The Design of Industry*

by

Sharon Novak
Massachusetts Institute of Technology
Sloan School of Management
snovak@mit.edu

and

Birger Wernerfelt
Massachusetts Institute of Technology
Sloan School of Management
bwerner@mit.edu

April 21, 2006.

JEL Codes: D23, L22, L62

Keywords: Make-or-buy, Simulated GMM, Adjustment-costs

*We are indebted to seminar participants at MIT and IIOC, as well as Ken Corts, Mercedes Delgado Garcia, Bob Gibbons, Whitney Newey, Jim Orlin, Roberto Rigobon, Duncan Simester, Eric Van den Steen, Tom Stoker, Dan Whitney, Dennis Yao, Florian Zettelmeyer, and Ezra Zuckerman for helpful discussions, and to Dmitri Byzalov, Chris Ruggerio, and especially Carol Meyers for excellent research assistance. Of course, the usual disclaimer applies.

The Design of Industry

Abstract

We study the division of labor within an industry, formulate it as a generalized make-or-buy problem, and characterize the optimal allocation of work as that minimizing the sum of adjustment-costs within and between firms. Using a unique dataset on eight segments of the global automobile industry, we test the theory from several angles: We first show that any two tasks are more likely to be performed by the same firm if mutual adjustments between them are needed on a sufficiently frequent basis. To take indirect effects into account, we then look at the entire industry and find that a disproportionate number of adjustments are managed inside firms. We finally use simulated GMM to estimate a structural model in which industry design is portrayed as the solution to an integer program aimed at minimizing industry-wide adjustment-costs. The program is extremely complex and while there is a significant heterogeneity in the eight actual designs, they all fit the model very well. The main substantive contribution of the paper is thus to present robust evidence consistent with the view that the firm is a low variable-, but high fixed cost way to govern adjustments. A more methodological contribution is to introduce industry-level estimates and show that they outperform the firm-level estimates used in other studies of make-or-buy decisions.

I. Introduction

The sub-tasks of a production process can be allocated between firms in many ways: each could be undertaken by a separate firm, the same firm could perform all of them, or we could have any one of a very large number of intermediate possibilities. This grouping of tasks, which we will call the design of the industry, can be seen as an implication of the theory of the firm. We will here use it to test the adjustment-cost theory, according to which the firm governs adjustments in a high variable-, but low fixed cost way (Wernerfelt, 1997). Specifically, we use a unique dataset from the automobile industry to test: (1) that any two tasks are more likely to be performed by the same firm if mutual adjustments are needed on a sufficiently frequent basis, (2) that a disproportionate share of all adjustments are managed inside firms, and (3) that industry design can be portrayed as the solution to an integer program aimed at minimizing the sum of adjustment-costs within and between firms. Consistent with the complexity of the program, the patterns of integration observed in the data differ widely. In spite of this, our results are very strong and robust, suggesting that the theory captures a significant and important effect.

The structural specification has no precedence in the literature, and a secondary contribution of the paper is to introduce it as a new way to test theories of the firm. By testing at the industry-level, we take advantage of information contained in the entire division of labor within the industry. Past tests have been conducted at the firm-level in the context of several parallel "make-or-buy" questions with reference to a specific firm (Masten, 1984; Novak and Eppinger, 2001; Simester and Knez, 2002). However, since most firms use several inputs, a lot of information is lost by ignoring the extent to which the out-sourced parts are co-produced. The amount of information lost depends on how many parts are out-sourced. Suppose that there are N parts and that a fraction p of all $N(N-1)/2$ pairs of them are co-produced. On the average, this allows us to ask $N-1$ make-or-buy questions with reference to the (firm producing the) first part, $(N-1)(1-p)$ co-production questions with reference to the first out-sourced part, etc. etc., for a total of approximately $(N-1)/p$ "degrees of freedom." In our data, p is typically around .25, suggesting that our tests are based on about four times as much information as that used in a firm-level estimation. Consistent with this, we find that our industry-level model outperforms both GMM and logit models estimated at the firm-level.

There are also qualitative differences between the two levels of testing in the sense that they may evaluate particular industry designs quite differently. For example, at the industry-level, there is not that big a difference between "Toyota in-sourcing the interior" and "GM out-sourcing the interior," as long as the same firm makes all the parts of the interior. In contrast, a firm-level make-or-buy analysis portrays these two cases as maximally different. We will here test only the adjustment-cost theory, but our procedure could be applied to any alternative theory.

While our test is more powerful, data requirements and computational difficulties are obstacles to its use. For each pair of parts, we need data on the frequency with which mutual adjustments are needed; and for each part, we need to know its producer. We are lucky to have a very good data set, but depending on the independent variables used, it can be very time-consuming to collect appropriate data. The second problem concerns to estimation procedure. Even for small sets of parts, there are a very, very large number of possible industry designs, each of which needs to be considered as a counterfactual. More importantly, it is not possible to reduce the dimensionality by looking at each pair by itself: if parts 1 and 2 are co-produced, and parts 2 and 3 are co-produced, one can not ask if parts 1 and 3 should be co-produced. Since these dependencies cause extreme non-linearities, we use simulated GMM to estimate the model. A very large integer programming problem has to be solved in each simulation, and we believe that our 36 parts are close to the limit of what is currently practical. On the other hand, the efforts at data collection and computation have given us a very well-fitting structural model.

After a brief review of the theoretical and empirical literature in Section II, we derive the hypotheses in Section III. The data are described in Section IV, and estimation techniques and results are presented in Section V. In Section VI, we briefly compare estimates found at the industry- and firm-level, and the paper concludes with a discussion in Section VII.

II. Literature

In spite of its central importance to the field, economists have not yet agreed on a theory of the firm. The several competing theories include, but are not limited to, those of Grossman and Hart (1986), Holmstrom and Milgrom (1994), Wernerfelt (1997), and Williamson (1979). A significant amount of empirical work has failed to settle the issue, possibly because many of the theories are difficult to conclusively falsify with available data. As a result, many economists evaluate the alternative theories on essentially a priori grounds. We have no illusions that the present paper can change this in a major way, but we hope to contribute to the thinking in a very important area.

The adjustment-cost theory of the firm (Wernerfelt, 1997) compares alternative game forms in light of the costs of change (adjustment).[1] In its simplest form, the theory looks at a trading relationship between two players in which maximization of joint payoffs occasionally requires that the actions of one or both players change. If trade is sufficiently frequent, we can assume that efficient adjustments will be implemented, and the question reduces to one of finding the cheapest gameform with which the players can agree on the adjustments and associated changes in payments. The choice of gameform is non-contractible, but is sustained by an implicit contract. There are a sea of possibilities, but it is illuminating to compare three specific alternatives: "Negotiation-as-needed" in which the players have a bilateral discussion about each adjustment and associated changes, a "Pricelist" in which the consequences of all possible adjustments are agreed upon ex ante, and an "Employment Relationship" in which one player ex ante agrees to let the other dictate what happens in case of adjustment, while both retain the right to terminate the relationship. Purchases governed through these three gameforms could be a house renovation, a beauty treatment, and secretarial services. The claim is now that there are some, perhaps very small, costs negotiation. Since the employment relationship implements individual adjustments without negotiation, no other gameform can have lower variable costs. On the other hand, employment entails some fixed costs ex ante because the parties have to reach agreement on the arrangement. There are no such fixed costs with negotiation-as-needed because no ex ante agreement is needed to open negotiation ex post. So among these three gameforms, "Negotiation-as-needed" is most efficient when adjustments are rare, a "Pricelist" is best when the set of possible adjustments is small, and the "Employment Relationship" is the lowest cost solution when adjustments are both frequent and diverse. Among all possible gameforms, none can implement adjustments at lower variable (per-adjustment) cost than the Employment Relationship, implying that it is asymptotically efficient under the stated assumptions. Different implications of the theory have been tested by Simester and Knez (2002) and Wernerfelt (1997), but the present paper is the most extensive test by far.

