

# Why Do Firms Own Production Chains?\*

Ali Hortaçsu  
University of Chicago  
and NBER  
[hortacsu@uchicago.edu](mailto:hortacsu@uchicago.edu)

and

Chad Syverson  
University of Chicago Booth  
School of Business  
and NBER  
[syverson@uchicago.edu](mailto:syverson@uchicago.edu)

## Abstract

Many firms own links of production chains—i.e., they own both upstream and downstream plants in vertically linked industries. We use broad-based yet detailed data from the economy’s goods-producing sectors to investigate the reasons for such vertical ownership. It does *not* appear that vertical ownership is usually used to facilitate transfers of goods along the production chain, as is often presumed. Shipments from firms’ upstream units to their downstream units are surprisingly low, relative to both the firms’ total upstream production and their downstream needs. Roughly one-third of upstream plants report no shipments to their firms’ downstream units. Half ship less than three percent of their output internally. We do find that manufacturing plants in vertical ownership structures have high measures of “type” (productivity, size, and capital intensity). These patterns primarily reflect selective sorting of high plant types into large firms; once we account for firm size, vertical structure per se matters much less. We propose an alternative explanation for vertical ownership that is consistent with these results. Namely, that rather than moderating goods transfers down production chains, it instead allows more efficient transfers of intangible inputs (e.g., managerial oversight) within the firm. We document some suggestive evidence of this mechanism.

---

\* An older version of this paper previously circulated as “Vertical Integration and Production: Some Plant-Level Evidence.” We thank Daron Acemoglu, Luis Garicano, Austan Goolsbee, Tom Holmes, Tom Hubbard, Lynn Riggs, Chris Snyder, Steve Tadelis, and seminar participants at the ASSA meetings, Chicago, Chicago Fed, Harvard, HEC Montreal, IIOC, LSE, NYU Stern, SITE, UC Berkeley, and the U.S. DoJ for helpful discussions and comments. Margaret Triyana provided excellent research assistance. Syverson thanks the NSF (award no. SES-0519062), the John M. Olin Foundation, and the Stigler Center for funding. The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Chicago Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed. Support for this research at the Chicago RDC from NSF (awards no. SES-0004335 and ITR-0427889) is also gratefully acknowledged. Hortaçsu: Department of Economics, University of Chicago, 1126 E. 59<sup>th</sup> St., Chicago, IL 60637; Syverson: University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637.

## I. Introduction

Firms often own links of production chains. That is, they operate both upstream and downstream units, where the upstream plant operates in an industry that makes an input for the downstream plant's industry. We explore the reasons for such ownership using two detailed and comprehensive data sets on ownership structure, production, and shipment patterns throughout broad swaths of the U.S. economy.

We find that most vertical ownership does *not* appear to be primarily concerned with facilitating physical goods movements along a production chain within the firm, as is commonly presumed. Upstream units ship surprisingly small shares of their output to their firms' downstream plants. One-third of upstream plants report no internal shipments. The median internal shipments share across plants is 3.8 percent, if shipments are counted equally, and 2.6 percent in terms of total dollar values or weight. Even the 90<sup>th</sup> percentile internal shippers are hardly dedicated makers of inputs for their firms' downstream operations, with 42 percent of the value of their shipments sent outside the firm. (However, a small fraction of upstream plants—around 2 percent—are operated as dedicated producers of inputs for their firms' downstream operations, and these plants also tend to be quite large. We will discuss this further below.) These small shares are robust to a number of choices we made about the sample, how vertical links are defined, and whether we measure internal shares as a percentage of the firm's upstream production or its downstream use of the product.

This result raises a puzzle. If firms don't own upstream and downstream units so the former can provide intermediate materials inputs for the latter, why do they own them? Certainly, much of the literature on vertical integration—stretching back to the landmark paper by Coase (1937), with other notable later contributions like Stigler (1951), and Grossman and Hart (1986)—couches firms' motives for integrating in terms of facilitating movement of products along a production chain.<sup>1</sup> (Of course, in some contexts like hotel or business services franchising, vertical integration often does not involve transfers of physical goods. Our paper, however, focuses on vertically integration and shipments in the *goods-producing* sectors of the

---

<sup>1</sup> The size of the literature precludes comprehensive citation. Surveys include Perry (1989), Salop (1998), Joskow (2005), and Lafontaine and Slade (2007). Much of the recent industrial organization research on integration has focused on foreclosure (market power) implications. Examples of recent theoretical and empirical work with broader views of the determinants of integration within and across industries include Antras (2003), Acemoglu, Aghion, Griffith, and Zilibotti (2004) and Acemoglu, Johnson, and Mitton (2005).

economy, like manufacturing. Our view is that a fair reading of the parables and case studies in the vertical integration literature would imply that many, if not most, researchers would consider moderating physical goods transactions a key motive for vertical ownership.)

We propose an alternative explanation that is consistent with small amounts of shipments within vertically structured firms—and even with an absence of internal shipments altogether. Namely, we surmise that the primary purpose of ownership is to mediate efficient transfers of *intangible* inputs within firms. Managerial oversight and planning strike us as important types of such intangibles, but these need not be involved. Other possibilities include marketing and sales know-how, but any information-based input might be transferred readily across upstream and downstream units.<sup>2</sup>

That vertical integration is often about transfers of intangible inputs rather than physical ones may seem unusual at first glance. However, as observed by Arrow (1975) and Teece (1982), it is precisely in the transfer of nonphysical knowledge inputs that the market, with its associated contractual framework, is mostly likely to fail to be a viable substitute for the firm. This, of course, does not preclude integration from also involving physical input transfers in some cases. Indeed, we find a small number of plants that are clearly dedicated producers for their firms' downstream production units. However, these are the exception rather than the rule. Thus, it appears that the “make-or-buy” decision (at least referring to physical inputs) can explain only a fraction of the vertical ownership structures we observe in the economy.

We also find other patterns in the data that are consistent with the intangible inputs explanation. We document that plants in vertical ownership structures have higher productivity levels, are larger, and are more capital intensive than other plants in their industries. We go on to show that these disparities, which we interpret as embodying fundamental differences in plant “type,” primarily reflect persistent differences in plants that are started by or brought into firms with vertical structures. In other words, while there are some modest changes in plants' type measures upon integration, most of the cross sectional differences reflect selection on pre-existing heterogeneity. Second, controlling for firm size explains most of these type differences. That is, plants of similarly-sized firms have similar types, regardless of whether their firm is structured vertically, horizontally, or as a conglomerate.

---

<sup>2</sup> These inputs might be just as likely to be transferred from the firm's “downstream” units to its “upstream” ones as vice versa. The names reflect the flow of the physical production process, not necessarily the actual flow of inputs within the firm.

These patterns evoke the equilibrium assignment view of firm organization advanced by Lucas (1978), Rosen (1982), and more recently by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007). To the extent that intangibles are complementary to the physical inputs involved in making vertically linked products, equilibrium assignment typically entails the allocation of higher-type intangible inputs to higher-type plants in each product category. If plant size is restricted by physical scale constraints, better intangible inputs will also be shared across a larger number of plants.

