is to estimate a production function using OLS and use the residuals from such a regression as the measure of productivity. The problem with this approach was pointed out long ago by Maarshak and Andrews (1944): input choices are likely to be correlated with unobserved productivity. To the extent that this happens, the OLS estimates will be biased and will yield a biased measure of productivity. The usual approach to deal with this simultaneity problem is to use Instrumental Variables estimators. However, with plant-level data it is very hard to find valid instruments, because most variables that are correlated with input choices are correlated with productivity. A solution to the unavailability of appropriate instruments has often been to adopt a fixed effects estimator, but this estimator assumes that firm-level productivity is constant over time.

As an alternative, Cornwell et. al. (1990) use a plant-specific and time-varying efficiency that can be described as a quadratic function of time. This methodology is also used in Liu (1993), and Liu and Tybout (1996). It requires one to first estimate the production function by fixed effects in order to obtain the input coefficient vector. The residuals are calculated by subtracting the actual values from the predicted values of output. For each plant, these residuals are regressed on a constant, time, and time squared. The productivity measure is constructed using the estimates of the coefficients from the last regression. This approach improves on the fixed-effects methodology, but since it requires a parametric specification of productivity many degrees of freedom are lost in the estimation process. Moreover, this procedure still uses fixed effects estimation in the first step that provides the residual for the construction of the productivity measure. So although the measure does vary over time, it is still likely to be based on biased coefficients in the presence of simultaneity.

Olley and Pakes (1996) provide a methodology that deals explicitly with the simultaneity problem described above. They also deal with the selection problem arising from the fact that firms' exit decisions depend on their perceptions of future productivity, which are in turn partially determined by their current productivity, making balanced panel data sets to be selected in part on the basis of unobserved productivity realizations. Their contribution falls into both of the categories in this literature to the extent that it models firm behavior in a dynamic framework with the empirical goal of obtaining unbiased measures of plant-level productivity. Their model combines features of both the Ericson and Pakes (1995) model with features of the Hopenhayn and Rogerson (1993) model, and allows for firm-specific efficiency differences that exhibit idiosyncratic changes over time. To sort out the simultaneity problem, the model specifies the information available when input decisions are made. To control for the selection induced by exit, the model generates an exit rule.

Olley and Pakes (1996) provide a framework for analyzing the biases in traditional estimators that result from selection and simultaneity, and for building alternative algorithms that circumvent these biases. Variations of this model may, for instance, be applied to uncover the relationship between firm productivity and vertical integration in the presence of possibly endogenously determined variables such as firm size or market concentration.

4.3 The Model

My empirical goal is to learn about what causes a firm to vertically integrate backwards into the production of an intermediate input. In particular I want to learn about productivity, reveal its role in this decision process, and understand its relationship with vertical integration, accounting for the unobserved firm-specific features that underlie the firm's decision process. To do so I will employ a panel data set describing, at the plant-level, the integration and production decisions of a group of paper firms. Such firms must typically decide whether to integrate backwards into the production of pulp. A la Olley and Pakes (1996), I intend to learn about an unobserved attribute of these firms, i.e. their productivity, by exploiting the entry/exit features of my data. I maintain that productivity is not only a major determinant of whether a firm exits or enters the market, but it is also a major determinant of whether a firm vertically integrates backward or specializes in paper production. To ignore its role in a reduced form model can give rise to two problems. First, to the extent that differences in productivity are known to firms when they make their decision whether to vertically integrate or not (and productivity of a given firm has been shown to be highly correlated over time (Olley and Pakes, 1996)), there is a classic simultaneity problem. Second, the entry and exit observed in the industry over the period being studied raise the issue of how to handle attrition from, and additions to, the data. If firms' exit decisions depend on their perceptions of their future productivity, and if their perceptions are partially affected by their current productivity, then a panel sample will be partly selected on the basis of the unobserved productivity realizations, and there will be a selection bias in the estimation of the model.

A dynamic model of firm behavior that allows for firm-specific productivity differences allows me to deal with the selection and the simultaneity problems. To sort out the simultaneity problem, the model specifies the information available when the vertical integration decision is made. To control for the selection induced by exit decisions, the model generates an exit rule.

Moreover, it seems sensible to frame the decision of a firm to vertically integrate as a dynamic problem if we believe that firm managers are forward looking decision-makers, who take into account the effect of their choices today on the performance of their firms tomorrow, and the day after tomorrow, and so on. In pursuit of profit maximization, firm managers must act not only upon the current information available, but they must also take into consideration their expectations about the future.

The Incumbent Paper Mill Problem

I envision the incumbent paper mill as solving the following general optimization problem. At the beginning of every period, the firm first chooses whether to exit or to continue in the market. In making this decision, it compares its sell-off value in dollars to the expected present discounted value of the stream of profits that will accrue to it if it continues in operation. This present value, in turn, depends on whether the mill is specialized or vertically integrated. If the mill exits, it receives its sell-off value and never reappears again. If it decides to continue in operation it then chooses whether or not to produce its own input (whether to specialize, or vertically integrate). The mill will vertically integrate if its expected present discounted value of profits from integration exceeds the expected present discounted value of profits it does not integrate.

The current period operating profits of the nonintegrated paper mill

The current period profits of the nonintegrated paper mill i operating in market m are formally defined as

$$(1) \qquad \Pi^N_{im} = p_m(Q_m)*f^N_{im}(L_{im}, P_{im,} \omega_i) - w_m L_{im} - \pi^N_{im} (F_{im,} N_{i,}, M_m)P_{im}$$

where firm it's paper output, $f^N_{im}(L_{im},P_{im,}\omega_i)$ is written as a function of labor, L_{im}, pulp, $P_{im,}$ and a mill-specific index of unobserved productivity, ω_i, and $p_m(.)$ denotes the market price for paper, determined by the total market output of paper Q_m. In order to keep my model computationally tractable, I am assuming that mill i is small and takes the other mills' outputs as given (i.e. there is no multi-agent dynamic problem in this model).

On the costs side, the nonintegrated firm i pays the market wage rate, w_m, in exchange for labor, and the price π^N_{im} in exchange for pulp. The price for pulp is written as a function of asset-specific factors, which are costly to the nonintegrated firm, and of concentration in the pulp market. F_{im} represents closeness to forestland. F_{im} is a dummy variable that takes the value of 1 whenever the firm is located in the South or West regions, where most of the forestland in the U.S. is found, and it proxies for site specificity. N_i represents the production of newsprint, a standardized paper grade, that is usually produced on a larger scale than other specialized finer grades of paper which are produced in smaller batches (to order), and requires a less complex mix of pulps than the finer grades. It is also a dummy variable, and it proxies for asset specificity, because newsprint production requires the specialization of the pulping assets to conform to the requirements of the papermaker. M_m is a measure of market concentration in the input market. I expect that the costs associated with investment specificity as well as the costs associated with market concentration should be reflected in a higher pulp cost for the mill that is not integrated.

This profit function is written under the assumption that the firm produces a single paper grade, and uses a single pulp type in its production process. While this is true in many cases, in others it is a simplification that we should be aware of.

The Current Period Operating Profits of the Integrated Paper Mill

Similarly, under the assumption that vertically integrated firms do not participate as suppliers in the pulp market, the current period profits of the vertically integrated firm i operating in market m can be written as

$$(1) \qquad \Pi^I_{im} = p_m(Q_m)*f^I_{im}(L_{im}, P_{im,} \omega_i) - w_m L_{im} - \pi^I_{im} (H_{im}(\theta_m))P_{im}$$

As before, $f^I_{im}(L_{im}, P_{im,} \omega_i)$ represents firm i's production and p_m denotes the market price for paper. On the costs side, firm i again pays the market wage, w_m for labor, L_{im}. Now, however it pays a different price for pulp that is denoted as $\pi^I_{im.}$ This price is assumed to be lower than π^N_{im}. It is written as a function of $H_{im}(.)$, which represents the latent liability costs and other costs arising from the environmental regulation under the Clean Water Act of 1972 and its amendments. Environmental regulation under the Clean Water Act affects the pulp industry since it is a highly water polluting industrial activity.[8] These costs are assumed to be an increasing function of the water quality standards set by the state in which the integrated firm operates, $\theta_{m.}$ To explain the reasoning that underlies this assumption, in what follows I briefly summarize the dispositions of the Clean Water Act.

The Clean Water Act required all municipal and industrial wastewater to be treated before being discharged into waterways. Industries were given until July 1977, to install "best practicable control technology" (BPT) to clean up waste discharges. By March 1989, industries were required to use the "best available technology" (BAT) that was economically achievable. In addition, states were required to implement control strategies for water that was expected to remain polluted even after industrial dischargers had installed the best available cleanup technologies required under the law. The federal government sets the agenda and standards for pollution abatement, while the states carry out the day-to-day activities of implementation and enforcement. The states are responsible for establishing water quality standards, consistent of a designated use (recreation, water supply, industrial, or other), and a numeric or narrative statement identifying maximum concentration levels of various pollutants that would not interfere with the designated use.

Delegated responsibilities under the Act include authority for qualified states to issue and enforce discharge permits to industries and municipalities.[9] All discharges into the nation's waters are unlawful, unless specifically authorized by a permit. Permits specify the control technology applicable to each pollutant, the effluent limitations a discharger must meet, and the deadline for compliance. They are issued for 5-year periods and must be renewed thereafter to allow continued discharge. The Environmental Protection Agency may issue a compliance order or bring a civil suit in a U.S. district court to violators of the terms of a permit. The penalty for such violations can be as high as $25,000 per day. Stiffer penalties are authorized for criminal violations of the Act.

States with higher water quality standards do not allow as much polluting discharges and cause pulp producers to be more at risk of committing a violation of their polluting permits. Hence, vertically integrated firms face higher latent liability costs when operating in states with relatively high water quality standards. In addition, it is to be expected that firms will want to avoid as much as possible appearing in the news as environmental law-breakers. This could be particularly harmful to large multi-plant firms that operate in several states. So the environmental regulation is likely to be more costly to large vertically integrated firms. Finally, the environmental regulation under the Clean Water Act often implies that firms must invest in new cleaner technologies. To this extent the more productive firms are expected to face lower environmental costs, since they are expected to be already employing the best available technologies.

---

[8] Federal Water Pollution Control Act Amendments, Public Law 92-500, 1972; Clean Water Act of 1977, Public Law 95-217; Municipal Wastewater Treatment Construction Grants Amendments, 1981, Public Law 97-117; Water Quality Act of 1987, Public Law 100-4.

[9] As of December 1998, 43 states had been delegated the permit program. The Environmental Protection Agency (EPA) issues discharge permits in the remaining states.

The Dynamic Problem

Define Y_t as a binary variable such that $Y_t = 1$ if the incumbent firm is vertically integrated at time t, and $Y_t = 0$ if the incumbent firm is specialized at time t. Also suppose a firm incurs a cost of F dollars when it decides to integrate, and receives a sell-off value of Z dollars when it decides to disintegrate.