As noted in the Introduction, another contribution of our paper is to show how a theory of the firm can be tested by looking at the allocation of tasks within an entire supply chain. While the study of supply chains is unusual in economics, it has a long history in operations management, primarily in the context of the "Design Structure Matrix" (Steward, 1981; Eppinger, 1991; and Baldwin and Clark, 2000). This matrix summarizes the direction and importance of information flows between pairs of tasks, and is a tool for managing new product development

[1] Bajari and Tadelis (2001) present a related argument.

processes. It contains more and different information than that used here, but cases with "important bi-directional information flows" are suggestive of the frequent mutual adjustments that we here focus on - although "important" is different from "frequent". The question of firm boundaries receives little attention in this literature, but it is interesting to note that Baldwin and Clark (2000, p. 368ff.) adopt arguments similar to those tested here and suggest that firm coordination carries low variable costs compared to market coordination.

III. Theory and Hypotheses

There is no new theory in the paper, but since we ultimately will estimate a structural model, we need to be rather explicit in describing how the adjustment-cost theory applies to our data. The product consists of N parts, and for each of the $N(N-1)/2$ pairs i and j we know the frequency with which there has to be a coordinated adjustment in them. We denote this variable by a_{ij}. Since the parts are complex in themselves, there is no automatic transitivity in the need for adjustment. That is, it is possible that $a_{12} > 0$, $a_{23} > 0$, and $a_{13} = 0$. The first two sets of adjustments could be necessary without the third being so (as if the first is about color and the second is about size).

The N parts can be made by anywhere from one to N firms and we define an industry design as an allocation $S = \{S_1, S_2, \ldots, S_n\}$ of all N parts into $n \leq N$ non-empty, non-overlapping sets. We take the frequency of needs for adjustments as exogenously given and look for the most efficient design of the industry.[2] To

[2] While we thus treat technology, represented by the frequency of needs for mutual adjustments, as exogenously given, it is obviously endogenous in the very long run. The theory implies that

economize on notation, we denote the set of N parts by N. We use i and j as generic parts produced in firms I and J, respectively, while b and c are generic firms. The variable adjustment-costs, per adjustment per pair, are m in the market and $f(s_b) < m$ in a firm of size $s_b \equiv | S_b |$, where the f's are non-decreasing. Total fixed adjustment-costs are $w(s-1)$ in a firm of size s (and 0 in the market). To keep things simple, we assume that all adjustments are implemented, regardless of the costs thereof. Recalling that I is the set of parts made by the same firm as part i, total expected costs are therefore $(N-n)w + \frac{1}{2}\Sigma_{i \in \mathrm{N}}[f(s_I)\Sigma_{j \in I}\, a_{ij} + m\Sigma_{j \notin I}\, a_{ij}]$, and the cost-minimizing design solves the Partitioning Problem (P):

$$\mathrm{Min}_n \mathrm{Min}_S\ (N-n)w + \tfrac{1}{2}\Sigma_{i \in \mathrm{N}}[f(s_I)\Sigma_{j \in I}\, a_{ij} + m\Sigma_{j \notin I}\, a_{ij}], \text{ s.t.}$$

$$\bigcup_{b=1}^{n} S_b = \mathrm{N},\ \bigcap_{b,c \neq b} S_b\, S_c = \emptyset, \text{ and } S_b \neq \emptyset \text{ for all } b. \qquad (P)$$

There are a finite number of possible designs, so while it is hard to characterize solutions, they trivially exist.

To gain some intuition about solutions, it is useful to start thinking about truly modular cases in which the symmetric NxN matrix $\{a_{ij}\}$ is block-diagonal with n blocks. It is natural to propose a design with n firms, corresponding to the n blocks, with expected costs $(N-n)w + \frac{1}{2}\Sigma_{i \in N} f(s_I)\Sigma_{j \in I}\, a_{ij}$. Alternatives with fewer firms are dominated because they will add w per firm eliminated (employee added) and offer no benefits. Alternatives with more firms will save fixed costs of w per firm added (employee eliminated), but add extra costs of market adjustments in the

firms can influence the industry design by investing in more or less modular technologies. It is likely that the incentives to do this will depend on the number of firms competing, as well as economies of scale in each module.

form $\frac{1}{2}\Sigma_{i\in \mathrm{N}}[m-f(s_I)]\Sigma_{j\notin I}\, a_{ij}$. So if the marginal costs of market over firm adjustment $[m-f(s_I)]$ are sufficiently large relative to the fixed costs of firm adjustment (w), the proposed design will be optimal.

Beyond the block-diagonal case, things are much more complicated. If we add inter-block entries to $\{a_{ij}\}$, designs with fewer firms offer the possibility of saving $[m-f(s_I)]$ on inter-block adjustments. Similarly, if we eliminate some intra-block entries, designs with more firms become more attractive because there will be fewer additional market adjustment costs. More firms always imply lower fixed costs, while a "merger" of two firms implies lower variable costs. However, the variable cost savings are weighted by the elements of $\{a_{ij}\}$. This means that it may be very complicated to identify the design that minimizes the expected sum of variable costs - even if we keep the number of firms constant. For example, given firms of size $s_1, s_2, \ldots s_n$, parts can be assigned to firms in $N!/(s_1!\, s_2! \ldots s_n!)$ possible ways. There are $\Sigma\Sigma\, s_b\, s_c/2 \equiv t$ designs in which one pair is switched and $t(t-1)^{p-1}/p!$ designs in which p pairs are switched.

To understand the first-order effects of increasing needs for adjustments, we consider a candidate solution in which i and j are produced by two different firms. Formally, the candidate has n firms, $i \in I$, $j \in J$, $I \neq J$. The expected adjustment costs associated with i and j in the candidate solution are

$$f(s_I)\Sigma_{k\in I/i}\, a_{ki} + m\Sigma_{k\in J/j}\, a_{ki} + m a_{ij} + f(s_J)\Sigma_{k\in J/j}\, a_{kj} + m\Sigma_{k\in I/i}\, a_{kj}. \qquad (1)$$

If i is moved to J, the expected costs change to

$$m\Sigma_{k\in I/i}\, a_{ki} + f(s_J+1)[\Sigma_{k\in J/j}\, a_{ki} + a_{ij} + \Sigma_{k\in J/j}\, a_{kj}] + m\Sigma_{k\in I/i}\, a_{kj}. \qquad (2)$$

Comparing *(1)* and *(2),* we see that the candidate solution can be improved by putting i in J if

$$a_{ij}>\Sigma_{k\in I/i}a_{ki}[m-f(s_I)]/[m-f(s_J+1)]+\Sigma_{k\in J/j}a_{kj}[m-f(s_J)]/[m-f(s_J+1)]-\Sigma_{k\in J/j}a_{ki}-\Sigma_{k\in I/i}a_{kj}, (3)$$

with a parallel condition for the option of moving j to I. That is, the candidate is suboptimal if a_{ij} outweighs the advantages of having i in I instead of J or vice versa. The chances of this happening are, ceteris paribus, increasing in a_{ij}.