Simply put, higher-quality intangible inputs (e.g., the best managers) are spread across a greater set of productive assets. Some of these assets can be vertically linked plants, but their vertical linkage need not necessarily imply the transfer of physical goods among them. Furthermore, there may not be anything special about vertical structures per se. The evidence below suggests that firm size, not structure, is the primary reflection of input quality. Larger firms just happen to be more likely to contain vertically linked plants.

In this way, vertical expansion by a firm may not be altogether different than horizontal expansion. A typical horizontal expansion involves the firm starting operations in markets that are new but still near to its current line(s) of business, under the expectation that its current abilities can be carried over into the new markets. Physical goods transfers among the firm's establishments are not automatically expected in such expansions, though inputs like management and marketing are expected to flow to the new units. Vertical expansions may operate similarly. The industries immediately upstream and downstream of a firm's current operations are obviously related lines of business. Firms will occasionally expand into these lines, expecting their current capabilities to prove useful in the new markets. And, just as with horizontal expansions, transfers of managerial and other non-tangible inputs will be made to the new establishments. Yet no physical good transfers from upstream to downstream establishments need occur.

The upshot is that the assignment view of the firm is consistent with large firms composed of high-type plants operating (often) in several lines of business. Common ownership allows the firm to efficiently move intangible inputs across its production units. Many of these units will be vertically related, making these segments "vertical" in that the firm owns each end of a link in a production chain. But the chain need not exist for the purpose of moderating the flow of physical products along it.

This scenario is consistent with the evidence we document in the paper, and in particular with our primary result about the lack of goods shipments within vertically structured firms. The remainder of the paper lays out the evidence and tests the hypothesis in more detail. It is organized as follows. The next section gives an overview of the two data sources we use in the paper. We then explain in Section III how we use them to measure vertical integration and shipments internal and external to vertical chains within firms. Section IV reports the empirical results. Section V discusses our proposed explanation for the results in more detail. We conclude in Section VI.

## **II. Data**

We use microdata from two sources: the U.S. Economic Census, and the Commodity Flow Survey. We discuss each dataset in turn.

*Economic Census.* The Economic Census (EC) is an establishment-level census that is conducted every five years, in years ending in either a “2” or a “7”. Establishments are unique locations where economic activity takes place, like stores in the retail sector, warehouses in wholesale, offices in business services, and factories in manufacturing. Our sample uses establishments from the 1977, 1982, 1987, 1992, and 1997 censuses. We specifically use those establishments in the Longitudinal Business Database, which includes the universe of all U.S. business establishments with paid employees. The data has been reviewed by Census staff to ensure that establishments can be accurately linked across time and that their entry and exit have been measured correctly. We exclude data from before 1977 because plant-level data was available almost exclusively for the manufacturing sector before this time. This precludes proper classification of vertical ownership status for manufacturing plants owned by firms that are in fact vertically structured, but only into non-manufacturing sectors (say, for example, a firm that owns a manufacturing plant and a retail store that sells the product the plant makes).

Critically, the Economic Census contains the owning-firm indicators necessary for us to identify which plants are vertically integrated. (We discuss in Section III below how we make this classification.) Additionally, the Census of Manufactures portion of the EC also contains considerable data on plants’ production activities. This includes information on their annual value of shipments, production and nonproduction worker employment, production worker hours, book values of capital equipment and structures, intermediate materials purchases, and

energy expenditures. We use this production data to construct plant-specific output, productivity, and factor intensity measures; details are discussed further below and in the Data Appendix. In some cases, we augment the base production data with microdata from the Census of Manufactures materials supplement, which contains, by plant, six-digit SIC product-level information on intermediate materials expenditures.<sup>3</sup>

*Commodity Flow Survey.* The Commodity Flow Survey (CFS) collects data on shipments originating from mining, manufacturing, wholesale, and catalog and mail-order retail establishments.<sup>4</sup> Shipments in the survey are defined as “an individual movement of commodities from an establishment to a customer or to another location in the originating company.” The CFS takes a random sample of an establishment’s shipments in each of four weeks during the year, one in each quarter. The sample generally includes 20-40 shipments per week, though establishments with fewer than 40 shipments during the survey week simply report all of them.

For each shipment, the originating establishment is observed, as well as its destination ZIP code (exports report the port of exit along with a separate entry indicating the shipment as an export), the commodity, the mode(s) of transportation, and the dollar value and weight of the shipment.

We use the microdata from the 1993 and 1997 CFS; the former contains roughly 120,000 establishments and 11 million shipments, and the latter 60,000 establishments and 5.5 million shipments. As with the Economic Census, each establishment has an identification number denoting the firm that owns it. Both the establishment and the firm numbers are comparable to those in the EC, so we can merge data from the two sources. We match the 1993 CFS to the 1992 EC; this will inevitably lead to some mismeasurement of ownership patterns, but we expect this will be small given the modest annual rates at which plants are bought and sold by firms.

---

<sup>3</sup> For very small EC plants, typically those with less than five employees, the Census Bureau does not elicit detailed production or materials expenditure data from the plants themselves. It instead relies on tax records to obtain information on plant revenues and employment and then imputes all other production data. We exclude such plants—called Administrative Records (AR) plants—from those analyses below that use plant-level measures constructed from the Census of Manufactures (e.g., productivity), since they would otherwise be constructed from imputed data. While roughly one-third of plants in the Census of Manufactures are AR establishments, they typically comprise a much smaller share of industry-level output and employment aggregates because of their small size.

<sup>4</sup> Hillberry and Hummels (2003, 2008) use the CFS microdata to investigate various affects of distance on trade patterns. They do not make the within- and without-of-firm distinctions that we do here. These are the only other studies using the CFS microdata that we are aware of.

### III. Measuring Vertical Ownership and Shipments within Firms' Production Chains

This section explains how we use the Economic Census and the Commodity Flow Survey microdata to measure key inputs in our analysis: which businesses are vertically integrated, and whether the shipments of such establishments are used within their firm or sent to external buyers.

The first step in determining which businesses are in vertical ownership structures is to ascertain the industry affiliation of every establishment in the Economic Census. We use the 1987 Input-Output Industry Classification System, the taxonomy used by the Bureau of Economic Analysis (BEA) for constructing the Benchmark Input-Output tables. Within the manufacturing sector especially, this system closely mimics the SIC 4-digit system, though there is some aggregation of SIC industries, and more rarely, SIC industries are split among input-output (I-O) industries. Aggregation is more common outside of manufacturing.<sup>5</sup> While the EC data does not contain establishments' I-O industry classifications, it does contain their SIC codes, so reclassification is straightforward using the BEA's published concordance.<sup>6</sup>

The next step is identifying in which industries firms operate. The Economic Census microdata contains owning-firm identification numbers for virtually every plant in the nonfarm private sector, which makes it easy to observe the industries in which a firm owns establishments. One limitation of our data is that only one owner is assigned to each establishment (in the EC, the owner is the legal entity that the Internal Revenue Service considers responsible for the payroll tax; essentially it is an Employer Identification Number). If other firms have partial ownership of establishments, we do not see it.