I assume that firm managers plan their market participation and their vertical integration status to satisfy:

$$V_{t_o} = \max\left\{ Y_{(t-1)}\Phi^I + (1-Y_{(t-1)})\Phi^N,\ \sup_{|Y_{(t+\tau)}|_{\tau=0}^{\infty}} E_{t_o}\left[\sum_{\tau=t_o}^{\infty} \delta^{(\tau-t_o)}\begin{bmatrix} Y_{(t+\tau)}\left[\Pi^I_{(t+\tau)} - \left(1-Y_{(t+\tau-1)}\right)F\right] \\ + \left(1-Y_{(t+\tau)}\right)\left[\Pi^N_{(t+\tau)} + \left(Y_{(t+\tau-1)}\right)Z\right]\end{bmatrix}\right]\right\}$$

where the profits of the integrated and the nonintegrated mill are functions of the state variable ω, whose evolution can affect the mills' choices to vertically integrate or not, as well as to remain active, and of X_t^I and X_t^N, the sets of exogenous variables that help determine the profits of each type of mill respectively:

$$\Pi_t^I = \Pi_t^I\left(\omega_t, X_t^I\right): \Omega \times X \to \Re$$
$$\Pi_t^N = \Pi_t^N\left(\omega_t, X_t^N\right): \Omega \times X \to \Re$$

Φ^I and Φ^N represent the exit fees of the integrated and of the nonintegrated firm respectively, E_t is an expectations operator conditioned on the set of information available at time t, and δ is the one period (i.e., 5 year) discount rate.

The Bellman equation for an incumbent mill can then be written as:

$$\begin{aligned} V_t\left(\omega_t, Y_{(t-1)}, X_t^I, X_t^N\right) = \max\{\ & Y_{(t-1)}\Phi^I + (1-Y_{(t-1)})\Phi^N, \\ & \sup_{y_t}\left[Y_t\left[\Pi_t^I\left(\omega_t, X_t^I\right) - \left(1-Y_{(t-1)}\right)F\right] + \right. \\ & \left.(1-Y_t)\left[\Pi_t^N\left(\omega_t, X_t^N\right) + Y_{(t-1)}Z\right] + \delta E_t\left(V_{t+1}/\omega_{t+1}, Y_t, X_{t+1}^I, X_{t+1}^N\right)\ \right]\ \} \end{aligned}$$

This equation states that the value today is composed of current returns plus the expected discounted value of tomorrow given optimal behavior in the interim. The solution to this functional equation generates an exit rule and a vertical integration policy function that can be combined into a single policy rule for each mill type. A mill has three possible choices of status: it can choose to exit, it can choose to stay in operation and be disintegrated, and it can choose to stay in operation and be integrated. If a function R_t is defined to be equal to zero

(0) if the firm exits, one (1) if the firm chooses to be in and disintegrated, and two (2) if the firm chooses to be in and integrated, then the policy rule that determines the mill's choice of status can be written as

$$(3) \qquad R_t(Y_{(t-1)}) = \begin{cases} 0 & \omega_t \leq \underline{\omega}_t^E\left(X_t^I, X_t^N, Y_{(t\text{-}1)}\right) \\ 1 & \underline{\omega}_t^E\left(X_t^I, X_t^N, Y_{(t\text{-}1)}\right) < \omega_t \leq \underline{\omega}_t^I\left(X_t^I, X_t^N, Y_{(t\text{-}1)}\right) \\ 2 & \omega_t > \underline{\omega}_t^I\left(X_t^I, X_t^N, Y_{(t\text{-}1)}\right) \end{cases}$$

The firm behavior characterized above implies that disintegrated producers vertically integrate whenever

$$(5) \qquad \Pi_t^I(\omega_t, X_t^I) - \Pi_t^N(\omega_t, X_t^N) + \delta[E_t(V_{(t+1)}/(Y_t = 1)) \text{ - } E_t(V_{(t+1)}/(Y_t = 0))] > (1 - Y_{(t-1)})F + (Y_{(t-1)})Z$$

Vertically integrated incumbents that choose to stay in operation choose to remain vertically integrated whenever current net operating profits from vertical integration plus the expected discounted future payoff from remaining vertically integrated exceeds the sell-off value from disintegration, Z. Nonintegrated incumbents choose to integrate whenever this sum, net of the fixed cost to vertically integrate, F, is positive.

4.4 The Model Simplified for Estimation

Paper mills are seen as maximizing their average profits or, equivalently, minimizing their costs per unit for a given amount of production. This is a simplified view of how they operate that places the focus on productivity maximization and allows me to temporarily overlook the mill's choice of capacity. In this section I derive the functional forms that will be taken to the data.

The Paper Mill's Production Function

I assume that the industry uses a Cobb-Douglas technology to produce paper. In the short-run perspective, where capital is fixed, the production function of paper firm i can be written as:

$$Q_i = f(\omega_i)L_i^{\alpha}P_i^{\beta}$$

where L_i is labor, P_i is pulp, and $f(\omega_i)$ is a function of an index of unobserved productivity, ω_i, which is a firm-specific state variable. I further assume that the technology used in paper production is constant returns to scale, so $0<\alpha<1$, $0<\beta<1$, and $(\alpha + \beta) = 1$.

This production function is written under the assumption already stated that the mill produces a single paper grade, and uses a single pulp type in its production process. I am also maintaining the assumption that vertically integrated mills do not participate as suppliers in the pulp market.

The Paper Mill's Average Cost Function

Paper firm i, operating in market m, faces the following cost minimization problem:

$$\min\left(w_m L_i + \pi P_i\right) \qquad \text{s.t. } \bar{q}_i = f\left(\omega_i\right) L_i^{\alpha} P_i^{\beta}$$

where w_m is the wage rate in market m, π is the cost of pulp in market m for firm i, and $\bar{q}_i$ is the fixed amount of output that firm i wants to produce. The standard solution to this problem yields the following cost function for firm i:

$$C(w_m, \pi, \bar{q}_i, \omega_i) = \frac{\bar{q}_i (w_m)^{\alpha} (\pi)^{\beta}}{f(\omega_i)} [\beta/\alpha]^{-\beta} [1 + \beta/\alpha]$$

The mill's average cost function can then be written as

$$\Rightarrow \frac{C(w_m, \pi, \bar{q}_i, \omega_i)}{\bar{q}_i} = \frac{(w_m)^{\alpha} (\pi)^{\beta}}{f(\omega_i)} [\beta/\alpha]^{-\beta} [1 + \beta/\alpha]$$

I expect that the integrated and the nonintegrated paper mills differ in the way their unobserved productivity enters their cost functions. Hence,

$$f^N(\omega_i) \neq f^I(\omega_i)$$

where the superscript N denotes the nonintegrated paper firm, and the superscript I denotes the integrated paper firm. If it is true that unobserved productivity explains in part why some paper firms choose to be vertically integrated and some do not, then this must be true. I further expect that $f^I(\omega_i) > f^N(\omega_i)$. This should be demonstrated by the data.

The average cost functions of the integrated and the nonintegrated paper firms will also differ in the price firms ultimately pay for pulp, π. In the case of the nonintegrated firm, π^N is meant to reflect not only the market price per ton of pulp, but also additional costs incurred by the firm such as transaction costs from site or asset specificity, or costs from having to operate in a concentrated pulp market. In the case of the integrated firm, π^I is meant to reflect the marginal cost of producing a ton of pulp.

The Paper Mill's Current Period Average Profit Function

The integrated mill

$$\frac{\Pi^{I}\left(w_m, \pi^{I}, \bar{q}_i, \omega_i\right)}{\bar{q}_i} = \bar{p} - \frac{(w_m)^{\alpha}\left(\pi^{I}\right)^{\beta}}{f^{I}(\omega_i)}\left[\beta/\alpha\right]^{-\beta}\left[1+\beta/\alpha\right]$$

The disintegrated mill

$$\frac{\Pi^{N}\left(w_m, \pi^{N}, \bar{q}_i, \omega_i\right)}{\bar{q}_i} = \bar{p} - \frac{(w_m)^{\alpha}\left(\pi^{N}\right)^{\beta}}{f^{N}(\omega_i)}\left[\beta/\alpha\right]^{-\beta}\left[1+\beta/\alpha\right]$$

where $\bar{p}$ is the price per unit of paper.

The Dynamic Problem

As before, Y_t is a binary variable set equal to 1 if the incumbent mill is vertically integrated at time t, and equal to 0 if the incumbent mill is specialized at time t. Also as before, a nonintegrated mill incurs a cost of F dollars when it decides to integrate. For identification purposes, the salvage value of Z dollars that an integrated mill receives if it chooses to disintegrate is set to zero.

Managers plan their market participation and their vertical integration status to satisfy:

$$V_{t_o} = \max\left\{\Phi, \sup_{\left|Y_{(t+\tau)}\right|_{r=0}^{\infty}} E_{t_o}\left[\sum_{\tau=t_o}^{\infty}\delta^{(\tau-t_o)}\begin{bmatrix} Y_{(t+\tau)}\left[\Pi^{I}_{(t+\tau)}/\bar{q} - \left(1-Y_{(t+\tau-1)}\right)F\right] \\ +\left(1-Y_{(t+\tau)}\right)\left[\Pi^{N}_{(t+\tau)}/\bar{q}\right]\end{bmatrix}\right]\right\}$$

where Φ represents the exit fee to be paid by the mill that decides to leave the industry, E_t is an expectations operator conditioned on the set of information available at time t, δ is the one period (i.e., 5 year) discount rate, and $\prod^{I}_{(t+\tau)}/\bar{q}$ and $\prod^{N}_{(t+\tau)}/\bar{q}$ denote the average profit functions of the integrated and of the nonintegrated paper mills respectively,

At the beginning of every period the incumbent mill has two decisions to make. The first is to decide whether to exit or continue in operation. If it exits, it receives a sell-off value of Φ dollars, and never reappears again. If it continues, it chooses its integration status ($Y_t = 1$, or $Y_t = 0$). If at time (t-1) the paper mill was not integrated ($Y_{(t-1)} = 0$) and at time t it chooses to integrate ($Y_t = 1$), its value function will be composed of its current period average profits as an integrated mill minus a cost to integrate of F dollars, plus the present discounted value of its expected future average profits as an integrated mill, given the current information. If it chooses to remain disintegrated ($Y_t = 0$), its value function will be composed of its current period average profits as a disintegrated mill, plus the present discounted value of its expected future average profits as a disintegrated mill, given current information. If at time (t-1) the mill was integrated ($Y_{(t-1)} = 1$) and at time t it chooses to disintegrate ($Y_t = 0$), its value function will be composed of its current period average profits as a disintegrated mill plus the present discounted value of its expected average profits as a disintegrated mill, given the current information. Finally, if the mill was integrated at time (t-1) ($Y_{(t-1)} = 1$) and it chooses to remain integrated ($Y_t = 1$), its value function will be composed of its current period average profits as an integrated mill plus the present discounted value of its expected future average profits as an integrated mill, given current information. Therefore, both the exit and the integration decisions will depend on the mill's perceptions of the future, given current information. The exit and integration decisions generated by these perceptions will, in turn generate the market structure (i.e. productivity and integration status pair for each active mill) for the future years.