The above property only characterizes the optimal design in a very imprecise way. The fact that i and j are more likely to be co-produced if the frequency of mutual adjustment a_{ij} is large can be thought of as a direct (or first-order) effect. It is not unlike the forces studied in the firm-level tests discussed in the Introduction. However, there are also several indirect (or higher-order or network) effects. An example of a second-order effect is that the parts i and j are more likely to be co-produced if there exists a third part k such that both a_{ik} and a_{jk} are large. While this condition might seem less likely to be satisfied, there are $N-2$ possible ways to meet it. A third-order effect would be the existence of a k,h pair such that a_{ik} and a_{jh} and a_{hk} all are large, etc. The only way to capture all indirect effects is to take all the elements of $\{a_{ij}\}$ into account by solving *(P)*.

We test the theory from three different angles. We start by checking the first-order implication that two parts i and j are more likely to be made by the same firm if a_{ij} is higher. To account for indirect effects, we then look at the entire industry and find that a disproportionate number of adjustments are managed inside firms. We finally estimate a structural model in which industry design is found as a solution to a program like *P* aimed at minimizing industry-wide adjustment-costs.

IV. Data

We test the theory by looking at industry designs and adjustment frequencies among parts going into cars. We have data from the late 1980s on eight supply chains pertaining to eight different cars in the luxury-performance segment of the automobile market. We look at each supply chain as a separate industry and will henceforth use that term and label the industries A, B, C, D, E, F, G, and H. Since there is virtually no overlap between the participants in the eight industries, we treat them as independent, but identical.[3] That is, we assume that the same technology and the same set of parts, and therefore the same adjustment frequencies, drive the design of all eight industries. So we can estimate on a per-industry basis, or pool the data across all eight.

A car consists of more than ten thousand parts. Since we need information on *N(N-1)/2* pairs of parts, computational constraints and problems associated with administering very long questionnaires force us to divide the set of parts into 36 "megaparts", henceforth "parts", for the purposes of the study. For each of the resulting 630 pairs of parts, we know the frequency with which mutual adjustment is needed as well as whether or not they are co-produced (made by the same firm). During the time period covered by the data, in all eight cases, every component of

[3] Across the 36x8 part-industry combinations, a total twentynine were subject to inter-industry linkages: one firm made a part for 4 industries, three made a part for 3 industries, and eight made a part for 2 industries. In each of these cases, the firm was required to have separate employees and often even separate production facilities to serve each customer.

each part was made by a single firm. So no inconsistency is introduced by conducting the study at the megapart level, although we do lose a significant amount of information. In particular, we can not make use of the fact that the sub-components of the parts are co-produced, even though our theory speaks directly to it. The (mega)parts are listed in Table 1 below.

Table 1

List of Parts

Body in white	Airbag controller	Intake manifold	Alternator
Body sheet metal	Airbag	Crankshaft	Speed control
Headlining	Power steering gear	Camshaft	Automatic transmission
Bumpers	Steering linkage	Piston	Suspension
Safety belts	Steering column	Intake valves	Drive shaft
Lock cylinders	Steering wheel	Radiator	ABS system
Door handles	Power steering pump	Starter	Spindle assembly
Windshield washers	Cylinder head	Distributor	Upper and lower arms
Seat system	Engine block	Instrument panel	AC assembly

The data on industry design (patterns of co-production), as well as some of our information about adjustment frequencies, come from interviews conducted by one of the authors. These interviews were very extensive and wide-ranging – involving more than 1000 employees of the eight firms. They were focused on sourcing, and included specific questions identifying the producer of every part in every industry. In spite of the fact that the industries use virtually identical technologies, there are very large differences in co-production practices between them. We have too few data points to explain these differences, but they do not appear to be related to societal factors. For example, the number of co-produced pairs ranges from 11% to 79% among four industries anchored in a single country.

The Interview subjects were not systematically asked about adjustment frequencies, but the topic was repeatedly touched upon and the interviewer was able to use her notes to construct estimates for all (part, part) pairs. Specifically, for each pair of parts, she rated "the frequency with which there needs to be a mutual adjustment in this pair" on a seven point scale from 0 - 6. As mentioned above, we think of these as fundamental engineering relationships that should apply to all the cars. Table A-1 in the Appendix displays the interview data on adjustment frequencies along with the co-production data for industry A.

A possible problem is that the interviewer rated the adjustment frequencies with knowledge of the hypothesis to be tested. As an ex post check, we therefore used a questionnaire to collect a second set of adjustment frequencies from an industry expert who was unaware of the hypothesis. Our expert is Dan Whitney, who for many years has played a major role in MITs International Motor Vehicle Program. After getting a table with the 630 (=36 x 35/2) pairs of parts, this expert was asked to think of a typical luxury-performance car in the late 1980s, and answer the question: "Please consider a pair of parts and rate, on a scale from 0-6, the frequency with which there needs to be a mutual adjustment in this pair." It turns out that the questionnaire ratings are extremely similar to the interview ratings. The expert could have calibrated the seven point scale differently than the interviewer, but at .88 and .87 the means are almost identical, as are the fraction of frequencies rated zero (.72 and .74). Most importantly, there is a highly significant Spearman rank-order correlation of .915 between the two data sets. To keep the argument as clean as possible, all results reported in the body of the paper are based on the expert's questionnaire responses. Analog analyses, based on our interviews, are reported in the Appendix. As can be seen there, the results are essentially the same for both measures.

Some descriptive statistics are given in Table 2, in which the second column indicates the fraction of each industry's parts that are co-produced, the third gives the sizes of all clusters of co-produced parts, and the *r'*s in the fourth column are Spearman rank-order correlations.

Table 2

Descriptive Statistics by Industry[1]

Industry	**Fraction co-produced**	**Size-distribution**	***r*(co-production, adjustment frequency)**
A	.29	19, 5, 2, 2, 2	.53
B	.26	16, 9, 4, 2	.51
C	.14	13, 5, 2, 2, 2	.61
D	.18	15, 3, 3, 2, 2, 2, 2	.58
E	.79	32	.51
F	.17	14, 5, 3, 3, 2, 2	.54
G	.12	12, 4, 3, 2, 2, 2	.63
H	.11	11, 4, 3, 3, 2	.64

[1]N=630.

Looking first at the fraction of parts that are co-produced, we notice that all the industries are designed differently, with E as an extreme outlier.[4] This heterogeneity ought to make us pessimistic about our tests, since we hypothesize that all industries are solutions to the same optimization problem, but observe apparently very different designs. As a first clue that this pessimism would be unfounded, we see in all eight cases a very strong rank-order correlation between co-production and adjustment frequency. In spite of the differences, all the designs reflect a strong influence of adjustment frequency. If all observations were independent, the Spearman correlations would have *t*-values (~25*r*) above 10, but the news is not that good. As noted earlier, the data exhibits complicated dependencies, because "parts 1 and 2 are co-produced" and "parts 2 and 3 are co-produced" imply that "parts 1 and 3 are co-produced". So the correlations in Table 2 do in some sense overstate the degrees of freedom and are just suggestive.