---

<sup>5</sup> The SIC industries that are aggregated in the input-output taxonomy are typically those that sell different outputs to a "final demand" sector (e.g., personal consumption expenditures or gross private fixed investment) and use similar intermediate materials inputs and production processes. The input-output classification system is primarily concerned with intermediate goods and services transfers, so it places less importance on distinguishing products that vary only from the standpoint of final demanders. Since we share the focus on within-production-chain transfers here as well, the input-output classification system is appropriate for our analysis. One of the largest such aggregations in the 1987 input-output system, in terms of the number of industries involved, is industry 180400, "apparel made from purchased materials." This one input-output industry consists of the 23 four-digit SIC industries in 231x-238x. These SIC industries use similar inputs and production processes to make various apparel products primarily for personal consumption. Examples include industries like "mens' and boys' neckwear," "women's, misses', and juniors' dresses," and "robes and dressing gowns."

<sup>6</sup> A given plant is assigned to a unique industry. Some plants do produce final products that fall under different four-digit SIC industries, however. The Census Bureau classifies such plants based on their primary product (almost always the product accounting for the largest share of revenue).

We next determine if a firm owns establishments in industry pairs on both ends of a substantial link in a vertical production chain. We define a “substantial link” as existing between one industry and another based on the relative volume of trade flows between those two industries. Specifically, a substantial link exists between Industry A and any industry from which A buys at least five percent of its intermediate materials, or any industry to which A sells at least five percent of its own output. The industry pairs that comprise such links are determined using the BEA’s Benchmark Input-Output Tables.<sup>7</sup>

Finally, we find all establishments that the firm owns on both ends of a substantial vertical link and classify them as being in vertical ownership structures. If there are multiple vertical links within a firm, all establishments in the relevant industries are classified as integrated. While we only use manufacturing plants in some of our empirical work below because some of the detailed production data we use is limited to that sector, we use ownership information across all industries to determine which plants are and are not integrated.

As an example of how integration status is determined, consider a plant in I-O industry 490100, a.k.a. pumps and compressors. According to the Benchmark Input-Output Tables, this industry receives at least five percent of its total intermediate inputs from three upstream industries: 370200 (iron and steel foundries), 530400 (motors and generators), and 690100 (wholesale trade). Of its customers outside of final demand sectors, it sells more than five percent of its output to only a single I-O industry: 110000 (construction). A pump-compressor plant is labeled as vertically integrated, then, if its firm also owns a steel foundry, a motor-generator plant, an establishment housing a wholesaling operation, or a construction office. The corresponding plant(s) in the vertically linked upstream or downstream industry (industries) are also considered vertically integrated. Notice that integration is defined at the plant, not firm, level. If an integrated plant’s owning firm also owns other establishments that are not in a vertical production chain, these plants are not considered vertically integrated simply because the firm owns some plants that are. This distinction will be preserved in most of the empirical work below, though a few necessary exceptions will be noted.

---

<sup>7</sup> We use the 1987 tables. Given that the I-O structure of the economy is fairly stable over time, we do not expect those intertemporal differences in vertical commodity flows that we miss by using a single table over our whole sample to have a large impact. The five-percent cutoff used to define substantial vertical links is of course arbitrary. We have checked our major findings using ten- and even fifty-percent cutoffs and found few differences (the overall level of integration is of course lower in these more stringent cases).



To classify shipments in the Commodity Flow Survey from vertically integrated establishments as internal or external to the firm, we first must merge the CFS and EC data. This can be done straightforwardly using the two datasets' common establishment and firm identifiers. We then find the CFS establishments that we know from the EC data are in vertically integrated structures (determined as above), and furthermore, are on the upstream end of production chains within their firm. Whether or not an establishment is on the upstream end of a production chain is determined using the Input-Output tables: an establishment A is upstream if its firm owns another establishment B in an industry that buys at least five percent of the output of A's industry, or for which A's industry accounts for at least five percent of B's industry's input costs. (Note that A could be classified as *both* upstream and downstream if the firm also owns an establishment C in an industry from which A's industry either buys at least five percent of its inputs or buys at least five percent of C's industry output. In such a case, C would also be classified as upstream.) Next, for each upstream establishment in a vertical ownership link, we use the EC to find the ZIP codes of all the downstream plants owned by that same firm. We then compare the destination ZIP code of the CFS shipment to those of the firm's downstream plants. If the destination ZIP matches *any* of the downstream establishments' ZIP codes, we classify the shipment as internal to the firm.<sup>8</sup> The CFS contains shipment-specific sample weights that indicate how many actual shipments in the population each sampled shipment represents. We use these weights when computing the shares of internal shipments (be it by count, dollar value, or weight).

#### IV. Empirical Results

---

<sup>8</sup> Notice that we do not require that the shipment be destined to an establishment that is in an industry directly downstream to the shipping establishment, only that the destination be a plant that is on the downstream end of *any* vertical link in a firm. For instance, suppose a firm has two upstream establishments  $U_1$  and  $U_2$ , and two downstream establishments  $D_1$  and  $D_2$ , and  $U_1$ - $D_1$  and  $U_2$ - $D_2$  are separate vertical links. We would classify a shipment from  $U_1$  as internal if it is destined to *either*  $D_1$  or  $D_2$ 's ZIP codes, not just  $D_1$ 's. In this way, we are being liberal in defining internal shipments. Furthermore, because we only see destination ZIP codes, we are also assuming that a shipment to downstream establishment's ZIP code is indeed sent to that establishment rather than one outside the firm in the same ZIP code. This again will lead us to overstate the fraction of internal shipments (though likely only slightly, since in many industries, it is unlikely that there will be more than one establishment in a particular ZIP code). One factor leading to understatement, on the other hand, is that the EC is missing ZIP codes for about 10 percent of establishments. Therefore, intra-firm shipments to establishments for which we do not have ZIP codes will be misclassified as external. If these ZIP codes are randomly missing—there is no indication otherwise—then we can quantify the bias: internal shipments would be about 10 percent higher than reported. We further explore below many other measurement issues with classifying internal versus external shipments.

We begin by looking at the patterns of shipments within firms' vertical links. As described in the previous section, we find establishments in the Commodity Flow Survey that are a) in vertical ownership structures and b) upstream links within those structures. We then compare the destination ZIP codes of these plants' shipments to the locations of their firm's downstream plants. Matches are considered internal shipments.

#### *A. Vertically Integrated Establishments' Shipments—Benchmark Sample*

The combined 1993 and 1997 CFS yield a core sample of 29,931 plant-year observations of upstream establishments in firms' production chains. These establishments report a total of 2,826,296 shipments in the CFS. Panel A of Table 1 shows the prevalence of internal shipments within this sample. It reports quantiles of the distribution of internal shipment shares across plants, measured as the fraction of the total number, dollar value, and weight of the establishment's shipments.<sup>9</sup>

Overall, only a small share vertically integrated upstream establishments' shipments are to downstream units in the same firm. Looking across the roughly 30,000 establishments, the median fraction of internal shipments is 3.8 percent; the median internal shares by dollar value and weight are even smaller, at 2.6 percent. A third of these plants report no internal shipments at all. Even the 90<sup>th</sup> percentile plant sells over 40 percent of its output outside the firm.