The Bellman equation for an incumbent mill can then be written as:

$$
\begin{aligned}
V_t\left(\omega_t, Y_{(t-1)}, w_t, \pi_t^I, \pi_t^N\right) = \max\Big\{ \Phi,\ & \sup_{y_t}\Big[Y_t\left[AVG\Pi_t^I\left(\omega_t, w_t, \pi_t^I\right) - \left(1 - Y_{(t-1)}\right)F\right] \\
& + (1 - Y_t)\left[AVG\Pi_t^N\left(\omega_t, w_t, \pi_t^N\right)\right] \\
& + \delta E_t\left(V_{t+1}/\omega_{t+1}, Y_t, w_{t+1}, \pi_{t+1}^I, \pi_{t+1}^N\right)\Big]\Big\}
\end{aligned}
$$

where $AVG\Pi_t^I(.)$ and $AVG\Pi_t^N(.)$ denote the average profits of the integrated and the nonintegrated mill respectively.

The status choice rule generated by the solution to this problem is unchanged relative to that presented in the previous section:

$$
\text{(3)} \qquad R_t(Y_{(t-1)}) = \begin{cases} 0 & \omega_t \le \underline{\omega}_t^E\left(w_t, \pi_t^I, \pi_t^N, Y_{(t-1)}\right) \\ 1 & \underline{\omega}_t^E\left(w_t, \pi_t^I, \pi_t^N, Y_{(t-1)}\right) < \omega_t \le \underline{\omega}_t^I\left(w_t, \pi_t^I, \pi_t^N, Y_{(t-1)}\right) \\ 2 & \omega_t > \underline{\omega}_t^I\left(w_t, \pi_t^I, \pi_t^N, Y_{(t-1)}\right) \end{cases}
$$

Stokey, Lucas and Prescott (1989) give in detail the conditions under which any V(.) which solves the Bellman equation is equal to V* defined as the supremum of the value function evaluated at the optimal policy path. They also provide the conditions under which any $R_t(.)$ that solves the Bellman equation, is an optimal policy for V*. A sufficient condition for both of the above to hold, is that V*(.) be bounded.

Define B as the set of bounded profit functions of both types of mills on $\Re$ (assume that profits are bounded from above), and define $\rho(.)$ as the sup norm i.e. $\rho(f,g) = \sup_{x\in\Re}|f(x)-g(x)|$. The Bellman equation is a fixed point to the operator T where T: B→ B and is defined pointwise by

$$
\begin{aligned}
Tf\left(\omega', w', \pi^{I}{}', \pi^{N}{}', Y\right) = \max \{ \; & \Phi, \\
& \sup_{y_t} \{Y'\left[AVG\,\Pi^{I}\left(\omega', w', \pi^{I}{}'\right) - (1-Y)F\right] \\
& + (1-Y')\left[AVG\,\Pi^{N}\left(\omega', w', \pi^{N}{}'\right)\right] \\
& + \delta\int_{\omega'', w'', \pi^{I}{}'', \pi^{N}{}''} f\left(\omega'', w'', \pi^{I}{}'', \pi^{N}{}'', Y'\right) p(\omega'' \mid \omega', w', \pi^{I}{}', \pi^{N}{}', Y) \; \}
\end{aligned}
$$

The exogenous variables in this problem enter as state variables for which the mill manager is assumed to have myopic expectations. The stochastic process generating the sequence of realizations $\{\omega_t\}$ is a Markov Process, the distribution of $\omega_{(t+1)}$ being determined by ω_t. If X denotes the set of exogenous state variables, then there is a set of distribution functions for each possible (ω, X) couple. This determines the distribution of $\omega_{(t+1)}$ given ω_t:

The mill behavior characterized above implies that paper producers stay in the market as long as the present discounted value of their expected average profits if they stay in operation is more than the fee they receive if they choose to exit. If they choose to stay in operation then, their choice to vertically integrate is driven by

$$
\text{(5)} \qquad AVG\,\Pi_t^{I}(\omega_t, X_t^{I}) - AVG\Pi_t^{N}(\omega_t, X_t^{N}) + \delta[E_t(V_{(t+1)}/(Y_t = 1)) - E_t(V_{(t+1)}/(Y_t = 0))] > (1-Y_{(t-1)})F
$$

A vertically integrated mill will choose to continue operating as vertically integrated whenever the profit difference above is greater than zero (its sell-off value from disintegration). A nonintegrated mill will choose to integrate whenever this amount is greater than the fixed cost to vertically integrate, F.

The dual problem

Equivalently, the problem facing the mill can be approached directly from the cost side. Productivity maximization is equivalent to minimization of costs per unit. The mill's behavior can alternatively be characterized by:

$$
V_{t_o} = \min\left\{ \; -\Phi, \; \inf_{|Y(t+\tau)|_{\tau=0}^{\infty}} \left[E_{to}\left[\sum_{\tau=t_o}^{\infty} \delta^{(\tau-t_o)} \left[Y_{(t+\tau)}\left[\overline{C^{I}_{(t+\tau)}} + \left(1-Y_{(t+\tau-1)}\right)F\right] + \left(1-Y_{(t+\tau)}\right)\left[\overline{C^{N}_{(t+\tau)}}\right]\right]\right]\right] \right\}
$$

where Φ represents the exit fee to be paid by the mill that decides to leave the industry, E_t is an expectations operator conditioned on the set of information available at time t, δ is the one period (i.e., 5 year) discount rate, and $\overline{C_t^I}$ and $\overline{C_t^N}$ denote the average cost functions of the integrated and of the nonintegrated paper mills respectively,

$$\overline{C^I} = \frac{C^I\left(w_m, \pi^I, \overline{q}_i, \omega_i\right)}{\overline{q}_i} = \frac{(w_m)^\alpha \left(\pi^I\right)^\beta}{f^I(\omega_i)} [\beta/\alpha]^{-\beta} [1+\beta/\alpha]$$

$$\overline{C^N} = \frac{C^N\left(w_m, \pi^N, \overline{q}_i, \omega_i\right)}{\overline{q}_i} = \frac{(w_m)^\alpha \left(\pi^N\right)^\beta}{f^N(\omega_i)} [\beta/\alpha]^{-\beta} [1+\beta/\alpha]$$

The Bellman equation for an incumbent mill is written as:

$$\begin{aligned} V_t\left(\omega_t, Y_{(t-1)}, w_t, \pi_t^I, \pi_t^N\right) = \min\Big\{ & -Y_{(t-1)}\Phi^I - (1-Y_{(t-1)})\Phi^N, \\ & \inf_{y_t}\Big[Y_t\left[\overline{C_t^I}\left(\omega_t, w_t, \pi_t^I\right) + \left(1-Y_{(t-1)}\right)F\right] \\ & + (1-Y_t)\left[\overline{C_t^N}\left(\omega_t, w_t, \pi_t^N\right)\right] + \delta E_t\left(V_{t+1}/\omega_{t+1}, Y_t, w_{t+1}, \pi_{t+1}^I, \pi_{t+1}^N\right)\Big]\Big\} \end{aligned}$$

The status choice rule generated by the solution to this problem is identical to the one discussed above.

Chapter 5

Simulation Exercise

5.1. Motivation

The aspect of firm behavior that the dynamic model presented in the previous section captures relies on the belief that productivity is not only a major determinant of whether a firm exits or enters the market, but also a major determinant of whether a firm vertically integrates backward or specializes in paper production. By explicitly allowing for the estimation of firm level unobserved productivity, the dynamic approach proposed in this dissertation takes care of the estimation biases that may arise from simultaneity and selection in a reduced form model, which ignores its role in the firm's decision process.

In order to show the workings of my dynamic approach, and to better motivate the estimation exercise that is presented in the following chapter, I have used my model to simulate data for 500 firms over 5 periods. In this section I present a brief description of the process by which the data were simulated, and I present the results obtained when these data are used in the estimation of a reduced form model of vertical integration, similar to the model presented in Chapter 3.

5.2 Simulation Process

I use my dynamic model to build an artificial dataset of 500 mills over ten periods of time. I do this by simulating the behavior of each mill, i.e. each mill's choice of policy for every point in time, where the choices available to the mill at the beginning of each period are to stay in operation and produce its own input, to stay in operation and obtain its input through outsourcing, or to exit the industry. A mill that exits never reappears in the industry.

The model is set so that every period the previous integration status of a mill affects the policy path that mill will follow (the previous period integration status enters as a state variable in the value function iteration). So for each particular mill the process of simulating a chain of data starts with a random draw for the previous period integration status, and a random draw for each of the variables that enter the mill's average cost function that the mill minimizes when it makes its status choice. The value function for this mill at time t is calculated for these specific values, using a given set of parameters. The outcome of the value function iteration is a policy choice for each of the possible values of the unobserved productivity index, ω[10]. The next step is to make a draw for the mill's unobserved productivity index in period t, ω_{it}, and to use the solution of the value function to determine the behavior of the mill, given the values of the simulated data.

[10] Details about the value function estimation are presented in Chapter 6.

The process is repeated for periods (t+1) and after, with the difference that now the previous period integration status is no longer a random draw, but is instead taken from the previous period simulated policy decision made by the mill. Also, the previous period draws for the variables that change over time, will now affect the distribution from which the new draws will be made. The assumption made in most cases is that the exogenous variables follow a random walk, in which the coefficient on the previous period observation of each variable is set equal to 1 (i.e. the previous period observation for variable j at time t, becomes the mean of the distribution of variable j at time (t+1)). The updates in all exogenous variables are assumed to be normally distributed. The only variable whose distribution is set to follow a more complicated Markov process over time is the market concentration measure. I allow the period t unobserved productivity draw for mill i, to affect the path followed by the market concentration faced by that mill from time (t+1) on. I do this by letting the productivity shock at time t (ω_{it} – mean (ω_{it})) alter the mean of the distribution of the market concentration measure at time (t+1). My intention is to replicate the endogeneity of market concentration and vertical integration, that is a concern from the reduced form model estimation presented in Chapter 3. I want to build artificial data that will induce the biases in estimation that I found when dealing with the real data. This will then allow me to show how the structural dynamic model that I am proposing takes care of these biases.

For each period the mill’s value function is estimated, and a set of policy choices for each possible ω value is obtained. A new draw for the mill’s current period ω is used, as before, to determine the behavior of the mill, given the values of the simulated data. Unobserved productivity for each mill is assumed to follow a parameterized Markov process.