V. Estimation Techniques and Results

In this Section we will present three successively deeper tests of the relationship between adjustment frequency and co-production. To assess the robustness of the theory, we have done most of these tests on an industry-by-industry, rather than pooled, basis. We start by looking at direct effects only, asking whether two parts are more likely to be co-produced if mutual adjustments are needed more frequently. This is done for each pair, while incorporating corrections for the interdependencies between the pairs. To take account of

[4] E represents a corporate form not used in the U.S., and it is possible to argue that the 32 parts are produced by three, rather than one, firm. However, we wanted to be as conservative as possible. It should

indirect effects, we then go to the industry-level and compare the sum of internalized adjustment ratings against the distribution of the same measure in random industry designs. We finally estimate a structural model in which industry design is portrayed as the outcome of a maximization problem aimed at internalizing a weighted average of adjustments and random noise. This allows us to evaluate the importance of adjustment costs relative to other forces in industry design.

V.i. Tests at the pair-by-pair level

To draw statistical inferences about the relationship between co-production and adjustment frequency at the pair-by-pair level, we resort to simulation. There are many ways to look at these relationships, but we have chosen to compare the actual design (the actual pattern of co-production) to those in randomly generated allocations of the 36 parts to firms of the same size. It may seem more natural to compare against "completely" random designs generated by allocating each of the 36 parts to 36 equiprobable firms without the constraint on firm sizes. The problem is that all eight industries have quite skewed size-distributions, with the owner of the brand name being very large. Such large firms, and thus instances of co-production, would be relatively rare in a set of completely random designs, implying that our results would be much stronger. However, since we do not feel comfortable attributing the skewed size-distributions to our theory alone (the brand name manufacturers are becoming less and less integrated, but are still making a very large share of all parts); we prefer the more conservative test. The results are given in Table 3 below.

also be noted that three of the other industries are anchored in the same country as E.

Table 3

Probability of Pairwise Co-production by Adjustment Frequency[1]

Industry	0	1	2	3	4	5	6
A	.232 *(.001)*	.364 *(.234)*	.500 *(.089)*	.517 *(.020)*	.286 *(.598)*	.500 *(.012)*	.867 *(.001)*
B	.221 *(.008)*	.273 *(.474)*	.250 *(.605)*	.448 *(.026)*	.314 *(.272)*	.406 *(.050)*	.600 *(.019)*
C	.105 *(.002)*	.242 *(.139)*	.150 (.533)	.379 *(.004)*	.257 *(.056)*	.188 *(.306)*	.333 *(.095)*
D	.141 *(.006)*	.152 *(.681)*	.200 *(.491)*	.345 *(.045)*	.257 *(.173)*	.313 *(.057)*	.733 *(.000)*
E	.761 *(.087)*	.939 *(.113)*	.800 *(.662)*	.897 *(.164)*	.771 *(.683)*	.844 *(.301)*	.933 *(.283)*
F	.150 *(.047)*	.121 *(.795)*	.150 *(.650)*	.345 *(.029)*	.229 *(.245)*	.281 *(.089)*	.333 *(.140)*
G	.090 *(.005)*	.091 *(.745)*	.050 *(.887)*	.241 *(.070)*	.229 *(.059)*	.250 *(.038)*	.667 *(.001)*
H	.077 *(.004)*	.030 *(.963)*	.200 *(.194)*	.172 *(.202)*	.257 *(.007)*	.125 *(.458)*	.600 *(.000)*

[1] *p*-values in parentheses refer to tests versus random allocations, relative to 100,000 randomly designed industries.

Since there is a lot of information in Table 3, we will briefly look at the interpretation of a couple of cells. The .232 in the "A, 0" cell means that 23.2% of all pairs rated 0 were co-produced in industry A. This is significantly less than the 29% (see Table 2) one would expect from a random design. Similarly, the .867 in the "A, 6" cell means that 86.7% of all pairs rated 6 were co-produced in industry A. This is significantly more than the 29% one would expect from a random design.

Adjustment ratings from 1 to 6 are only given to about twenty pairs each, but the larger ratings are still individually significant compared to random allocations. The strong significances in the first column reflect the very large number of 0's in the data. The lacking significance for industry E is perhaps surprising in light of the high Spearman correlation (.51) from Table 2. However, the high level of co-production "eats up degrees of freedom" and causes the performance difference between a random and an optimal design to be very small, leaving us with less statistical power. We will see this more clearly in the industry-level analysis presented below.

Looking beyond the individual entries at the overall pattern of results in Table 3, we see that most rows show a monotonic increase indicating that pairs are more likely to be co-produced if the adjustment frequency is higher. We could also perform a statistical test of this relationship, but since a more correct analysis takes indirect effects into account, we will not offer any analyses of direct effects beyond those reported above.

In spite of the fact that we used the aforementioned conservative test, Table 3 show very strong support for the hypothesis that a pair of parts is more likely to be produced by the same firm if mutual adjustments are needed on a more frequent basis. The results are also broadly based in the sense that we see the same general pattern across all the industries in spite of their apparent differences.

V.ii. Tests at the industry level

The analysis in Table 3 is incomplete because it only takes direct effects into account. In fact, it is closely related to the one-part-at-a-time studies criticized in the Introduction. To capture indirect effects, we need a test that takes account of the entire matrix of adjustment frequencies. To this end, we use as our measure the sum of importance-weighted internalized adjustments. This can be formally expressed as $\Sigma_{ij}z_{ij}a_{ij}$, where z_{ij} is a 0-1 indicator of co-production. (We will later see that this is the correct measure of performance if the number of firms is held constant and the variable adjustments-costs are independent of firm size.) To test the hypothesis that a disproportionately large number of adjustments are internalized, we compare the actual value of this measure against those in 100,000 random designs.[5] Following the arguments made before Table 3, we again opt to be conservative and constrain the random designs to have the same firm sizes as the actual.

Table 4 below contains the results of the resulting test. The second column give the means of the measures ($E\Sigma_{ij}z_{ij}a_{ij}$) based on the simulations, the numbers in the fourth column ($Max_z\Sigma_{ij}z_{ij}a_{ij}$) are its theoretical maximum (the best that can possibly be done within the constraints of the actual size-distribution) and is found by integer programming. The third column reports the actual measures of internalized adjustments ($\Sigma_{ij}y_{ij}a_{ij}$), where y_{ij} is the value taken by z_{ij} in the data (a 0-1 indicator of actual co-production).

[5] Another way to approach the problem is to use an entirely different statistical technique. A particularly interesting candidate may be cluster analysis, the use of which has some history in sociological and ecological studies of network effects (Frank, 1995).