The exception to this general pattern is the small set of establishments that are clearly dedicated to serving the downstream needs of their firm. These are the 2.1 percent of the sample that report only internal shipments. The unusualness of this specialization is even more apparent in the histogram of plants' internal shipment shares shown in Figure 1, panel A. The histogram echoes the quantiles reported above: the vast majority of upstream plants make few internal transfers. The fractions of establishments fall essentially monotonically as internal shipment shares rise—until the cluster of internally dedicated establishments. While not apparent in the histogram, another factor in the unusualness of these internal specialist plants is that they are larger on average. This, along with the internal share distribution being highly skewed, explain why the aggregate internal share of upstream plants' shipments (the across-plant sum of internal shipments divided by the across-plant sum of total shipments) is 19 percent. This is well above

---

<sup>9</sup> For data confidentiality reasons, the reported quantiles are actually averages of the immediately surrounding percentiles; e.g., the median is the average of the 49<sup>th</sup> and 51<sup>st</sup> percentiles, the 75<sup>th</sup> percentile is the average of the 74<sup>th</sup> and 76<sup>th</sup> percentiles, and so on.

the median share across plants. Thus internal shipments are more important on a dollar-weighted than an ownership-decision-weighted basis, but are the exception in either case.

These results imply the traditional view that firms choose to own plants in upstream industries to control input supplies may be off target. Clearly other motivations for ownership *must* apply for those plants making no internal shipments. Even for those that do serve their own firm, though, their typically small internal shipments shares suggest that this role may not be primary.<sup>10</sup>

### *B. Robustness Checks*

The level of the disconnect—at least in terms of physical goods transfers—between the upstream plants and their downstream partners is stark. We conduct several checks on the robustness of the result.

First, it is appropriate to review some details of how the Commodity Flow Survey is conducted, specifically with regard to its ability to capture intra-firm transfers. The CFS definitely seeks to measure them, and it makes no distinction between intra- and inter-firm transfers in its definition of “shipment.” In fact, the survey instructions (U.S. Census Bureau (1997)) state explicitly that respondents should report shipments “to another location of your company,” save for incidental items like “inter-office memos, payroll checks, business correspondence, etc.”

There are several reasons to believe the implied shipments totals are accurate. First, the Census Bureau audits responses by comparing the establishment’s implied annual value of shipments from the CFS with that from other sources. If the disparity is well beyond statistical variance, the Bureau contacts the respondent and reviews the responses for accuracy. If integrated establishments were systematically underreporting internal shipments because of confusion or by not following directions, the auditing process would help catch this.

Second, most establishments do report *some* internal shipments, indicating that they have not interpreted the definition of shipments as precluding intra-firm transfers. This is also reflected in the small share of establishments that report nothing but internal shipments. Moreover, there is no mechanical reason why we should find the bump up in internal shipment

---

<sup>10</sup> It is possible in some production chains that an upstream plant could serve its firm’s downstream needs completely with only a small fraction of its output. We will explore whether this factor might be driving our results shortly.

shares near one, as seen in Figure 1. We take this as further evidence of survey takers' comprehension of the CFS instructions.

Finally, for CFS establishments in the manufacturing sector, we can compare their implied annual shipments (the weighted sum of their sample shipments using the shipment-specific sample weights included in the CFS) to the same plants' reported values of shipments from the Census of Manufactures (CM). These two shipments totals are reported on separate survey instruments, and often by different individuals at the plant. We find a strong correlation between them. The correlation coefficient between plants' logged shipments from the CM and the CFS is 0.85 across our matched sample of 15,043 plant-years, and a regression of the former on the latter yields a coefficient of 0.973 (s.e. = 0.004).

### *B.1. Robustness: Sample*

In our first series of robustness checks, we consider the impact of modifications to our core sample of upstream vertically integrated plants. The corresponding distributions of establishments' internal shipments are shown in Table 4, panel B. Each row is a separate check. We show only the distributions of the dollar value shares for the sake of brevity; similar patterns are observed in the shares by shipment counts or total weight.

The first robustness check uses only establishments reporting at least the median number of shipments across all establishments in the sample. The point is to exclude those for which sampling error could be higher and for whom extreme values like zero are more likely. This leaves us with a sample of 15,161 establishment-years making just under 1.9 million shipments. (This is greater than half the establishment-years in the benchmark sample because several plants report exactly the median number of shipments.) The results are in the first row of panel B. Extreme values are in fact rarer in this sample: 29.4 percent report making no internal shipments, down from 34 percent in the full sample, and 0.7 percent report exclusively internal shipments, down from 2.1 percent. The remainder of the distribution is not much different, however. The median fraction of internal shipments is 3.9 percent (3.1 percent by value or weight), and the 90<sup>th</sup> percentile establishment is less likely to ship internally than that in the full sample, with just under half of shipments being intra-firm.

The second check drops any establishment that reports any shipments for export. In the CFS, the destination ZIP code of shipments for export is for the port of exit, with a separate note

indicating the shipment's export status and its destination country. Thus internal shipments to a firm's overseas locations would be misclassified as outside the firm, unless by chance the firm has a downstream establishment in the port's ZIP code. Focusing on the 21,219 establishments reporting no exports among their roughly 2 million shipments avoids this potential mismeasurement. The results are in the second row of panel B of Table 4. The entire distribution is close to the benchmark results above, with the median internal share being 2.7 percent and 35.7 percent of establishments reporting zero intra-firm shipments. Missing export destinations are not the source of our results.<sup>11</sup>

The next check drops all wholesale establishments from the sample. Many industries have the wholesale sector as both an upstream and downstream link. Therefore, we may be misclassifying some wholesalers as upstream links in a firm, when they are in fact downstream ones. Since the latter type are more likely to ship to final demanders like consumers (who are obviously external to the firm), they would indicate a deceptively small share of internal shipments for upstream establishments. Moreover, because of the nature of their business, wholesalers comprise a significant portion of the CFS sample. This amplifies the impact of any such classification problem. The no-wholesaler sample consists of 16,646 establishment-years and 1.6 million shipments. The across-plant distribution of internal shipment shares, in the third row of panel B, does indicate that non-wholesalers tend to be slightly more internally focused, though again most plants ship little inside their firms. The median fraction of internal shipments is 5.4 percent, and the 90<sup>th</sup> percentile is 69 percent. Just under one-quarter of plants report no internal shipments, lower than in the entire sample, and the fraction that are completely internally specialized is higher, at 2.7 percent.

The fourth check counts shipments destined for the ZIP code of *any* plant in the firm as internal, not just those going to locations of downstream links of vertical chains. It is possible that some vertical production may occur outside those chains we identify using the Input-Output tables. Some may even occur within a given industry, when a particularly complex production process is broken up across multiple establishments. Here, we are taking the broadest possible view toward shipments that represent intra-firm transfers of physical goods along a production chain. As seen in row 4 of panel B, all quantiles involve internal shipment fractions higher than

---

<sup>11</sup> We will discuss below how the fraction of international trade that is within firms could be so much larger than the intra-firm domestic trade we document here.

the benchmark numbers, as they must. The median is 6.5 percent, and the 90<sup>th</sup> percentile 67.2 percent. About 20 percent of establishments still have no shipments to a ZIP code of any plant in their firm, and exclusively internal establishments number 2.5 percent. A histogram of internal shares across plants for this definition of intra-firm shipments is shown in panel B of Figure 1. While internal shares are of course higher, it has the same shape as the benchmark distribution in panel A.