The process described above was replicated 500 times, and the outcome is the artificial dataset described at the beginning of this section. Since the working of my dynamic model is based upon the explicit estimation of the unobserved productivity of each mill and the artificial data were constructed using the model, this dataset has a particular feature: for each mill, for each point in time, and for each set of artificial variables used to simulate mill behavior, there is also a mill-specific productivity measure, ω_{it}. I point this out, because this is what will allow me to show the usefulness of the dynamic model approach proposed in the previous chapter, for dealing with the biases from simultaneity and selection that can arise in the estimation of a reduced form model.

5.3 Functional Forms and Parameters

In this section I introduce the functional forms, and the parameter values I have used to perform this simulation exercise.

Unobserved Mill Productivity

(1) $f^N(\omega_i) = A^N (\omega_i)^{\eta 1}$

(2) $f^I(\omega_i) = A^I (\omega_i)^{\eta 2}$

where ω_i is the index of productivity, known to the firm but unobserved by us.

The productivity functions of both types of mills are assumed to be nonnegative (A^N => 0, A^I => 0) and concave (0 <= η1<= 1, 0 <= η2 <= 1) over the relevant ω_i range in order to ensure the convexity of the value function.

Markov Process Governing Unobserved Productivity

The unobserved productivity index, ω_{it}, is assumed to evolve over time according to an exogenous Markov process:

$$\omega_{i(t+1)} = \gamma_o + \gamma_1\,\omega_{it} + \varepsilon_{it},$$

where ε_{it} is a normally distributed error term.

Price of pulp

(1) $\pi^N(M) = \lambda_0 + \lambda_1 M + \lambda_2 SW + \lambda_3 AS$

(2) $\pi^I(.) = \lambda_4$

where M is a measure of market concentration, SW is a forest land dummy thought as a proxy for site specificity, and AS is some asset specificity measure.

Current Period Average Costs

Incorporating the above, I arrive at the average cost functions used in the process of data simulation:

$$\overline{C^I} = \frac{C^I\left(w_m, \pi^I, \overline{q}_i\right)}{\overline{q}_i} = \frac{(w_m)^{\alpha}\left(\pi^I\right)^{\beta}}{f^I(\omega_i)}\left[\beta/\alpha\right]^{-\beta}\left[1+\beta/\alpha\right]$$

$$\overline{C^N} = \frac{C^N\left(w_m, \pi^N, \overline{q}_i\right)}{\overline{q}_i} = \frac{(w_m)^{\alpha}\left(\pi^N\right)^{\beta}}{f^N(\omega_i)}\left[\beta/\alpha\right]^{-\beta}\left[1+\beta/\alpha\right]$$

Parameters

Table 5.1 presents the parameter values from which the artificial data were created.

5.4 Simulated data

Table 5.2 presents summary statistics of the artificial data used in the reduced form estimation of section 5.5. The wage was simulated because it is a variable required to estimate the static average cost functions that enter the structural model. It is not used in section 5.5. Table 5.2 does not include the observations that correspond to exit choices. Table 5.3 presents some more telling integration and exit status statistics to further describe the simulated dataset. There is no entry in my artificial data. In period 1 there are 500 firms and I track their integration and exit behavior overtime. This is an important difference with the real paper industry data. Also, the simulated data show a smaller proportion of disintegrated mills each period.

5.5 Bias in Reduced Form Estimation

I used the artificial dataset to estimate a reduced form model of vertical integration similar to that presented in Chapter 3:

$$VI_i = \alpha_0 + \alpha_1 CONC_i + \alpha_2 SW_i + \alpha_3 ASSET\ SPECIFICITY_i + \varepsilon_i$$

This version of the model was estimated for the pooled dataset and also as a panel, allowing for random and fixed effects, and in each case the estimation was repeated with a variation: the inclusion of the mill-specific productivity measure among the explanatory variables. Table 5.4 presents the results.

By explicitly accounting for unobserved productivity, whose role is typically ignored in a reduced form model, I am able to produce a strong variation in the size and significance of the coefficient on the variable that is by construction endogenous to vertical integration. This will also be true for the real data, if market concentration is truly endogenous. As well, the explicit accounting for the unobservable takes care of the possible bias from selection, although this is harder to show in a reduced form framework. Recall that a selection bias in the estimation of the model arises from firms' exit decisions depending on their perceptions of their future productivity, and their perceptions being partially affected by their current productivity. The structural model looks explicitly at the mill-specific unobserved productivity realizations and provides an exit rule, based upon which a mill will decide to exit or stay in operation. Estimation in the dynamic model framework explicitly accounts for the role of unobserved mill productivity in exit. By doing so, it produces unbiased parameter estimates.

Chapter 6

Empirical Estimation

6.1. Functional Forms

In this Chapter, I describe in detail the steps required for the estimation of the structural model proposed in Chapter 4, and present some baseline results. The purpose of the empirical exercise presented here is to provide an idea of what the model allows. Because even the estimation of the model in its simplified version is computationally demanding, here I am using one of the most basic possible versions of it. The functional forms specified are identical to those used in the simulation exercise presented in Chapter 5, with the exception of the price of pulp for the nonintegrated firm, restricted even further for estimation:

Price of pulp

(1) $\pi^N(M) = \lambda_0 + \lambda_1 M$

(2) $\pi^I(.) = \lambda_2$

where M is a measure of market concentration.

6.2 Estimation Strategy

The estimation of the model involves dynamic programming techniques nested inside a nonlinear search algorithm over its parameters. The estimation strategy employed can be broken into steps, as follows.

Step 1: Derivation of the optimal policy rules

The unobserved mill productivity index, ω_i, can take any value in a range between any two numbers. I selected 0 and 30 as limiting values and constructed a grid spaced at increments of 0.1 over this range. I then used this grid to calculate the transition probability of ω, under the assumption that it is normally distributed. The selection of the limit values of the grid range is arbitrary, and as good as any other for the purposes of my estimation.

I started with a parameterized simple[11] version of the static average cost functions for both the mill that was integrated and the mill that was not integrated in the previous period, and calculated the static average costs of both types of mills for all the possible values of ω. Using these static cost functions, a guess of the first period's value of the mill, and a guess of the mill's status at time t (out, in and disintegrated or in and integrated), I was able to

[11] The only variable that entered these first-stage cost functions was the endogenous productivity index. All other variables were temporarily made into constants.

calculate the value function of both mill types for all possible ω values. I then used value function iteration to obtain the optimal choice of status policy rule, R_t, for both the mill that was integrated at time (t-1) and the mill that was not integrated at time (t-1).

The optimal policy rules obtained from the convergence of the value function iteration allow me to identify two sets of cut-offs on the grid of ω: the cut-off values of ω for which a nonintegrated firm will switch to integration and an integrated firm will disintegrate, and the cut-off values of ω below which an integrated and a nonintegrated firm will choose to exit the industry. These cut-offs are, however, bound by construction to fall on grid points. Since there is no reason for the correct policy cut-offs to fall exactly on grid nodes rather than in between any two of them, there is a need to restart the value function iteration in the "optimal" cut-off range, now allowing for the possibility of a solution in continuous space. This is possible by using linear splines between the grid nodes. The solution to this value iteration is still in the form of discrete cut-offs, but now not restricted to fall on grid nodes.

Figure 6.1 shows the choice of status policy rules of a formerly integrated ($Y_{(t-1)} = 1$) and of a formerly disintegrated ($Y_{(t-1)} = 0$) paper firm, with the corresponding exit and integration ω cutoffs. In this picture, the policy rule of each firm takes the value of zero (0) when the firm is out, one (1) when the firm is in and disintegrated, and two (2) when the firm is in and integrated.

It seems reasonable to expect that, when compared to a disintegrated mill of the same size, the integrated mill will exit and switch integration status at lower values of ω than the disintegrated. It takes a greater negative productivity shock to make the integrated mill change status because it is assumed to have started out in a higher productivity range. There is no reason, however, to expect that the pattern depicted above should hold in general for all integrated vs. nonintegrated mills, and ultimately we will allow the data to tell us what sort of pattern holds.

Step 2: Monte Carlo Integration of the expected likelihood function

I start by simulating for each mill a string of ω_i's that supports the data. I use a temporarily parameterized Markov process for ω, and the cutoffs from the choice of status rules that solve my value function, to take sequences of random draws for ω from a truncated normal distribution, where truncation each time occurs at the cutoff ωs. This procedure guarantees that my draws in each case come from the set of unobservables that support the observed data. Each mill's sequence of ω_i's is then such that its observed policy path can be exactly replicated. I use the following example to show how this is done.

Example

Denote the exit and integration cutoffs from the choice of status rules that solve my value function as ω_I^E and ω_I^I for the mill that was integrated at time (t-1) $\left(Y_{(t-1)} = 1\right)$. Correspondingly denote as ω_N^E and ω_N^E the exit and integration cutoffs for the mill that was not integrated at time (t-1) $\left(Y_{(t-1)} = 0\right)$. Then suppose we observe the status of a mill over five time periods. We start with time (t-1) in which the mill is active and not integrated, and we observe the following chain of status choices: at time t the firm remains active and disintegrated, at time (t+1) the mill is active and it has switched to integration, at time (t+2) the mill is active and back to being disintegrated, and at time (t+3) the mill is no longer in the dataset having chosen to exit. Table 6.1 tracks these changes of status and the corresponding path for this mill's unobserved productivity index, ω, and Figure 6.2 illustrates how the drawing from a truncated normal distribution is carried out.

Notice that the data from time (t-1), i.e. the first period for which I observe the mill, is used only to obtain the information about the mill's previous period's integration status, which is then used to determine what cutoffs are to be taken as truncation limits for the next period's random draw. The mean ω at time t is a parameter to be estimated. For the following time periods the mean is obtained directly from the Markov process that governs ω. In order to perform these draws, the other piece of information that I need is the variance of the ω distribution. This is yet another parameter that needs to be estimated.

I take 500 vectors of simulated ω paths for each mill in my data, such that if they had actually occurred they would have supported the policy paths that I observe. From the normal distribution and the information I have about the mean, the variance and the range in which ω falls each period, I am able to compute the probability of actually observing each ω path. I do this for each of the 500 vectors of simulated ω's that correspond to each mill, and average over them to obtain an accurate probability measure. I then convert these average probabilities into logarithms, and obtain the overall likelihood numerically, as the summation of these probabilities in logarithmic form.

More formally, the likelihood function of my model can be written as

$$L = \sum_{i=1}^{n} \left(\ln \left[\prod_{t=1}^{\tau} f\left(R_{it}\left(X_{it}, Y_{i(t-1)}\right), \theta\right) \right] \right)$$

where X_{it} represents the exogenous variables entering the value function, θ denotes the parameter vector, and $f(.)$ is the (average) normal density function giving us the probability for firm i at time t of the observed status actually occurring, i.e. the probability that we would actually observe a realization of ω_{it} in the ω_{it} range obtained from the choice of status rules that solve the value function.

Step 3: Inclusion of exogenous variables

Once the mechanics of estimation are working smoothly, the exogenous variables can be added. The full version of the model requires the inclusion of wage data and concentration measures to the average cost functions.