Table 4

Sum of Internalized Adjustments by Industry[1]

Industry	$E\Sigma_{ij}z_{ij}a_{ij}$	$\Sigma_{ij}y_{ij}a_{ij}$	$Max_z\Sigma_{ij}z_{ij}a_{ij}$
A	161	275 *(.0001)*	395
B	142	221 *(.002)*	397
C	79	143 *(.005)*	288
D	100	195 *(.0001)*	316
E	433	468 *(.120)*	521
F	95	147 *(.017)*	329
G	68	152 *(.00000)*	269
H	59	134 *(.0001)*	251

[1] *p*-values in parentheses refer to tests versus random internalization, relative to 100,000 randomly designed industries.

Consistent with the results at the pair-by-pair level, we see that the industry-level measures, except for industry E, are very significant. The result is also robust in the sense that all industries show similar patterns. So with the exception of industry E, there is strong statistical support for the claim that a disproportionate number of adjustments are managed inside firms.

One way to measure the magnitude of the effect is by comparing the performance of random designs ($E\Sigma_{ij}z_{ij}a_{ij}$), actual designs ($\Sigma_{ij}y_{ij}a_{ij}$), and optimal designs ($Max_z\Sigma_{ij}z_{ij}a_{ij}$). We can interpret ($\Sigma_{ij}y_{ij}a_{ij}$ - $E\Sigma_{ij}z_i a_{ij}$)/(Max$_z\Sigma_{ij}z_{ij}a_{ij}$ - $E\Sigma_{ij}z_{ij}a_{ij}$) as the actual "excess internalization" divided by the highest possible "excess internalization" (what would be observed if the theory explained everything and our measures were perfect). The average of this ratio is around *.33,* suggesting that one third of the forces captured by our measures are reflected in the actual design. However, the sheer size of the optimization problem raises an almost philosophical question about the use of full optimality as a benchmark. Clearly, no precise solution has been feasible until very recently and it is difficult to think that the industry, based on just 100 years of competition or experience, would have found the most efficient structure among so many possibilities. Industry G internalizes more adjustments than all of the 100,000 random designs created to test the hypotheses in Table 4, and yet scores less than one half on the $(\Sigma_{ij}y_{ij}a_{ij} - E\Sigma_{ij}z_i a_{ij})/(\mathrm{Max}_z\Sigma_{ij}z_{ij}a_{ij} - E\Sigma_{ij}z_{ij}a_{ij})$ measure. On the other hand, it is obviously very hard to argue for any specific benchmark other than full optimality.

V.iii. A structural model

The results in Table 4 tells us that actual industry designs internalize many more adjustments than one would expect if designs were random. They do not tell us why this is. To this end, we will use simulated GMM to estimate a structural model of industry design, aiming to show that it can be portrayed as the solution to a program like *(P)*. This will allow us to measure the extent to which ours is the right model of industry design. Specifically, we will formulate a program in which adjustment-costs plus noise are minimized, and show that the former can not be ignored. One way to do this is by writing the objective function in *(P)* as $(N-n)w + \frac{1}{2}\Sigma_{i\in \mathrm{N}}[f(s_I)\Sigma_{j\in I}(\beta a_{ij} + e_{ij}) \ + m\Sigma_{j\notin I}(\beta a_{ij} + e_{ij})]$, where once again N is the number of parts, n is the number of firms, $w(s-1)$ are fixed adjustment-costs in a firm of size s, $f(s)$ are the variable adjustment-costs in a firm of size s, m are the variable adjustment-costs in the market, a_{ij} is the frequency of mutual adjustment, and e_{ij} is random noise. The idea is now to evaluate the importance of adjustment frequencies by looking at the magnitude and significance of β. Since the theory allows the $f(s_b)$'s to be constant, we simplify a bit by restricting them to be independent of firm size. This allows us to write the objective function of *(P)* in terms of the two parameters f/w and m/w, thus reducing the dimensionality of the estimation problem. On the other hand, since it is hard to believe that the $f(s_b)$'s are constant, we achieve the formal simplicity at the cost of estimating a very coarse and presumably less well fitting model.

There are, however, significant computational barriers to even this plan. To estimate the model with simulated GMM, we start with a provisional parameter value β' (say 1) and solve *P* for a number of randomly drawn $\{e_{ij}\}$ matrices. After using the appropriate moment conditions to evaluate β', we repeat the procedure for a new parameter value, continuing until we find the best estimate. We might thus end up solving *(P)* more than 2000 times. But for $N = 36$, the number of feasible solutions to *(P)* is comparable to the number of seconds passed since the big bang, resulting in what presently are insurmountable computational demands. Rather than reducing the number of parts, we have once again chosen to constrain the simulated solutions to the same size-distribution of firms as that in the data, denoted by $(s_1^o, s_2^o, ..s_n^o)$. That is, we test that parts are allocated to minimize adjustment-costs under the assumption that the industry has to follow an exogenously imposed size-distribution.

A fixed size-distribution implies that the number of firms is constant, and the objective function in *(P)* reduces to $\frac{1}{2}\Sigma_{i\in \mathrm{N}}[f\Sigma_{j\in I}(\beta a_{ij} + e_{ij}) + m\Sigma_{j\notin I}(\beta a_{ij} + e_{ij})]$. This can be expressed as $\frac{1}{2}(f\text{-}m)\Sigma_{i\in \mathrm{N}}\Sigma_{j\in I}(\beta a_{ij} + e_{ij}) + \frac{1}{2}m\Sigma_{i\in \mathrm{N}}\Sigma_{j\in \mathrm{N}}(\beta a_{ij} + e_{ij})$, where the last term is a constant and $(f\text{-}m) < 0$. So total adjustment-costs are minimized by having as many intra-firm adjustments as possible, while respecting the actual size-distribution. Consequently, *(P)* is equivalent to the following much simpler Partitioning Problem

$$\text{Max}_S\, \Sigma_{i \in \text{N}} \Sigma_{j \in I} (\beta a_{ij} + e_{ij}), \text{ s. t. } s_b = s_b^o, \text{ for all } b. \qquad (P')$$

In our context, we can interpret *(P')* as a problem in which 36 parts have to be put into n firms of predetermined size in such a way that the sum of the intra-firm benefits, $(\beta a_{ij} + e_{ij})$, is maximized. If $\beta = 1$ and $e_{ij} = 0$, this is the same measure used in the simulations underlying Table 4.

While it is much, much smaller than *(P), (P')* is still a very large integer programming problem. In the typical industry, it has more than 10^{22} possible solutions, but can be formulated as a linear integer program with about 2000 variables and 4000 constraints. If we use a number of tricks to speed up the program, the CPLEX IP code allows us to find an optimal solution in a reasonable amount of time. Most of our individual runs take ten to fifteen minutes, although the time varies from a few seconds (for industry E) to several days or weeks (for industries F and B). The results reported below are based on a total of roughly 25,000 optimizations.

Because the actual size-distributions are quite skewed, we have argued that the tests reported in Tables 3 and 4 should be done against alternative designs with the same firm sizes. Without this constraint, large firms would rarely result from randomly generated data, and the actual designs would appear to perform extremely well. There is no similar problem with the structural model. For appropriate values of *w, m,* and *f(s),* we could easily find that a skewed size-distribution performs well, even with essentially random inputs (for very small values of β). So while it is nice to estimate the structural model under the same assumptions as those used in the simpler tests, we would prefer to estimate without the constraint on size-distribution; were it not for computational considerations. However, the theory does imply that the actual industry design is better than any other with the same size-distribution. So we are in some sense testing an implication of the theory, although our estimates are constructed in the restricted model only.