In the fifth check we make the generous assumption that a shipment is internal if it goes to any *county* in which the firm has a downstream establishment. While unrealistic, this approach accounts for almost any problems with ZIP code reporting errors or missing ZIP codes. The results of this exercise are in row 5 of panel B. Not surprisingly, the shares of shipments considered intra-firm are considerably higher, given the easier criterion for being defined as internal. There are more internal specialists or near-specialists: the 90<sup>th</sup> percentile internal share is 96 percent, and 6.3 percent of establishments have all of their reported shipments being internal. Even so, a substantial fraction of establishments—17 percent, more than twice the number of internal specialists—report no shipments to counties where downstream plants in their firms are located. The median internal share across plants is 28 percent.

The sixth check restricts the sample to plants in the twenty-five manufacturing industries with the least amount of product differentiation, as measured by the Gollop and Monahan (1991) index of product differentiation. The concern is that even the detailed Input-Output industry classification scheme may be too coarse to capture the true extant vertical links. For instance, it might be that while two industries are substantially linked at an aggregate level, this actually reflects the presence of (say) two separate vertical links within a six-digit I-O industry. In this case, we wouldn't expect many shipments to go from upstream plants in one link to downstream plants in another, even though we might infer the two are vertically linked just from comparing the industry-level trade patterns. By selecting industries with undifferentiated products, we hope to remove most of the product heterogeneity within detailed I-O industries, raising the probability that the industry-level links in the I-O table hold at a disaggregate level. There are 1216 plant-years in this subset of industries in the CFS. We find that median internal shipment shares are higher for these plants at 6.1 percent, but the right tail is actually less internally focused; the 90<sup>th</sup> percentile plant's internal share is 39.3 percent.

The remaining of the first set of robustness checks explore the impact of varying the five

percent cutoff for defining substantial vertical links. Row 7 (row 8) of the table shows the results using a 10 percent (50 percent) cutoff. Both of these changes restrict attention to increasingly interrelated vertical links at the industry level. The 10 percent cutoff sample contains 20,818 plant-years, and the 50 percent sample 3187 plant-years. Yet even here, internal shipment shares are low. They are, in fact, declining with the stringency of the definition of what constitutes a vertical link.

All in all, our benchmark results appear robust to several sample and variable construction changes.

### *B.2. Robustness: Is Geographic Closeness Important?*

It's quite likely that some of the low internal shares we see above arise because a firm's plants are too spatially separated to make internal shipments practical. Of course, if this is the case, this may be a *result* as much a *cause* of the lack of within-firm goods transfers along a production chain. If moving physical products down a production chain was the only reason for vertical ownership, after all, no firm would own vertically related plants that were located too far from one another to make intra-firm shipments impractical. The fact that firms do own vertically linked producers that are far apart suggests other motives for ownership.<sup>12</sup>

Nevertheless, it is interesting to quantify how much distance matters. We take two approaches. The first is to compute the distribution of internal shipment shares for firms whose plants are all located close to one another. The second is to compare plants' shipment distances to the distances they are from other plants in their firms.

To see shipment patterns of closely-spaced firms, we use the subset of upstream plants from our CFS sample where *all* of the plants that their firm owns are in the same county. (This is determined from the Economic Census, which includes state and county codes for virtually all establishments. This location information is not subject to the limitations of the EC ZIP code data, where codes for 10 percent of plants are missing.) This subset is small—680 plant-years and 60,847 shipments—and contains a large number of two-plant firms with one upstream and

---

<sup>12</sup> Hortaçsu and Syverson (2007) document examples of vertically integrated cement and concrete firms that own clusters of ready-mixed concrete plants that are remotely located from their upstream cement plants. These firms, in fact, do not internally supply these clusters with cement. The downstream concrete plants report instead buying cement in the local market from the firm's upstream competitors. We find evidence that the firms' motives for owning these concrete plant clusters is to harness logistical efficiencies in a business that shares a common final demand sector (construction) with cement.

downstream plant each. Nevertheless, it offers a rough gauge the role of distance.

The results are consistent with the patterns above. Just over half (51.3 percent) of the upstream plants report no shipments to downstream units in their firm. The 90<sup>th</sup> percentile plant ships 41.3 percent of the value of its shipments internally. The fraction of plants with all shipments staying in the firm is slightly above that in the benchmark sample, however, at 3.1 percent. Thus it appears that vertically structured firms with closely located plants are less likely to make internal shipments on average, but somewhat more likely to contain internally dedicated upstream plants.

We next compare the shipment distances of our entire sample of upstream plants in the CFS to their distances from other plants in their firms (both measured in great circle terms). It's clear from pooling shipments across plants that internal shipments go shorter distances. In fact, the average external shipment is sent roughly three times as far as the average internal shipment. This may reflect upstream plants “bypassing” their downstream units with some of their shipments, but it may also reflect composition effects if internally dedicated, high-volume upstream establishments are located close to downstream units in their firm.

We can decompose these contributions to the pooled numbers by looking within plants. We find that for 18.2 percent of upstream shipment plants, their farthest-traveling shipment does not go as far as the distance to the nearest downstream plant in their firm. These plants account for just over half of the one-third of our upstream plants that report no internal shipments, showing the importance of distance. But this also means the other half of plants reporting no internal shipments do send output at least as far as their nearest plant. This pattern isn't unusual across the broader sample. Looking across plants, the average of the within-plant medians of reported shipment distances is 256 miles, while the average distance to the closest downstream plant within the firm is 242 miles.

### *B.3. Robustness: Accounting for Actual Downstream Use*

We measure internal shipments above as an upstream plant's internal shipments as a share of its total shipments. There are cases where this ratio might be misleading as to the extent of intra-firm product movements. Consider a hypothetical copper products company with two plants: an upstream mill that produces copper billets, and a downstream plant that processes billets into pipe. Suppose the downstream plant needs \$10 million of billets to operate at



capacity. Now say the upstream mill produced \$100 million of billets in a year. If the mill shipped \$10 million of billets to the pipe-making plant and the remaining \$90 million elsewhere, we would compute its internal shipment share as 10 percent. Yet the firm would be completely supplying its downstream needs internally. The difference in the scales of operations upstream and downstream creates this misleading internal share.

Like the geographic closeness issue above, such patterns may be a result, rather than a cause, of low internal shipments within firms' vertical ownership structures. Nevertheless, in this section we take firms' upstream and downstream magnitudes of operation as given and create an alternative measure of internal shipment shares.

Instead of using upstream plants' total shipments as the denominator in the internal shipment share measure, we instead calculate (in ways described below) firms' downstream use of products they make upstream. We then construct internal shipments shares as intra-firm shipments of upstream plants divided by the *minimum* of two values, either the firm's total upstream shipments (similar to what we do above), or the firm's reported downstream use of the upstream product. Hence the internal share of the hypothetical copper firm above would be 100%, rather than 10% as before, since the firm completely provides all the copper it uses downstream.

While the CFS offers a random sample of establishments' shipments, we unfortunately do not have a random sample of establishments' incoming materials. This precludes us from directly measuring "internal purchase shares" in the same way as we measure internal shipment shares. But we can, for a subset of firms, construct internal shipments as a fraction of downstream use. To do so, we must first restrict our CFS sample. For our purposes here, we need to see *all* the upstream shipments of a firm, at least for a given product. If firms served downstream needs from upstream plants not in the CFS, we would not observe their non-CFS plants' shipments, and therefore would not know they are internal. Hence we look here only at CFS plants where we observe all the firm's plants in a particular industry. The Economic Census is used to find this subset of establishments, which ends up being about one-fifth of our benchmark CFS sample, 6228 plant-years. If we calculate these plants' internal shipment shares as above, this subsample looks similar to the entire sample. For example, 34.7 percent of these plants report making no internal shipments, the median internal share is 2 percent, and the 90<sup>th</sup> percentile plant ships 60.7 percent of its output internally.