The addition of the exogenous variables builds upon the already existing computational structure. Recall that the exogenous variables enter the value function as state variables for which the mill manager has myopic expectations. For each possible combination of values taken by each of the exogenous variables there will be a new set of solution cut-offs. The computationally efficient way to go about this calls again for the use of linear interpolation splines. For each exogenous variable a set of representative values is selected for which the full value function iteration is performed and optimal exit and integration cut-offs obtained. Since for each mill there is an idiosyncratic combination of values in the exogenous variable data, the individual set of optimal cut-offs for each mill is then obtained by interpolation.

The overall likelihood of the model is obtained as before, only now the random draws for ω are obtained from a truncated normal distribution with different truncation points for each individual mill.

Step 4: Maximum likelihood optimization

The Monte Carlo integration procedure to obtain the expected likelihood function is used to perform a grid search over the model parameters in order to obtain good starting values, and is then nested in an optimization algorithm to obtain the final parameter estimates.

6.3 Baseline Results

Table 6.2 presents the results of my estimation exercise. As expected, the parameters of the functions of unobserved productivity that enter the average cost function of each type of firm, $f^{I}(\omega)$ and $f^{N}(\omega)$, are such that as mill productivity becomes higher, average costs decrease (i.e. for both types of firms $f(\omega)$ is an increasing function of ω). More importantly, the parameters of these functions of unobserved productivity are such that $f^{I}(\omega) > f^{N}(\omega)$ for mill productivity above a certain level (see Figure 6.3). The direction of the inequality is as expected, and the cost reducing impact of productivity becoming higher for the integrated mill only after a certain productivity level makes sense if one thinks about the cost in performance that could come about when a low productivity plant (a plant with poor management, for instance) decides to expand into its input production.

As for the parameters that enter the specification of the price of pulp for the non-integrated mill, λ_o is estimated relative to λ_2, the price of pulp of the integrated firm, with λ_2 set equal to 15, so the estimation outcome is a 1.03 to 1 ratio for the constant portion of the price. This is again as expected, since even before accounting for the implications of variables exogenous to the production process, you would expect the integrated firm to be able to get its own input at marginal cost. The parameter reflecting the impact of market concentration on the price of pulp faced by the non-integrated firm, λ_1, has the correct sign but is not significant.

6.4 Implications for Productivity

The empirical evidence presented in this Chapter is still not conclusive with respect to the impact of vertical integration on mill level productivity. The version of the model taken to estimation is stripped down to the very basics and to that extent cannot be expected to give a fully reliable account of reality. It does, however, contribute some intuitive results that allow me to show the usefulness of the model I propose, as a tool for unbiased estimation.

Table 6.3 shows the results from interacting in a regression framework the mill level unobserved productivity measure backed out of the estimation of the dynamic model presented in section 6.3, with the mills' integration decision. Results are presented for the pooled dataset, as well as for the random-effects, and the fixed-effects panels. The outcomes are not very different across these models. In all cases the coefficient on the integration status dummy variable is positive and significant, suggesting that indeed mill level productivity depends on the mill's integration status.

Table 6.4 provides a simple analysis of survival probabilities with my estimates of productivity growth. As theory predicts, the exit probability is negatively related to productivity change. These results confirm the importance of acknowledging the role of unobserved mill efficiency in estimation in order to avoid biases from selection.

Table 6.5 provides further evidence that the suspicions of endogeneity from the reduced form exercises presented in Chapter 3 may be well founded. The unobserved mill level productivity measures obtained in estimation are indeed positively correlated with firm size (as measured by production capacity), and negatively correlated to all of the alternative concentration measures considered.

Finally, table 6.6 gives a taste of the possible results that can be produced using the estimated unobserved mill level efficiency measures. Aggregate level productivity is calculated as the capacity weighted average of individual mill productivity.

Chapter 7

Extensions

7.1 Incorporating the Investment Decision

The dynamic model of vertical integration presented in Chapter 4, and taken to estimation in its more simplified version, overlooks the mill's choice of capacity (i.e. the mill's investment decision). For the purpose of this dissertation, such simplification is justified in terms of computational convenience, and because it still allows me to focus on the relationship between vertical integration and mill level productivity, which is what I intend to learn about. However, the characterization of mill behavior seems incomplete without taking into account its investment decision. In the remainder of this section I outline the structural model that would incorporate such decisions, allowing us to explore the three-fold relationship between investment choices, vertical integration and mill level productivity.

The New Dynamic Problem

As before, define Y_t as a binary variable such that $Y_t = 1$ if the incumbent firm is vertically integrated at time t, and $Y_t = 0$ if the incumbent firm is specialized at time t. Also suppose a firm incurs a cost of F dollars when it decides to integrate, and receives a salvage value of Z dollars when it decides to disintegrate.

Now we need an accumulation equation for capital, of the form:

$$k_{t+1} = \alpha k_t + i_t$$

where k_t represents the capital stock (i.e. installed capacity), and i_t represents the investment at time t.

Mill managers are seen as planning their market participation, their investment each period, and their vertical integration status to satisfy:

$$V_{t_o} = \max\left\{ Y_{(t-1)}\Phi^I + (1 - Y_{(t-1)})\Phi^N, \sup_{|Y_{(t+\tau)}|_{\tau=0}^{\infty}} E_{t_o}\left[\sum_{\tau=t_o}^{\infty} \delta^{(\tau - t_o)} \begin{bmatrix} Y_{(t+\tau)}\left[\Pi^I_{(t+\tau)} - \left(1 - Y_{(t+\tau-1)}\right)F\right] \\ + \left(1 - Y_{(t+\tau)}\right)\left[\Pi^N_{(t+\tau)} + \left(Y_{(t+\tau-1)}\right)Z\right] \end{bmatrix} \right] \right\}$$

where the profits of the integrated and the nonintegrated mill are functions of the state variable ω, whose evolution can affect the mills' choices to vertically integrate or not, as well as to remain active, the mill capacity level, k_t, which enters the value function as a new state variable, and X_t^I and X_t^N, the sets of exogenous variables that help determine the profits of each type of mill respectively:

$$\Pi_t^I = \Pi_t^I\left(\omega_t, k_t, X_t^I\right) : \Omega \times K \times X \to \Re$$
$$\Pi_t^N = \Pi_t^N\left(\omega_t, k_t, X_t^N\right) : \Omega \times K \times X \to \Re$$

Φ^I and Φ^N represent the exit fees of the integrated and of the nonintegrated firm respectively, E_t is an expectations operator conditioned on the set of information available at time t, and δ is the one period discount rate.

The Bellman equation for an incumbent mill can then be written as:

$$\begin{aligned} V_t\left(\omega_t, k_t, Y_{(t-1)}, X_t^I, X_t^N\right) = \max\{ & Y_{(t-1)}\Phi^I + (1 - Y_{(t-1)})\Phi^N, \\ & \sup_{y_t}[Y_t\left[\Pi_t^I\left(\omega_t, k_t, X_t^I\right) - \left(1 - Y_{(t-1)}\right)F\right] + \\ & (1 - Y_t)\left[\Pi_t^N\left(\omega_t, k_t, X_t^N\right) + Y_{(t-1)}Z\right] + \delta E_t\,(V_{t+1}/\omega_{t+1}, k_{t-1}, Y_t, X_{t+1}^I, X_{t+1}^N)\ ]\ \} \end{aligned}$$

This equation states that the value today is composed of current returns plus the expected discounted value of tomorrow given optimal behavior in the interim. As before, the solution to this functional equation generates an exit rule, and a vertical integration policy function. In addition, it now also generates an investment demand function that can be written as:

$$i_t = i_t\left(\omega_t, k_t, Y_{t-1}\right)$$

This function i_t (.) is determined as part of the Markov Perfect Nash equilibrium and will depend on the parameters determining equilibrium behavior.

7.2 Environmental Extension

Chapter 4 presents some thoughts about the role of the environmental regulation under the Clean Water Act of 1972 in the decision of a paper mill whether or not to vertically integrate into pulp production. For the reasons exposed in Chapter 4, one would expect that mills operating under a stringent environmental regulation would choose not to integrate in order to avoid regulatory costs. An extension of this research would be to develop a version of the dynamic model to explicitly test this hypothesis.

A possible approach would be to model the pulp price perceived by the mill that integrates as a function of the latent liability costs from environmental regulation under the Clean Water Act of 1972 and its amendments, H(.):

$$\pi^I_t = \pi^I_t\,(H_t\,(k_t, \omega_t, S_t))$$

where H(.) is increasing in the mill's paper production capacity, k_t ,decreasing in the mill's unobserved productivity, ω_t, and increasing in the water quality standards set by the state in which the mill operates, S_t,.

Chapter 8

Conclusions

My data show mills of similar characteristics (i.e. size, paper type produced, location, etcetera) appearing to make different choices about their integration status. I have maintained that these differences in choices can be explained by mill level productivity, an unobserved, serially correlated, state variable that is a determinant of both survival probabilities and vertical integration choices. By making the efficiency differences an integral part of the process by which the industry behavior adjusts, the estimation methodology I propose in this dissertation takes care of the biases induced by selection and simultaneity in a reduced form approach. The simulation exercise presented in Chapter 5, shows how the inclusion of the unobservable backed out from the dynamics in the reduced form estimation corrects these biases.

Although the empirical evidence from the estimation of the simplified version of the dynamic model proposed in this dissertation is still far from conclusive with respect to the impact of vertical integration on productivity in the paper industry, it gives a taste of the workings of the estimation methodology proposed. It also suggests that unobserved mill productivity and market concentration are indeed correlated, which would explain the misbehavior of the coefficient on market concentration in the reduced form exercise. Also, my results show that exit behavior can be explained by unobserved mill level productivity. This is yet another proof that a reduced form approach that ignores the role of this unobservable will produce biased estimates.

To delve deeper into the evolution of mill level productivity in the paper industry and its interaction with vertical integration decisions, it is necessary to incorporate the mill's investment decision into the model, as suggested in Chapter 7. That is the first direction in which this research shall be extended.

Appendix – Tables and Figures

Table 2.1 – The U.S. pulp and paper industry 1975-1995

	1975	1980	1985	1990	1995
Number of mills					
Paper only	354	310	332	286	297
Pulp only	21	30	30	28	31
Pulp and paper	230	280	227	222	211
Number of firms	281	265	258	236	235
Percent of single mill firms	68%	64%	67%	70%	67%
Total paper production (tons/day)	143,450	171,425	184,096	215,841	245,019
Total pulp production (tons/day)	118,624	138,526	148,414	171,625	178,211
Total market pulp (tons/day)		18,144	18,630	24,104	24,663
Total paper capacity (tons/day)	170,506	196,400	213,259	225,324	249,958
Total pulp capacity (tons/day)	151,597	170,353	171,833	183,897	199,094
Total market pulp (tons/day)[1]	31,360	26,718	25,855	28,097	35,567
Production/Capacity ratios:					
Paper	84%	87%	86%	96%	98%
Pulp	78%	81%	86%	93%	90%
Average paper mill production (tons/day)	246	291	329	425	482
Average pulp mill production (tons/day)	473	447	577	686	736

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996); US Bureau of the Census, Current Industrial Reports, Pulp, Paper and Board, 1980, 1985

[1] Market pulp capacity is calculated as the difference between pulp and paper capacities for all mills that report any pulp production capacity, after substracting the pulp that goes to close-by paper mills belonging to the same owner.