It should also be said that the dependencies between the observations makes it impossible to rely on standard limit theorems to provide a theoretical justification our estimation techniques. In theory, we would need to develop some more general results, along the lines of Berry, Linton, and Pakes (2004). With these caveats, we offer the GMM results as weak complementary evidence.

We estimate β for individual industries from the moment condition:

$$E\Sigma_{ij}[y_{ij} - Prob(x_{ij}=1|\beta^*)]\, a_{ij} = 0, \qquad (4)$$

where y_{ij} and x_{ij} are 0-1 indicators of co-production, referring to the data and the simulations, respectively. We find a measure of $Prob(x_{ij}=1|\beta')$ as the average value of $x_{ij}(\beta')$ over 100 simulations based on β'. We arrive at this average as follows: We first draw 100 independent $\{e_{ij}\}$ matrices each consisting of 630 independent draws from $N(0, \sigma^2)$, where σ^2 is the variance of the a_{ij}'s. Given a provisional β' and the first matrix, we then solve P by finding the allocation that maximizes $\Sigma_{ij}x_{ij}(\beta')[\beta' a_{ij} + e_{ij}]$ subject to the constraint that the size-distribution is identical to the actual. After solving (P') for the other 99 $\{e_{ij}\}$ matrices, we assign

β' a score of $\Sigma_{100}\Sigma_{ij}\, x_{ij}(\beta')a_{ij}/100$.[6] We repeat the process for several other provisional β's, searching for a score equal to $\Sigma_{ij}y_{ij}a_{ij}$. To estimate the model on pooled data, we sum the left hand side of *(4)* over the eight industries to arrive at the moment condition:

$$E\, \Sigma_{industries}\{\Sigma_{ij}[y_{ij} - \Sigma_{100}x_{ij}(\beta^*)/100]\, a_{ij}\}=0, \qquad (5)$$

We calculated the standard errors of the estimates by parametric bootstrapping from the asymptotic distributions. Specifically, we randomly drew 100 $\{\beta^* a_{ij} + e_{ij}\}$ matrices and optimized against each of them to find 100 hypothetical data matrices. We then estimated a value of β for each of these, and found the standard error from the distribution of these 100 β's.[7] The results are reported in Table 5, where the row labeled POOLED refers to estimates with pooled data.

---

[6] Since individual optimizations of industry B took up to two weeks, the analysis of that is less rigorous. We used just 30 simulations for each value of β, and in some of them went with approximate solutions.
[7] Due to the computational difficulties, we generated just 30 simulated data points for industry B.

Table 5

Structural Models of Max $\Sigma_{ij}x_{ij}(\beta)[\beta a_{ij} + e_{ij}]$.[1]

Industry	β^*	s.e. β^*
POOLED	.314	.095
A	.40	.15
B	.29	.11
C	.27	.10
D	.37	.14
E	.26	.26
F	.21	.11
G	.35	.12
H	.33	.13

[1]Standard errors are bootstrapped.
N=5040 for the pooled model, 630 for industries A-H.

In spite of the fact that we estimated a very coarse model with very coarse data, the betas are, with the exception those for industry E, highly significant. In light of the very different industry designs documented in Table 2, the betas are also surprisingly similar. While each of the eight industries has solved the design problem in its own way, it appears that they all weigh the adjustment-costs to a similar degree.

Because the a_{ij}'s and the e_{ij}'s have the same variance, we can use the magnitudes of the betas to get another perspective on the influence of adjustment frequencies. We have portrayed the industry as maximizing sums of $\beta a_{ij} + e_{ij}$, so $\beta^* = .33$ suggests that our theory and measures capture about .33/[.33 + 1], or one fourth, of the forces going into the determination of industry design. The extent to which the model fits the data is surprising in light of the complexity of the optimization problem postulated by the theory. As discussed in connection with Table 4, one can reasonably question the practical possibility of full optimality and thus of its use as a benchmark.

Summarizing the results, we have provided robust and multifaceted evidence in support of the claim that the adjustment-cost theory provides a good explanation for several apparently very different industry designs.

VI. Comparing Industry- and Firm-Level Estimation

Given the difficulty of estimating the industry-level model, it is interesting to compare the results to those of a firm-level model based on direct effects only. Specifically, how much more precisely could we estimate the model by using the entire matrix of adjustment frequencies, as opposed to a single row? To look at this, we represent the firm by body-in-white - a part always made by the brand name manufacturer. Since the pair-wise ratings between body-in-white and the other parts have larger variance than those in any other row, this choice gives the firm-level model as much information as possible and thus leads to the most conservative evaluation.[8]

To minimize confounds, we compare the industry-level model to an analogously constrained and estimated firm-level GMM. Specifically, if body-in-white is co-produced with s_1 other parts, (e_{1j}) is a vector consisting of 35 independent draws from $N(0, \sigma^2)$, and σ^2 is the variance of all the a_{ij}'s, we portray the firm as solving

[8] If we had chosen to represent the firm by a part that is produced in isolation, the logit model would have been inestimable, while our model would be unchanged.

$$\text{Max } \Sigma_j x_{1j}(\beta a_{1j} + e_{1j}), \text{ s. t. } \Sigma_j x_{1j} = s_1. \qquad (P'')$$

In close analogy to *(4),* we estimate β from the moment condition

$$E\, \Sigma_j[y_{1j} - \Sigma_{100} x_{1j}(\beta^*)/100]\, a_{1j} = 0. \qquad (6)$$

For each value of a_{ij} (0, 1, 2, 3, 4, 5, and 6), this model gives us an estimate of $Prob\ (x_{1j}=1|a_{ij})$. To evaluate the extent to which the firm-level model fits the industry data, we insert these probabilities into the entire matrix, subtract $\{y_{ij}\}$, square each cell, and sum the squares to find the sum of squared residuals (SSR) over the 630 cells.

As there is little reason to estimate a firm-level model with GMM, we also estimate an (unconstrained) logit model of co-production with body-in-white. Since this model is based on a different functional form, it is a bit harder to compare with our industry model, but we can measure its relative performance by finding the SSR over the 630 cells as above. The SSRs from the industry model, the firm-level GMM model and the firm-level logit model are given in Table 6 below.

Table 6

Sum of Squared Industry Residuals (SSR) by Alternative Models.[1]

Industry	Industry GMM	Firm GMM	Firm Logit
A	117.2	136.2	133.5
B	117.4	151.1	150.5
C	74.4	89.3	88.6
D	84.8	99.3	97.4
E	105.9	115.1	114.6
F	88.6	100.6	99.2
G	60.5	72.9	71.9
H	56.7	70.8	70.8

[1] All entries are based on 630 pairs of parts.

The table shows that the industry-level model consistently fits better than either of the two firm-level models. Since it makes more intensive use of the data, it is not surprising that the industry-level model does better than the firm-level GMM model. We showed in Section II that its theoretical advantage is larger when fewer parts are co-produced, and we would expect the superiority of the industry-level model to differ between industries. The Table bears this out as the ratio between the SSRs generally is larger in industries like H, G, C, F, and D where fewer parts are co-produced, and smaller in industry E.