We then match these upstream plants' shipments to downstream usage within the firm. We construct three downstream usage measures. The first simply aggregates the materials purchases of all the firm's downstream manufacturing plants. These purchases are reported by every plant in the Census of Manufactures. We aggregate across plants and products; the firm's downstream use of upstream products is simply the sum of all intermediate materials purchases. We can compute these downstream use measures for 2835 firm-year observations. To compute internal shares, we add up the internal shipments of the firms' upstream plants to use as the numerator.<sup>13</sup>

The second measure of downstream usage matches upstream shipments to downstream usage by product. We use the detailed materials purchase information from the Census of Manufactures materials supplement, which collects plants' materials purchases by detailed product. We compute firm's upstream shipments by product using the shipment commodity codes available in the CFS. Product specific shipments are computed at the 2-digit level. (We use only 1993 CFS data here because a change in the commodity coding scheme made it difficult to match the 1997 CFS commodity codes with the materials codes in the Census of Manufactures.) We sum the same firm's reported downstream use of that 2-digit product from the Census of Manufactures. The internal shipment share is the ratio of the firm's internal shipments of the product divided by its reported downstream use of that product. We are able to match 2744 firm-material combinations from 1518 different firms.

The third and final measure of downstream materials usage repeats the procedure above, except matches at the more detailed 4-digit product level. Because the greater detail makes finding matches less likely, we have a sample of 1175 such firm-product combinations from 689 different firms.

The results from these exercises are shown in Table 2. Recall that we now compute

---

<sup>13</sup> There are two measurement problems with this first approach that will tend to bias our internal shares measures in opposite directions. First, because we only required that we observe all of a firm's plants making *a particular product* in the CFS, we might be missing internal shipments from firms' other upstream plants (this is much less of a problem in our other two downstream use measures below, since they are matched by firm-product, rather than just by firm). This will cause us to understate the true internal shipment share. The second measurement issue arises because we can only observe materials purchases for downstream establishments in the manufacturing sector. If some upstream products are used in the firms' non-manufacturing establishments, we will not include these in our downstream usage measures. This will lead us to overstate internal shipment shares. As a practical matter, both of these measurement concerns are probably second-order. Our restricted sample has a large fraction of firms with only a few establishments, so if a firm's upstream plant(s) is in the CFS and its downstream plant(s) in manufacturing, chances are those are all the plants the firm owns.

internal shipments as their share of the *smaller* of a) the firm's (or firm-product's) total upstream shipments or b) the firm's downstream usage. Again, only the dollar-value shares are shown for brevity. The first row shows shares computed using the firm-level match where internal materials usage is aggregated across all materials. The second row shows results from the sample of matched firm-products at the 2-digit level; the third shows the firm-product match sample at the 4-digit level.

All three measures downstream usage still imply that most vertical ownership structures are not about serving the downstream material needs of the firm. The median share across plants of internal shipments as a fraction of the smaller of the firm's upstream shipments and its downstream use is 2.7% in the first (firm-wide) downstream use measure. The share of this subsample reporting zero internal shipments is 36.9 percent. For the second measure of internal usage (firm-product matching at the 2-digit level), the median internal share is 1.7 percent and 43.6 percent of plants report no internal shipments. The third measure (firm-product matching at the 2-digit level) sample has a median internal share of 3.5 percent and 44.3 percent without internal shipments.

One thing to note about the results is that some shares are above one. It is possible that this reflects in part the fact that some of the upstream plants' shipments that we classified as internal because their destination ZIP code was where the firm owned a downstream plant in fact went to a plant outside the firm in the same ZIP code. But probably some of these shares reflect measurement error in firms' downstream materials use (because it is outside the manufacturing sector and we can't observe it, for instance). A summary measure of the extent of such measurement error is the fraction of observations with implied internal usage ratios above one. For the three downstream use measures above, these shares are 5.3, 11.3, and 16.7 percent, respectively.

Thus the small internal shares we were finding before do not seem to be simply reflecting the fact that most integrated structures have considerably larger upstream plant scales than downstream. In fact, we still find a large number of cases (over one-third of the sample) without any intra-firm shipments. In other words, we know a firm makes a particular product upstream, uses that same product as an input downstream, but does not ship any of its own upstream output to its downstream units.

#### *B.4. Robustness: Is There Vertical Integration Within Plants?*

Our definition of vertically integrated ownership links requires multiple plants by definition. A firm must own at least one plant each in vertically related upstream and downstream industries. This definition could be problematic if firms commonly vertically integrate production within a single plant. In such cases, the firm would be operating a vertically integrated production process and obviously supplying its own input needs. We would miss this type of integration, however, because we would not classify the plant as integrated. There would be no shipments from the upstream to downstream parts of the production process in the CFS, since those goods transfers never leave the plant.

To give a concrete example, consider the two following hypothetical firms. One has two plants. The upstream plant refines copper ore into billets which are then shipped to the downstream plant to be extruded into pipe. The second firm operates a similar production process in a single plant: one side refines ore into billets, and the other side turns billets into pipe. We would define the former plants as vertically integrated, but not the latter, even though each firm operates the same production processes.

How can we tell if this sort of misclassification is a big problem? We compare the materials purchase patterns of plants that we classify as being in vertical structures to those in the same industry not classified as such. In the context of the above example, we compare the two copper pipe plants. (Since plants are classified into industries in the Economic Census based upon their outputs, both the downstream plant in the first firm and the second firm's plant would be classified in the same industry, SIC 3351: Rolling, Drawing, and Extruding of Copper.) The pipe plant in the first firm—the one that we would have classified as in a vertical ownership structure—will list copper billets as an intermediate materials purchase in the Census of Manufactures materials supplement. The second plant, where billet production is inside the plant, will list copper ore as a materials purchase. Hence if we see substantial differences in materials use patterns across plants (in the same industry) that we classify respectively as vertically linked or not, we should be concerned that we are missing a lot of vertical production that occurs “under one roof.” On the other hand, a lack of significant differences suggests this sort of misclassification is less of a concern.

We make three such comparisons between the materials use patterns of what we classify as vertically linked plants and others in their industry. (Again our analysis is restricted to plants

in the manufacturing sector because of the detailed intermediate materials data requirements.) We first compute the share of each plant's intermediate materials purchases that is for "raw materials," which we define as the products of the agricultural, fisheries, forestry, or mining sectors—i.e., SIC product codes beginning with "14" or below. We then regress this share on a set of industry-year fixed effects and an indicator equal to one if we classify the plant as in a vertical ownership link. In essence, we test whether there are significant differences in the intensity of raw materials use across plants that we classify as vertical and non-vertical in the same industry. We would expect that if the "under one roof" misclassification problem were substantial, we would find that plants we designate as *non*-vertical would have a larger raw materials share, since a greater portion of the production chain would be operated within the plant. Again, to return to our example, the pipe plant in the second firm reports copper ore (a raw material) as a materials purchase, while the plant in the first firm purchases copper billets.