Table 2.2 – Vertical integration in the U.S. paper industry: 1975-1995

	1975	1980	1985	1990	1995
Number of paper mills	584	590	559	508	508
Number of pulp mills	251	310	257	250	242
Number of pulp and paper mills	230	280	227	222	211
Number of mills reporting capacities for both pulp and paper	230	280	227	222	211
% of paper mills	39%	47%	41%	44%	42%
% of pulp mills	92%	90%	88%	89%	87%
Number of integrated mills[1]	269	311	246	241	233
Integrated paper capacity[1] **(% total paper capacity)**	79%	83%	79%	82%	80%
Integrated pulp capacity[1] **(% total pulp capacity)**	94%	94%	93%	93%	93%

Sources: Post's Pulp and Paper Directory (1976,1981,1986), Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

[1] Including paper mills that are integrated not by in-site pulp production, but by the use of pulp from a close-by mill belonging to the same owner.

Table 2.3 – Vertical integration by region: 1975-1995[1]
(Percent of vertically integrated paper mills)

Region	1975	1980	1985	1990	1995
Northeast	26%	31%	24%	27%	26%
North Central	46%	53%	37%	42%	44%
South	73%	79%	70%	71%	64%
West	57%	60%	58%	61%	60%

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

[1] Including paper mills that are integrated not by in-site pulp production, but by the use of pulp from a close-by mill belonging to the same owner.

Table 2.4 – Paper mills with in-site pulp production: distribution of the pulp to paper capacities ratio (percent of mills in each range)

Year	Average[1] (Pulp/Paper)	(Pulp/Paper) < 0.5	0.5<=(Pulp/Paper) <1	(Pulp/Paper)>=1
1975	1.04	12%	26%	61%
	0.77			
1980	0.95	14%	31%	55%
	0.53			
1985	0.92	15%	40%	45%
	0.50			
1990	0.90	15%	43%	42%
	0.49			
1995	0.95	14%	41%	45%
	0.50			

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

[1] The standard deviation is presented in the second row.

Table 2.5 – Comparison of integrated versus nonintegrated paper mills[1]

	1975		1980		1985		1990		1995	
	Integrated	Only Paper	Integrated	Only Paper	Integrated	Only Paper	Integrated	Only Paper	Integrated	Only Paper
Number of mills	269	316	311	280	246	314	241	268	233	276
Average capacity (tons/day)	449	115	523	120	689	139	771	148	862	178
Percent of total mills:										
Survived to 1995	72%	53%	71%	63%	82%	74%	93%	92%		
Mills belonging to multi-plant firms	79%	60%	80%	64%	79%	64%	76%	62%	76%	66%
Producing more than one paper type	19%	11%	17%	11%	20%	10%	20%	8%	21%	8%
Newsprint	7%	1%	9%	0%	11%	0%	12%	1%	13%	0%
Book and Writing	25%	13%	23%	16%	29%	15%	30%	15%	30%	13%
Wrapping	12%	13%	10%	14%	10%	15%	11%	14%	11%	12%
Board	44%	43%	42%	44%	42%	48%	41%	52%	40%	53%
Sanitary	11%	12%	10%	10%	11%	10%	11%	5%	10%	7%
Industrial	9%	30%	15%	28%	11%	23%	9%	21%	6%	23%

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991,1996)

Table 2.6 – Distribution of paper capacity by relative size of mill capacity: 1975-1995

Mill size (% of total paper capacity)	Percent of industry capacity					Number of mills				
	1975	1980	1985	1990	1995	1975	1980	1985	1990	1995
Integrated mills										
Under 0.1	4.2%	5.3%	3.6%	3.3%	2.5%	76	101	60	58	52
0.1 to 0.2	7.9%	9.3%	7.3%	6.1%	5.7%	54	67	49	41	41
0.2 to 0.3	10.5%	11.1%	8.9%	9.7%	9.5%	42	45	36	38	39
0.3 to 0.4	9.0%	9.0%	8.9%	8.8%	7.2%	26	26	25	25	21
0.4 to 0.5	10.4%	7.9%	6.6%	6.8%	7.6%	23	18	15	15	17
0.5 to 0.75	14.4%	19.8%	22.6%	22.5%	19.9%	24	33	38	38	34
0.75 to 1.0	16.2%	10.2%	13.4%	13.9%	15.2%	19	12	16	16	18
1.0 to 1.5	4.4%	8.6%	6.5%	11.5%	12.7%	4	8	6	10	11
Above 1.5	1.7%	1.6%	1.6%	_	_	1	1	1	_	_
Nonintegrated mills										
Under 0.1	9.8%	8.1%	10.2%	9.2%	8.6%	255	227	255	221	213
0.1 to 0.2	6.8%	5.5%	6.3%	5.7%	6.1%	47	42	45	38	45
0.2 to 0.3	2.0%	1.6%	2.6%	1.7%	2.2%	8	7	11	7	10
0.3 to 0.4	0.6%	0.3%	0.3%	0.3%	2.3%	2	1	1	1	7
0.4 to 0.5	1.0%	_	0.5%	_	_	2	_	1	_	_
0.5 to 0.75	1.1%	1.6%	0.6%	0.6%	0.5%	2	3	1	1	1
0.75 to 1.0	_	_	_	_	_	_	_	_	_	_
1.0 to 1.5	_	_	_	_	_	_	_	_	_	_
Above 1.5	_	_	_	_	_	_	_	_	_	_

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991,1996)

Table 2.7 – Entry and exit of integrated mills versus specialized mills each period

Period	Total new	Integrated	Paper only	Pulp only
1975-1980	96	42	42	12
1980-1985	32	11	17	4
1985-1990	47	23	22	2
1990-1995	34	9	19	6
Period	**Total exit**	**Integrated**	**Paper only**	**Pulp only**
1975-1980	64	18	45	1
1980-1985	77	35	39	3
1985-1990	94	32	57	5
1990-1995	40	16	21	3

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

Table 2.8 – Integrated mills: entrants and established mills each decade, by region

Period and Region	Integrated entrants (% entering paper mills)	Integrated established mills (% mills operating entire period)
1975-1980		
U.S.	50.0%	54.0%
Northeast	8.0%	37.0%
North Central	52.0%	54.0%
South	63.0%	76.0%
West	42.0%	52.0%
1980-1985		
U.S.	39.0%	46.0%
Northeast	18.0%	28.0%
North Central	43.0%	38.0%
South	45.0%	72.0%
West	33.0%	54.0%
1985-1990		
U.S.	51.0%	47.0%
Northeast	23.0%	26.0%
North Central	60.0%	40.0%
South	69.0%	72.0%
West	38.0%	56.0%
1990-1995		
U.S.	32.0%	47.0%
Northeast	25.0%	28.0%
North Central	33.0%	44.0%
South	20.0%	69.0%
West	50.0%	56.0%

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

Table 2.9 – Paper mills survival

Type	Number of mills listed in 1975	Number of mills that continued to be listed in: 1980	1985	1990	1995
Integrated	230	224	183	159	150
Survival Rate[1]		97%	80%	69%	65%
Paper Only	354	215	199	155	140
Survival Rate[1]		61%	56%	44%	40%

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

[1] Reported as a percentage of the original number of firms listed in 1975.

Table 2.10 – Paper mills switching between integration and specialization

Period	Became integrated	Disintegrated
1975-1980	27	8
1980-1985	3	27
1985-1990	8	5
1990-1995	6	9

Note: Seventeen of the mills that integrated in the 1st period disintegrated in the 2nd period, one of the mills that integrated in the 1st period disintegrated in the 3rd period, and another disintegrated in the 4th period. Two of the mills that integrated in the 2nd period disintegrated again, one in the 3rd period and the other in the 4th. Two mills integrated in the 3rd period and returned to disintegration in the 4th period. Four mills that disintegrated in the 1st period integrated back, two in the 2nd period, one in the 3rd, and one in the 4th. One mill that disintegrated in the 2nd period integrated back in the following period. Finally, two mills that disintegrated in the 2nd period, integrated in the 3rd period and disintegrated again in the 4th period.

Table 2.11 – Concentration in the pulp and paper industry, 1975-1995, by region

Year and Region	Herfindahl Indices for the paper market	Herfindahl Indices for the pulp market
1975		
Northeast	0.012	0.149
North Central	0.012	0.115
South	0.014	0.050
West	0.031	0.077
1980		
Northeast	0.014	0.233
North Central	0.011	0.089
South	0.013	0.043
West	0.032	0.062
1985		
Northeast	0.016	0.274
North Central	0.012	0.135
South	0.013	0.042
West	0.029	0.063
1990		
Northeast	0.018	0.240
North Central	0.013	0.201
South	0.013	0.049
West	0.032	0.066
1995		
Northeast	0.017	0.149
North Central	0.014	0.141
South	0.013	0.036
West	0.034	0.071

Sources: Post's Pulp and Paper Directory (1976,1981,1986); Lockwood-Post's Directory of the Paper and Allied Trades (1991, 1996).