As mentioned above, it is harder to compare the industry GMM with the logit model because the difference in data sets is confounded with a difference in functional forms. To (imperfectly) decompose the effects, we can start by comparing the two firm-level models. From Table 6, we see that the logit model fits the questionnaire data bit better, but Table 6A in the Appendix shows that the firm-level GMM fits the interview data better. As one would expect, the relative advantages of the two functional forms depend on the data sets. On the other hand, the logit model clearly does less well than the industry-level GMM, suggesting that any advantages tied to functional forms are overwhelmed by those associated with more intensive use of the data.

One could argue that the above model comparison is biased against the firm-level models in favor of the industry-level model, because only the latter is estimated on the data used for comparison. However, we will claim that this standard of comparison is the only correct one, since it is consistent with the belief that all co-production decisions follow the same logic. We can nevertheless get another take on the model comparison by evaluating model performance relative to co-production with body-in-white only. These results are given in Table 7 below.

Table 7

Sum of Squared Firm Residuals (SSR) by Alternative Models.[1]

Industry	Industry GMM	Firm GMM	Firm Logit
A	7.41	8.52	8.29
B	7.31	5.01	4.75
C	7.25	6.59	6.35
D	7.26	8.01	7.88
E	3.57	3.62	3.53
F	8.58	8.45	7.95
G	6.13	7.29	7.14
H	6.36	7.15	7.06

[1] All entries are based on 35 pairs of parts.

Since the two firm-level models are estimated on that data only, this measure will be more favorable to them. In spite of this, the industry-level model still outperforms the two firm-level models in more than half of the industries. We conclude that there are significant advantages of estimating at the industry-level, but admit that it uses more data and that it entails non-trivial computational difficulties.

VII. Discussion

We have tested the adjustment-cost theory of the firm according to which the firm is a low variable- , but high fixed cost way to govern adjustments. We presented several successively deeper tests: We showed that any two tasks are more likely to be performed by the same firm if mutual adjustments are needed on a sufficiently frequent basis. At the industry-level, we found that a disproportionate number of adjustments are managed inside firms, and that industry design can be portrayed as the solution to an integer program aimed at minimizing the sum of adjustment-costs. The results are very robust; while the industries differ widely in their overall level of integration, they all follow the model fairly well.

Our innovative testing strategy contributed to the strong results. By taking an industry-level perspective, we have been able to extract much more information from the data, especially in the cases where many parts are out-sourced. Consistent with this, we found that our structural model fits better than a firm-level model, both theoretically and in this dataset. The testing strategy is not tied to the adjustment-cost theory and it should be possible to apply it to other theories of the firm as well.

Consistent with common practice in theories of the firm, we have confined our attention to the efficient solution and focused on the industry structure that minimizes the sum of adjustment-costs. This is perfectly reasonable in light of our data, since the eight industries produce substitutable final outputs (competing luxury cars). However, in other cases it may be more appropriate to look for an equilibrium, although this might be harder.

On the empirical side, we were very encouraged by the .915 correlation between out two measures of adjustment frequency. Even so, we can not generally rule out that “the frequency with which there needs to be a mutual adjustment", is picking up something other than adjustment frequency. It would be nice to conduct tests of specific competing theories.

It is possible to interpret our results as reflective of endogeneity with the idea being that there are more or fewer mutual adjustments between two parts *because* they are or are not co-produced. There are two reasons we are not too worried about this. First, we took steps to avoid it by asking the expert about "the frequency of *needs* for mutual adjustments", and explicitly differentiating it from actual adjustments. Secondly, our theory predicts that adjustment between co-produced parts is cheaper. In the interest of simplicity, we derived P under the assumption that all need adjustments are made. But in a more realistic formulation, without this assumption, we would find that more of the needed adjustments are implemented within than between firms. It is certainly possible to question our point estimates based on this line of reasoning, but the argument relies on the premise is that the theory is correct. So to the extent that reverse causality is driven by differences in adjustment-costs, we are not too unhappy about it.

It would be preferable to estimate the simulated GMM model without forcing the β's to be the same and/or without constraining the simulated solutions to the same size distribution as the data. In the current setting, this is complicated by the size of P, but there are three possible ways to reduce the computational burden. The first and most direct solution is to work with fewer parts, say ten. Since some sets of parts are partially co-produced at that level of aggregation, it is this is not a viable option in the auto industry, but would be in many other settings. A second possibility is to use near-optimal solutions to *P*. Because our moment condition *(4)* involves the expected average of the optimum and we average over 100 simulations, we could conceivably use solutions that are within, say 1% of optimal, and still get close to the true parameter values.[9] However, our experiments with this have not resulted in significant reductions in computation time. The third avenue is to define approximation by processor time and work with an inflated version of the best solution found after a fixed amount of time. Suppose that 100 complete runs, each taking more than twenty-four hours, give an average $\Sigma x_{ij}a_{ij}$ score of eleven when completed, but only ten after twenty-four hours. In such a case one could possibly stop all runs after forty-eight hours and inflate their scores by ten percent. We evaluated this for about 350 runs on industries B and C, and found that it worked fairly well.[10] However, it is hard to put bounds on the degree of approximation involved.

Another interesting issue, directly related to the problems of computational complexity discussed above, concerns the use of full optimality as a benchmark when interpreting measures of importance. Given the extremely large number of possible solutions, it seems unreasonable to expect the fully optimal outcome even in a world perfectly described by the model. One way to gauge the validity of this line of argument might be to repeat the analysis on

[9] If we had found our estimates from some other criterion, such as maximizing the number of correctly predicted instances of co-production, it would be much harder to gauge the effect of using approximate solutions.

[10] In these experiments, we limited processor time to the point where about 75% of the runs were stopped short of full optimality. Evaluated at β^*, the best solutions for the stopped runs gave a value of $\Sigma x_{ij}(\beta a_{ij} + e_{ij})$ that on the average was within one percent of that achieved by the optimal solutions for the same runs. The corresponding values of $\Sigma x_{ij}a_{ij}$ were on the average within two percent of those associated with the optimal solutions. However, taking into account that $\Sigma x_{ij}(\beta)a_{ij}$ is fairly steep around β^*, the approximations seem tolerable in comparison to the standard errors on β^*.

simpler products to see if the point estimates of importance are higher in industries with fewer parts.

In closing, we admit that the results presented here are tentative and that much work remains to be done. However, the robust patterns in the data should give pause to those who reject the adjustment-cost theory of the firm on a priori grounds.

References

Bajari, Patrick, and Steven Tadelis, "Incentives versus Transaction Costs: A Theory of Procurement Contracts", RAND Journal of Economics, 32, no. 3, Autumn, pp.387-407, 2001.

Baldwin, Carliss Y., and Kim B. Clark, Design Rules: The Power of Modularity, Cambridge MA.: The MIT Press, 2000.

Berry, Steven, Oliver B. Linton, and Ariel Pakes, "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems", Review of Economic Studies, 71, no. 3, No. 248, pp. 613-54, 2004.

Eppinger, Steven D., "Model-Based Approaches to Managing Concurrent Engineering," Journal of Engineering Design, 2, no. 4, pp. 283-90, 1991.

Frank, Kenneth A., "Identifying Cohesive Subgroups", Social Networks, 17, no. 1, pp. 27-56, 1995.