We run this regression on a sample of over 484 thousand plant-years from the Census of Manufactures. (We don't need the CFS for this.) The coefficient on the vertical ownership link indicator is 0.34 percent, with a standard error of 0.49 percent. Thus plants we classify as vertical do not use systematically different amounts of raw materials than other plants in their industry. Further, the point estimate of the share difference is small, less than one-twentieth the average raw materials share of 8.0 percent, and of the wrong sign (recall that we would expect plants we classify as vertically linked to use raw materials less intensively). Even if we restrict our comparisons only to those roughly 87,000 plants that report using positive raw materials shares, the vertically linked coefficient is 0.13 percent with a 0.18 percent standard error.

Our second check aggregates this raw materials use data to the industry level. We add up raw and total materials use of plants classified as vertical within an industry year, and compare the ratio of the two to the same share computed for non-vertical (again, under our classification) plants in the same industry. We then conduct a t-test for equality of means across our sample of 1714 industry-years. The mean difference is 0.09 percent, with a standard error of 0.16 percent. Again, there are no significant within-industry differences in raw materials usage intensity across the two types of plants.

Our final check is also done at the industry-year level. We separately aggregate materials purchases of our designated vertical and non-vertical plants for each industry year. We then order materials by decreasing intensity of use for each type of plant (as measured by their

aggregate share of purchases). This yields 73,668 industry-year-materials ranks for both vertical and non-vertical plants. We then compare these ranks within industry-year to see if there are systematic differences. The two ranks move together; the correlation coefficient is 0.73. Table 3 shows the frequency of relative rank orderings for the five most intensively used materials by industries' non-vertical plants. (Material 1 is the most intensively used.) Only ranks 1-7 of vertical plants are shown for parsimony. If materials usage patterns were exactly the same, we would only see entries on the diagonal of the table. The most intensively used material of an industry's vertical plants would be the most intensively used among its non-vertical plants; the second-most used would be so for both types of plants, and so on. Clearly, this is not the case. However, the general pattern holds. The diagonal is the largest element of a row or column, and the frequency of other pairings falls as they move further away from the diagonal. Hence these results suggest, as do the raw materials use tests above, that there are not systematic differences in the mix of materials used by plants we classify as in vertical ownership links and those we do not classify as such.

#### *B.5. Robustness: Changes in Shipment Patterns upon Vertical Ownership*

Some plants are in both the 1993 and 1997 CFS samples, so we can observe changes in shipment patterns over time for a small panel of plants. We use this panel to see if *becoming* part of a vertical ownership structure is associated with notable changes in shipment patterns. We select establishments that are in both the 1993 and 1997 CFS and that are not in vertical ownership structures in 1993. This yields a sample that is just shy of 12,500 plant-year observations. (Note that this sample necessarily includes plants outside of our benchmark sample of upstream establishments above.) We then regress several shipment patterns metrics for these plants on a set of plant fixed effects, an indicator for 1997, and an indicator for a plant in a vertical ownership structure (this can only be equal to one in 1997 by our sample selection criteria). This will show us whether plants that become vertically structured by 1997 see systematically different changes in shipment patterns than those remaining outside of vertical ownership links.

The results are shown in Table 4, panel A. We look at five metrics, each computed at the plant level (we take logs of each except the last): the average dollar value per shipment, average weight per shipment, average miles per shipment, the average dollar value per pound of

shipments, and the number of destination ZIP codes to which a plant sends shipments divided by the total number of shipments a plant makes (a measure of the location dispersion of a plant's shipments). There are significant changes in all of the metrics except average value per pound across the entire sample, as indicated by the coefficient on the 1997 dummy. However, the changes are systematically different for establishments that are brought into vertical ownership structure only for, curiously, the value per pound of shipments. This possibly reflects a change in the product mix or a quality shift in output within products, though we can't say for sure. There is no indication that plants start shipping to fewer (or more) destinations when they become part of a vertical ownership link.

We look more closely at the destinations of plants' shipments and their changes in panel B of Table 4. Here, we focus on those plants in our CFS panel that become part of vertical ownership structures; this includes 903 plants observed in each of two years. We make the following comparison. We first find all ZIP codes to which the plants make internal shipments in 1997 as parts of vertical ownership structures. We compute the each plant's internal shipment shares in the same way as with the benchmark results above. Next, we compute what fraction of the same plant's shipments was sent to those "internal" ZIP codes in the 1993 CFS. Comparing the fractions across the two years allows us to see if plants are more likely to ship to ZIP codes where their vertically structured firm owns downstream plants in 1997 after the plants become part of those vertical structures. Finding no difference would suggest that the plants were already shipping to their future downstream compatriots before they were ever co-owned. Quantiles of the two distributions—the "internal" shares in 1993 and the actual internal shares in 1997—are shown in panel B.

It does appear that becoming part of a vertical ownership structure is associated with some shift in the destinations of shipments. After integration, a greater fraction of the plants' shipments go to ZIP codes where their firm owns downstream plants. The fraction reporting no shipments to such ZIP codes falls from 57.7 to 44.5 percent, and the median (90<sup>th</sup> percentile) share rises from 0 (22.7) percent to 0.6 (40.8) percent. These are still by and large small internal shares in 1997—smaller than for the benchmark sample, in fact—but it's also clear that there has been some redirection. (Interestingly, the result in panel A above that the number of destination ZIP codes per shipment does not systematically rise indicates that these new internal shipment destinations are not simply added to previous destinations, but instead replace some of them.)

This fact—a modest redirection of shipments toward internal units for some plants that become vertically integrated—is the second of our findings consistent with vertical ownership existing to facilitate intra-firm goods flows along production chains. The other was the small share of upstream plants, on the order of two percent, that are clearly dedicated to producing inputs for their firms’ downstream operations. Yet, like that finding, this result suggests the goods transfer motivation is relatively weak; most newly integrated plants ship very little of their output to their downstream units.

### *C. Plants in Vertical Ownership Structures are High “Type” Plants*

It appears that the lack of movement of goods along production chains within most vertically structured firms is a robust feature of the data. As mentioned above, we propose that vertical ownership is instead typically used to facilitate movements of intangible inputs like management oversight across a firm’s production units. We begin in this section to document additional facts that are consistent with this theory.

We first focus on plant-level measures of “type.” We conceptualize plants’ types as combinations of idiosyncratic demand and supply fundamentals that affect plant profitability.<sup>14</sup> In industry equilibrium, these fundamentals are tied to plant observables like productivity, size, and (in some cases) factor intensities.

We use four such measures in our empirical work. They are not independent, but they differ enough in construction to allow us to gauge the consistency (or lack thereof) of our results. Two are productivity measures. Both measure plant output using the plant’s reported sales deflated by a price index for the plant’s four-digit SIC industry. The productivity metrics differ in their measure of inputs. One is output per worker-hour and the other total factor productivity. (Both are expressed as the log of the plant’s output-input ratio.) Our third type measure is simply the plant’s scale as measured by logged real revenue. The fourth metric is the plant’s logged capital-labor ratio (capital stock per worker-hour). Further details on the construction of these measures are in the Data Appendix. Because of data limitations, we can only construct these measures for the roughly 350,000 plants in each year’s Census of Manufacturers.