Table 3.1 – Probit regression of vertical integration in U.S. paper industry: 1975-1995

Dependent variable: VI

	1975 Probit Cofficients	dF/dx	1980 Probit Cofficients	dF/dx	1985 Probit Cofficients	dF/dx	1990 Probit Cofficients	dF/dx	1995 Probit Cofficients	1995 dF/dx
MODEL I										
Constant	-3.133	_	-2.964	_	-4.218	_	-3.667	_	-3.319	_
t-statistic	-8.425		-9.089		-9.474		-9.025		-7.991	
SIZE	0.722	0.288	0.654	0.258	0.873	0.343	0.734	0.293	0.719	0.287
t-statistic	10.484		10.812		11.242		10.530		10.106	
NEWSPRINT DUMMY*	0.173	0.069	1.330	0.386	1.446	0.483	0.916	0.331	1.344	0.434
t-statistic	0.413		2.061		2.103		2.042		2.159	
FORESTLAND DUMMY*	0.476	0.188	0.379	0.147	0.241	0.095	0.334	0.133	0.086	0.034
t-statistic	3.264		2.756		1.603		2.277		0.548	
CONCENTRATION MEASURE I	-1.159	-0.463	-0.703	-0.277	-0.972	-0.382	-3.667	-0.278	-1.131	-0.451
t-statistic	-5.051		-3.418		-3.997		-9.025		-4.592	
Number of observations	505		568		549		500		471	
Log likelihood function	-227.27		-272.91		-220.26		-220.05		-204.32	
LR (chi^2 = 4)	245.09		240.21		312.84		251.90		242.98	
MODEL II										
Constant	-3.625		-3.379		-4.937		-4.141		-4.050	
t-statistic	-11.365		-11.164		-12.054		-11.339		-10.818	
SIZE	0.731	0.291	0.644	0.253	0.862	0.340	0.723	0.289	0.692	
t-statistic	10.689		10.631		11.176		10.434		9.917	
NEWSPRINT DUMMY*	0.061	0.024	1.288	0.376	1.435	0.476	0.942	0.337	1.496	
t-statistic	0.153		2.052		2.243		2.167		2.539	
FORESTLAND DUMMY*	0.303	0.120	0.285	0.111	0.123	0.049	0.181	0.072	0.005	
t-statistic	1.832		1.754		0.721		1.116		0.031	
CONCENTRATION MEASURE II	0.002	0.001	0.0027	0.0011	0.004	0.002	0.003	0.001	0.003	
t-statistic	2.049		2.000		3.046		2.535		2.988	
Number of observations	505		568		549		500		471	
Log likelihood function	-238.40		-276.63		-223.16		-221.11		-210.13	
LR (chi^2 = 4)	222.84		232.77		307.05		249.77		231.36	

* dF/dx is for discrete change of dummy variable from 0 to 1

Table 3.2 – Panel analysis

	Pooled probit	Random effects probit	Pooled probit	Random effects probit	Pooled probit	Random effects probit	Pooled probit	Random effects probit
	Concentration Measure I		Concentration Measure II		Concentration Measure III		Concentration Measure III	
Data	All	All	All	All	All	All	All	All
Dependent variable	VI	VI	VI	VI	VI	VI	VI	VI
Constant	-3,302	-5,721	-3,977	-6,718	-3,934	-6,582	-3,270	-5,566
t-statistic	-19,312	-10,517	-25,53	-12,678	-25,166	-12,758	-12,494	-9,269
SIZE	0,713	1,226	0,706	1,189	0,698	1,160	0,697	1,193
t-statistic	23,661	11,906	23,763	11,814	23,415	11,893	23,066	12,426
NEWS	0,944	1,279	0,957	1,209	0,953	1,170	1,029	1,335
t-statistic	4,214	2,052	4,513	2,291	4,469	2,420	4,672	3,420
SW	0,294	0,766	0,259	0,922	0,186	0,807	0,087	0,548
t-statistic	4,549	3,704	3,697	4,221	4,469	3,605	1,154	3,420
Concentration Measure	-0,980	-1,702	0,0002	0,0001	0,003	0,004	0,002	0,003
t-statistic	-9,803	-7,448	4,529	1,598	5,559	3,150	4,500	2,226
Lagged Concentration Mea	1,31E-08	2,08E-08	1,54E-08	2,37E-08	1,59E-08	2,32E-08		
t-statistic	1,914	2,081	2,077	2,205	2,149	2,172		
Ruralness Measure							0,925	1,636
t-statistic							4,448	3,326
Personal Income (per cápita)							-0,0002	-0,0003
t-statistic							-4,331	-4,224
Number of groups		643		643		643		643
Number of observations	2593	2593	2593	2593	2593	2593	2593	2593
Log likelihood function	-1164,36	-890,68	-1203,16	-921,81	-1195,62	-918,58	-1157,63	-897,64
chi^2	1262,01	210,11	1184,4	195,24	1195,41	202,04	1275,47	243,84
Pseudo R^2	0,35		0,33		0,33		0,36	

Table 3.3 – Panel analysis, change data

	Pooled logit	**Fixed effects logit**	**Pooled logit**	**Fixed effects logit**
Data	Switchers who integrated Mills that did not change status	Switchers who integrated Mills that did not change status	Switchers who disintegrated Mills that did not change status	Switchers who disintegrated Mills that did not change status
Dependent variable	dVI	dVI	dVI	dVI
Constant	0.081	0.046	-0.030	-0.058
t-statistic	5.991	1.865	-2.013	-2.026
Change in Concentra	-0.216	0.001	-0.064	-0.039
t-statistic	-1.095	0.049	-2.978	-1.202
Lagged Concentratio	0.019	0.045	-0.032	-0.007
t-statistic	1.129	1.160	-1.724	-0.158
Dummy Variable for	-0.075	-0.046	-0.047	-0.025
t-statistic	-5.585	-3.497	-3.203	-1.657
Dummy Variable for	-0.059	-0.042	0.027	0.038
t-statistic	-4.370	-3.047	1.824	2.349
Dummy Variable for	-0.064	-0.038	0.016	0.031
t-statistic	-4.790	-2.726	1.061	1.886
Number of groups		640		634
Number of observati	1849	1849	1868	1868
F	F(5, 1843) = 8,86	F(5, 1204) = 3,51	F(5,1862) = 8,83	F(5,1229) = 5,50
R squared	0.02		0.02	

Table 5.1 – Parameter values used in simulation

Average Cost Function		**Markov Process governing omega**	
A^N	0.0040	γ_0	3.2789
A^I	0.0035	γ_1	0.1343
η^1	0.8160	μ	4.5952
η^2	0.9274	σ^2	10.7774
α	0.5101		
λ_0	15.4943	**Cost to integrate - F**	39.8800
λ_1	1.2000		
λ_2	0.5000	**Exit fee - Φ**	1636.2995
λ_3	0.7000		
λ_4	15.0000		

Table 5.2 – Summary statistics

Variable		Mean	Standard deviation	Min	Max
VI	overall	0.936	0.245	0.000	1.000
	between		0.186	0.000	1.000
	within		0.197	0.136	1.736
ω	overall	6.225	2.298	2.274	16.008
	between		1.310	2.525	10.517
	within		1.986	0.688	13.276
Wage	overall	3.310	0.552	2.501	5.262
	between		0.351	2.501	4.555
	within		0.437	1.839	5.174
Concentration Measure I (HH index of the 200-mile diameter pulp market)	overall	0.299	0.230	0.000	0.993
	between		0.150	0.054	0.962
	within		0.196	-0.238	1.047
Forestland Dummy (SW)	overall	0.151	0.358	0.000	1.000
	between		0.345	0.000	1.000
	within		0.000	0.151	0.151
Asset Specificity Measure	overall	0.384	0.253	0.000	1.000
	between		0.187	0.022	0.925
	within		0.179	-0.210	1.020
Number of Observations	1929				
Number of groups	500				

Table 5.3 – Integration and exit behavior in the artificial data

Period	Integrated mills	Disintegrated mills	All active mills
1	438	62	500
2	417	20	437
3	364	16	380
4	314	14	328
5	272	12	284
Period	**Mills that exited**	**Mills that integrated**	**Mills that disintegrated**
1 a 2	63	40	8
2 a 3	57	13	11
3 a 4	52	12	12
4 a 5	44	9	11
Mills active over the whole period	284		
Number of mills	500		
Number of Observations	2145		

Table 5.4 – Reduced form model using simulated data

	Pooled probit	**Pooled probit**	**Random effects probit**	**Random effects probit**	**Random effects logit**	**Random effects logit**	**Fixed effects logit**	**Fixed effects logit**
Data	Simulated data	Simulated data	Simulated data	Simulated data	Simulated data	Simulated data	Simulated data	Simulated data
Dependent variable	VI	VI	VI	VI	VI	VI	VI	VI
Constant	1.503	-7.336	1.736	-12.516	3.130	-22.829		
t-statistic	13.025	-10.490	10.941	-3.912	9.834	-3.610		
Concentration Measure I	-0.514	1.869	-0.609	2.963	-1.184	5.354	-1.640	5.902
t-statistic	-2.666	5.013	-2.751	3.332	-2.753	3.172	-3.053	1.799
Forestland Dummy	0.237	0.739	0.290	1.446	0.562	2.627		
t-statistic	1.642	2.778	1.564	2.300	1.492	2.184		
Asset Specificity Measure	0.426	1.219	0.455	1.826	0.936	3.307	0.628	30.373
t-statistic	2.235	3.662	2.034	2.731	2.083	2.622	1.071	1.088
Unobserved Productivity Index		1.993		3.450		6.310		8.127
t-statistic		11.685		3.915		3.622		2.533
Number of groups			500	500	500	500	91	91
Number of observations	1929	1929	1929	1929	1929	1929	406	406
Log likelihood function	-451.00	-176.85	-444.60	-170.37	-444.65	-170.31	-137.25	-6.01
LR (chi^2)	18.48	566.78					12.45	275.27
Wald (chi^2)			16.9	15.75	17.13	13.58		
Seudo R squared	0.02	0.62						

Figure 6.1 – Exit and integration cut-off values of ω

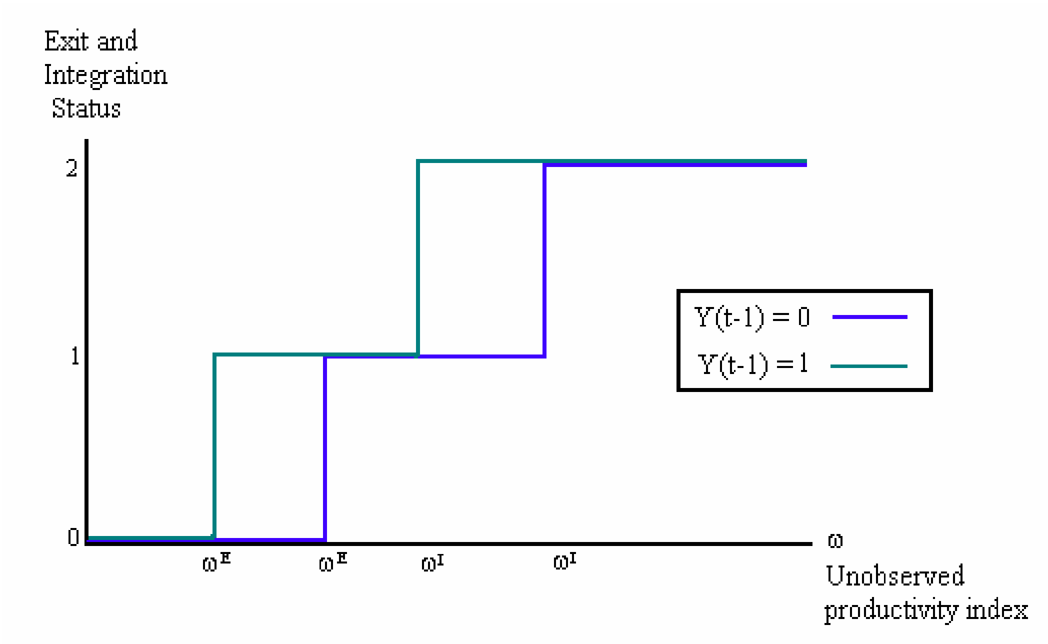

Table 6.1 – Example

Time	**Status of Mill i**	**Range for ω draw**	**Mean $\bar{\omega}$**
(t-1)	0		
T	0	$\omega_N^E < \omega_{it} \le \omega_N^I$	ω
(t+1)	1	$\omega_{i(t+1)} > \omega_N^I$	$\bar{\omega}_{i(t+1)} = \gamma_0 + \gamma_1 \omega_{it}$
(t+2)	0	$\omega_I^E < \omega_{i(t+2)} \le \omega_I^I$	$\bar{\omega}_{i(t+2)} = \gamma_0 + \gamma_1 \omega_{i(t+1)}$
(t+3)	2	$\omega_{i(t+3)} \le \omega_N^E$	$\bar{\omega}_{i(t+3)} = \gamma_0 + \gamma_1 \omega_{i(t+2)}$