Grossman, Sanford, and Oliver D. Hart, "The Costs and Benefits of Ownership," Journal of Political Economy, 94, no. 4, August, pp. 691-719, 1986.

Holmstrom, Bengt, and Paul Milgrom, "The Firm as an Incentive System," American Economic Review, 84, no. 5, September, pp. 972-91, 1994.

Masten, Scott E., "The Organization of Production: Evidence from the Aerospace Industry," Journal of Law and Economics, 27, no. 2, April, pp. 403-17, 1984.

Novak, Sharon, and Steven D. Eppinger, "Sourcing by Design: Product Complexity and the Supply Chain," Management Science, 47, no. 1, January, pp. 189-204, 2001.

Simester, Duncan I., and Marc Knez, "Direct and Indirect Bargaining Costs and the Scope of the Firm," Journal of Business, 75, no. 2, April, pp. 283-304, 2002.

Steward, D. V., "The Design Structure System: A Method for Managing the Design of Complex Systems," IEEE Transactions in Engineering Management, 28, no. 3, pp. 71-84, 1981.

Wernerfelt, Birger, "On the Nature and Scope of the Firm: An Adjustment-Cost Theory," Journal of Business, 70, no. 4, October, pp. 489-514, 1997.

Williamson, Oliver E., "Transaction-cost Economics: The Governance of Contractual Relations," Journal of Law and Economics, 22, no. 4, October, pp. 223-61, 1979.

Appendix: Results Based on Interview Data

Table A-1

Adjustment Frequencies from Interviews and Co-production Data for Industry A

	6	3	2	3	2	1	1	3	2	0	3	3	6	0	3	4	3	5	6	6	0	0	2	0	0	5	0	0	4	3	2	2	3	3	0
0		5	2	0	0	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
0	0		0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3
0	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0		6	0	0	5	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	1		0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0		0	2	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		4	4	2	1	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	0	0		6	3	2	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	3	0	0	0
0	0	0	0	1	1	1	0	0	1		3	0	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	1	0	0	0		4	6	6	5	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0		5	1	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0		5	5	0	0	0	0	0	0	0	0	2	0	4	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0		4	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0		0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1		5	2	6	6	4	2	0	2	2	0	2	3	3	2	3	2	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1		6	5	5	3	3	2	4	4	0	6	6	6	0	5	2	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1		3	3	0	1	4	0	1	0	3	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1		6	5	3	1	2	2	0	4	5	2	0	4	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1		5	3	1	2	2	0	4	5	2	0	4	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1		0	0	3	3	0	3	3	0	0	3	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		2	1	1	0	0	0	1	0	1	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0		0	0	0	0	0	2	0	0	0	0	0	1
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1		5	1	3	5	1	0	4	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	5	6	4	0	4	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0		0	1	0	0	3	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1		6	2	2	2	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		3	4	4	4	2	2	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0		1	0	2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		5	1	6	6	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0		0	3	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1	0		6	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0		0
1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1	0	1	0	

Table A-2

Descriptive Statistics by Industry[1]

Industry	Fraction co-produced	Number of parts co-produced	*r*(co-production, adjustment frequency)
A	.29	19, 5, 2, 2, 2	.52
B	.26	16, 9, 4, 2	.49
C	.14	13, 5, 2, 2, 2	.57
D	.18	15, 3, 3, 2, 2, 2, 2	.53
E	.79	32	.48
F	.17	14, 5, 3, 3, 2, 2	.53
G	.12	12, 4, 3, 2, 2, 2	.61
H	.11	11, 4, 3, 3, 2	.60

[1]N=630.

Table A-3

Probability of Pairwise Co-production by Adjustment Frequency[1]

Industry	0	1	2	3	4	5	6
A	.235 *(.002)*	.385 *(.218)*	.386 *(.134)*	.410 *(.114)*	.478 *(.045)*	.476 *(.061)*	.619 *(.005)*
B	.226 *(.016)*	.231 *(.697)*	.182 *(.917)*	.385 *(.073)*	.217 *(.754)*	.571 *(.004)*	.667 *(.001)*
C	.116 *(.017)*	.192 *(.321)*	.159 (.455)	.256 *(.067)*	.174 *(.424)*	.238 *(.182)*	.333 *(.033)*
D	.154 *(.028)*	.115 *(.866)*	.182 (*.567)*	.205 *(.199)*	.217 *(.402)*	.429 *(.010)*	.476 *(.003)*
E	.768 *(.150)*	.808 *(.506)*	.795 *(.525)*	.821 *(.448)*	.870 *(.249)*	.905 *(.173)*	.905 *(.186)*
F	.156 *(.099)*	.192 *(.468)*	.045 *(.998)*	.231 *(.247)*	.130 *(.801)*	.429 *(.007)*	.476 *(.003)*
G	.094 *(.009)*	.000 *(1.00)*	.045 *(.978)*	.282 *(.008)*	.130 *(.555)*	.429 *(.001)*	.476 *(.001)*
H	.083 *(.015)*	.115 *(.542)*	.091 *(.708)*	.205 *(.068)*	.174 *(.224)*	.238 *(.070)*	.286 *(.025)*

[1] *p*-values in parentheses refer to tests versus random allocations, relative to 100,000 randomly designed industries.

Table A-4

Sum of Internalized Adjustments by Industry[1]

Industry	$E\Sigma_{ij}z_{ij}a_{ij}$	$\Sigma_{ij}y_{ij}a_{ij}$	$Max_z\Sigma_{ij}z_{ij}a_{ij}$
A	162	264 *(.001)*	407
B	143	231 *(.001)*	420
C	80	132 *(.012)*	312
D	101	174 *(.004)*	340
E	436	476 *(.097)*	526
F	96	153 *(.011)*	356
G	69	154 *(.0000)*	295
H	60	112 *(.005)*	274

[1] *p*-values in parentheses refer to tests versus random internalization, relative to 100,000 randomly designed industries.

Table A-5

Structural Models of Max $\Sigma_{ij}x_{ij}(\beta)[\beta a_{ij} + e_{ij}]$.[1]

Industry	β^*	s.e. β^*
POOLED	.276	.069
A	.33	.13
B	.31	.08
C	.23	.09
D	.26	.11
E	.28	.36
F	.23	.09
G	.32	.10
H	.25	.10

[1]Standard errors are bootstrapped.
N=5040 for the pooled model, 630 for industries A-H.

Table A-6

Sum of Squared Industry Residuals (SSR) by Alternative Models.[1]

Industry	Industry GMM	Firm GMM	Firm Logit
A	115.2	131.8	132.8
B	116.3	149.8	149.3
C	73.7	92.1	93.0
D	87.9	98.2	98.5
E	114.6	107.0	107.1
F	88.2	96.1	96.6
G	62.4	68.2	68.6
H	59.2	66.7	66.1

[1] All entries are based on 630 pairs of parts.

Table A-7

Sum of Squared Firm Residuals (SSR) by Alternative Models.[1]

Industry	Industry GMM	Firm GMM	Firm Logit
A	7.29	8.26	8.14
B	6.77	4.48	4.54
C	7.16	5.72	5.66
D	7.78	7.42	7.31
E	3.88	3.57	3.50
F	8.69	7.71	7.73
G	6.51	6.83	6.74
H	6.98	6.92	6.80

[1] All entries are based on 35 pairs of parts.