Figure 6.2 – Drawing sequences of ω's from a truncated normal distribution

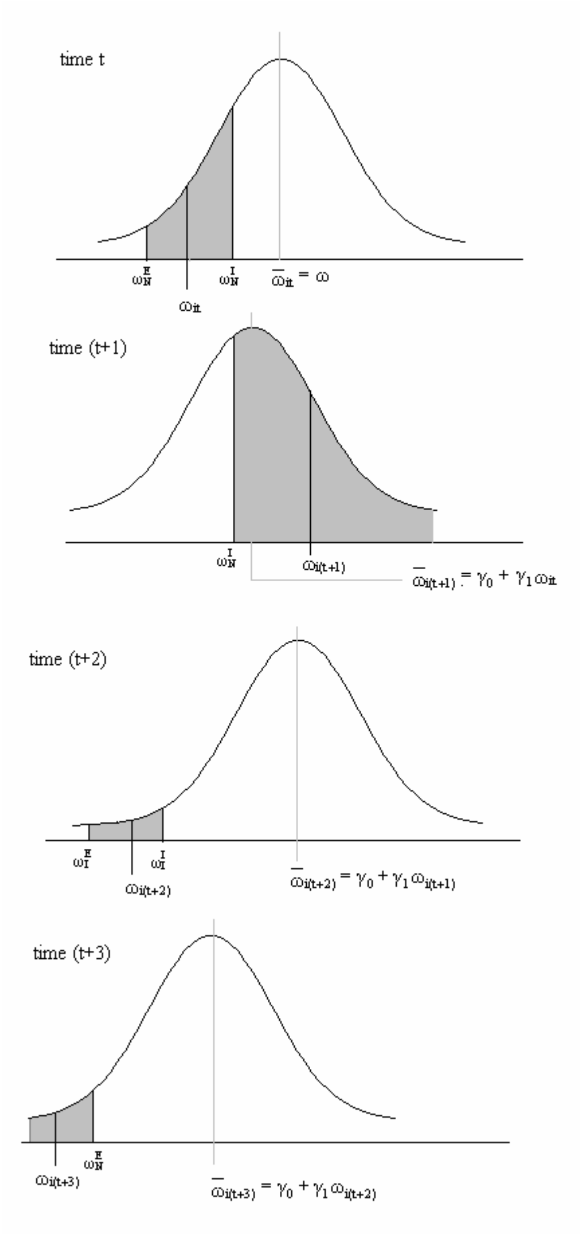

Table 6.2 – Parameter estimates
(t-statistics in parenthesis)

Average Cost Function		**Markov Process governing omega**	
A^N	0.0040	γ_0	3.2789
	(108.806)		(13.875)
A^I	0.0035	γ_1	0.1343
	(104.626)		(2.552)
η^1	0.8160	μ	4.5952
	(350.6601)		(28.421)
η^2	0.9274	σ^2	10.7774
	(366.680)		(1.204)
α	0.5101		
	(4.931)		
λ_0	15.4943	**Cost to integrate - F**	39.8800
	(316.455)		(4.912005)
λ_1	0.0021		
	(0.393)	**Exit fee - Φ**	1636.2995
			(7.063316)

Figure 6.2 – Estimated f(ω)

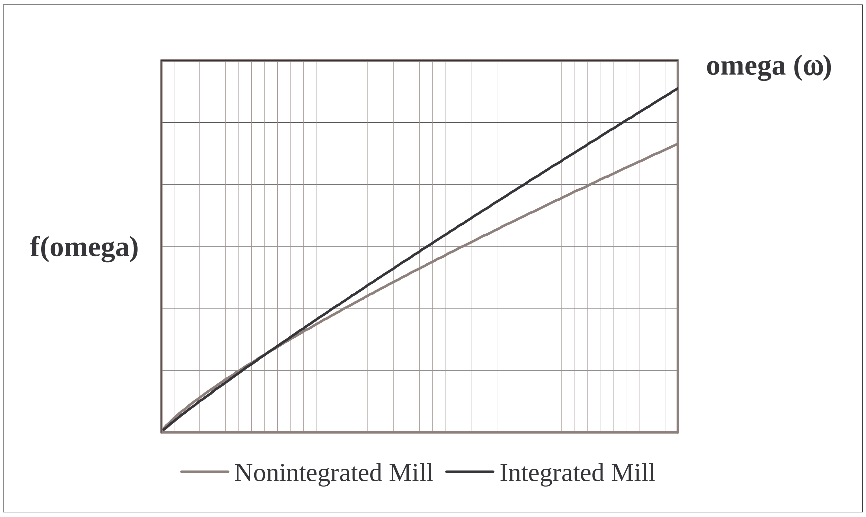

Table 6.3 – Mill status and productivity

	Pooled regression	**Random effects regression**	**Fixed effects regression**
Data	All active mills, estimated unobserved productivity	All active mills, estimated unobserved productivity	All active mills, estimated unobserved productivity
Dependent variable:			
Unobserved Productivity Index	ω	ω	ω
Constant	3.896	3.833	3.486
t-statistic	349.539	267.496	202.839
Integration Status Dummy	2.600	2.749	3.452
t-statistic	161.785	139.002	103.733
Number of groups		642	642
Number of observations	1949	1949	1949
F(1,1947)	26174.29		
F(1,1306)			10760.5
Wald $Chi^2(1)$		19321.57	
R^2		0.93	0.93
Adjusted R^2	0.93		

Table 6.4 – Models of exit probabilities

	Probit Model	**Logit Model**
Intercept	-15.556	-27.016
(t-statistic)	(-8.320)	(-7.273)
Productivity Change	-26.572	-46.035
(t-statistic)	(-8.450)	(-7.403)
# Observations	1479	1479
Log Likelihood	-55.2	-56.0

Table 6.5 – Partial correlations

Partial Correlation of unobserved mill productivity (ω) with:

Variable:	**Correlation**
Size (log of capacity)	0.464
Concentration Measure I (mill specific HH index for the 200-miles diameter market)	-0.121
Lagged Concentration Measure I (mill specific HH index for the 200-miles diameter market)	-0.077
Concentration Measure I (market-pulp per mill, in the 200-miles diameter market)	-0.029

Table 6.6 – Industry productivity growth rates

Period	Industry Level Average Productivity Change	Integrated Mills Average Productivity Change	Disintegrated Mills Average Productivity Change	Average size of mills with positive productivity growth (tons/day)	Average size of mills with negative productivity growth (tons/day)
1980-1985	1.97%	2.44%	7.59%	177.33	448.93
1985-1990	-5.80%	-7.09%	-2.83%	428.94	490.69
1990-1995	6.59%	6.90%	5.57%	509.63	520.34

Bibliography

Baltagi, Badi. *Econometric Analysis of Panel Data*. New York: Wiley, 1995.

Caves, Richard E., and Ralph M. Bradburd. "The Empirical Determinants of Vertical Integration." *Journal of Economic Behavior and Organization* 9 (April 1988): 265-79.

Clerides, Sofronis K., Saul Lach and James R..Tybout. "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco*." Quarterly Journal of Economics* (1998): 903-947.

Coase, Ronald H., "The Nature of the Firm", *Economica* 4 (Nov,1937):386-405, reprinted in G.J. Stigler and K.E. Boulding (eds.), *Readings in Price Theory* (Homewood, III. Richard D. Irwin, 1952).

Globerman, Steven, and Richard Schwindt. "The Organization of Vertical Related Transactions in the Canadian Forest Products Industry*." Journal of Economic Behavior and Organization* 7 (June 1986): 199-212

Greene, William H. *Econometric Analysis*. New York: Macmillan Publishing Company, 1990.

Greenstone, Michael. "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures", Working Paper #408, Princeton University

Hennart, Jean-Francois. "Upstream Vertical Integration in the Aluminum and Tin Industries." *Journal of Economic Behavior and Organization* 9 (April 1988): 281-99.

Johnston, J. and J. Dinardo. *Econometric Methods*, McGraw-Hill Companies, Inc., 1997.

Levy, David T. "The Transactions Cost Approach to Vertical Integration: An Empirical Examination." *Review of Economics and Statistics* 67 (August 1985): 438-45

Maddala, G.S. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press, 1983.

Maggat, Wesley A., Alan Krupnik, and Winston Harrington. *Rules in the Making. A Statistical Analysis of Regulatory Agency Behavior*. Washington D.C., Resources for the Future.

Miller Freeman Publications. *Post's Pulp and Paper Directory*. San Francisco, 1976, 1981, 1986.

__________. *Lockwood-Post's Directory of the Pulp Paper and Allied Trades*. San Francisco, 1991,1996.

Ohanian, Nancy Kane. *The American Pulp and Paper Industry, 1900-1940: Mill Survival, Firm Structure, and Industry Relocation*. Contributions in Economics and Economic History, no.140. Westport, C.T.: Greenwood Press, 1993.

__________. "Vertical Integration in the U.S. Pulp and Paper Industry, 1900-1940." *Review of Economics and Statistics* 76 (February 1994): 202-207.

Olley, G.S. and A. Pakes. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64(6) (November, 1996): 1263-1297.

Perry, M.K. "Vertical Integration: Determinants and Effects." In R. Schmalensee and R. Willig, editors, *Handbook of Industrial Organization*. Amsterdam, The Netherlands: North-Holland, 1989

Powell, J. "Symmetrically Trimmed Least Squares Estimation for Tobit Models," *Econometrica* 54 (1986): 1435-1460.

Spiller, P.T. "On vertical mergers." *Journal of Law, Economics, and Organizations* 1 (1985): 285-312.

Ringleb Al H. and Steven N. Wiggins. "Liability and Large-Scale, Long-Term Hazards", *Journal of Political Economy* 98(31) (1990): 574-595

Stigler, George J. "The Economies of Scale." *Journal of Law and Economics* 1 (October 1958): 54-71. Reprinted in *The Organization of Industry*. Chicago: The University of Chicago Press, 1968.

Temin, Peter. "Product Quality and Vertical Integration in the Early Cotton Textile Industry." *Journal of Economic History* 68 (December 1988): 891-907.

United States Department of Commerce, Bureau of the Census. Current Industrial Reports. Pulp, Paper and Board, 1980,1985.

Williamson, Oliver E. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press, 1975.

__________. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: The Free Press, 1985.