These empirical type measures have been shown in various empirical studies to be

---

<sup>14</sup> Foster, Haltiwanger, and Syverson (2008) present a model of industry equilibrium where producers differ along both demand and cost dimensions, and show that plant type can be summarized as a single-dimensional index of demand, productivity, and factor price fundamentals.



correlated with plant survival. Survival probabilities reflect plant type in many models of industry dynamics with heterogeneous producers, like Jovanovic (1982), Hopenhayn (1992), Ericson and Pakes (1995), and Melitz (2003). The productivity-survival link has perhaps been the most extensively studied empirically; see Bartelsman and Doms (2000) for a recent review of this literature. Plant scale and survival was the subject of much of Dunne, Roberts, and Samuelson (1989), and capital intensity's connection to survival was explored in Doms, Dunne, and Roberts (1995).

We first compare plant type measures across integrated and unintegrated producers by regressing plant types on an indicator for plants' integration status and a set of industry-by-year fixed effects. The coefficient on the indicator (which takes the value of one for vertically integrated plants and zero otherwise) captures the average difference between plants in and out of vertical ownership structures. By including fixed effects, we are identifying type differences across plants in the same industry-year, avoiding confounding productivity, scale, or factor intensity differences across industries and time. We estimate this specification for each of the four plant type proxies and report the results in Table 5, panel A.<sup>15</sup>

It is clear that plants in vertical ownership structures have higher types. They are more productive, larger, and more capital intensive. Their labor productivity levels are on average more than 40 percent higher ( $e^{0.347} = 1.416$ ) than their unintegrated industry cohorts. These are sizeable differences; Syverson (2004) found average within-industry-year interquartile logged labor productivity ranges of roughly 0.65; the gaps seen here are almost half of this. Total factor productivity differences, while still positive and statistically significant, are much smaller, at 1.7 percent. Vertical plants are much larger—4.5 times larger—than other plants in their industry in terms of real output. Capital intensities are substantially higher in integrated plants as well, explaining why their labor productivity advantage is so much bigger than the average TFP difference.

A natural question that follows from these results is the causal nature of vertically linked plants' type differences. There are three possibilities, and they are not mutually exclusive. The gaps could reflect the fact that newly built plants under vertical ownership are different than

---

<sup>15</sup> Sample sizes differ across the specifications because not all the necessary variables for construction of each are available for each proxy measure for every plant-year observation. In particular, capital information is not available in the 1963 and 1997 CMs. Below, we will focus on differences among the set of plants with each of the plant-level production measures (except TFP) available.

newly built plants in other ownership structures, and because types are persistent, this is reflected in the broader population. Or it may be that high-type firms that seek to merge new plants into their internal production chains choose plants that already have high types to add to the firm. Finally, becoming part of a vertical ownership structure might be associated with a change in an existing plant's type.

We can separately investigate these possibilities. To see if new vertically structured plants are different than newly built plants in other ownership structures, we regress all new plants' type measures on a dummy for their vertical ownership status and industry-year effects. New plants are defined as those appearing for the first time in the Economic Census. We exclude observations from the 1977 EC because of censored entry.<sup>16</sup> New plants are an important part of the formation of vertically integrated structures in the economy: entering integrated plants accounted for roughly 600,000 employees and \$50 billion in capital stock in a typical EC (respectively, about one-third and one-half of the total employment and capital stocks of new plants in a given EC).

Panel B of Table 5 shows the results. The differences among integrated and unintegrated new plants here are similar to those seen in the broader within-industry comparison discussed above. Labor productivity and capital intensity differences are over 30 percent. The TFP gap is smaller than the labor productivity differences, as before, but in the case of new plants here is about twice the gap among all plants. Scale differences are still quite large, though somewhat less pronounced than the gap seen in the overall sample. Thus many of the dissimilarities in type observed between plants in and not in vertical ownership structures reflect persistent differences present even at the time of the plants' entry.

The second possible source of vertically integrated plants' higher type measures is that firms comprised of high-type, vertically linked plants seeking to expand through merger or acquisition choose to match with unintegrated plants that are already high-type. We test whether or not this is the case in the data by regressing unintegrated plants' type proxies on a dummy indicating if a plant will *become* vertically integrated by the next EC. Again industry-year fixed effects are included. The estimated coefficient on the dummy captures how to-be-vertically-owned plants compare *before integration* to other plants in their industry that do not become

---

<sup>16</sup> A plant's first appearance in the EC is associated with the start of economic activity at its particular locations; i.e., these plants are greenfield entrants. Existing plants that merely change industries between ECs exist in earlier ECs, and as such are not counted as entrants in our sample.

integrated over the same five-year period.

The results, which are in panel C of Table 5, make it clear that soon-to-be integrated plants are different from other unintegrated plants in their industry. They are more productive; their labor productivity advantage is over 20 percent and they have about three percent higher TFP levels. They are also considerably larger and more capital intensive. (Here, as with the capital intensity comparisons above, integrated or to-be-integrated plants have considerably higher levels of both capital stocks and labor hours than unintegrated plants. It is simply that the capital gap is even larger than the labor gap.) Moreover, these differences are slightly smaller than the gaps measured above, but are still of a similar order of magnitude.

Finally, we investigate if becoming part of a vertical ownership structure is associated with unusually high growth in productivity, scale, or factor intensities. We compare changes in these values (computed as five-year differences between ECs) for plants that become vertically integrated to changes over the same period for industry plants remaining unintegrated. Operationally, we regress the growth in plants' type proxies on an indicator dummy for plants that become part of integrated production chains. We again include a full set of industry-year fixed effects to account for industry-specific growth patterns. We must restrict the sample here to continuing plants—i.e., those in both the current and prior ECs—that are not vertically owned in the prior EC.

The results are shown in panel D of Table 5. Becoming vertically part of a vertical structure is associated with four to five percent higher labor productivity growth than for continuing plants that remain unintegrated. Despite faster labor productivity growth, however, there are no statistically significant differences in TFP growth, and the point estimate is actually negative. This divergence between labor and total factor productivity growth reflects the fact that integrating plants see increases in capital intensity over those that stay unintegrated. This relative capital deepening raises the productivity of labor inputs but not the plants' overall factor-neutral efficiency. (We will see below that this capital deepening occurs because the plant experiences both growth in capital and declines in labor.) Interestingly, plants that become integrated do not grow significantly faster than their industry counterparts that remain unintegrated.

Comparing the type disparities in panels B, C, and D to those seen across all plants in panel A suggest that much of the heterogeneity between plants in and out of vertical ownership

structures reflects the effect of differences in the assignment of plant types to integration status. That is, vertically integrated plants are more productive, larger, and more capital intensive primarily because they were either born into integrated structures that way, or because those are the types of unintegrated plants that firms merge into integrated structures. What gaps not accounted for by these underlying differences are closed due to the faster growth in labor productivity, size, and capital intensity experienced by existing plants when they become integrated.<sup>17</sup>

#### *D. Firm Size and Plant Type Differences*

Plants in vertical ownership structures are different. This naturally leads to the question of whether *firms* with vertical structures are different. Figure 2 plots the densities of firm size (logged total employment, since revenue is unavailable outside of manufacturing) for three mutually exclusive and exhaustive sets of multi-establishment firms. One set includes firms with vertical ownership structures.<sup>18</sup> The other two multi-unit organizational structures are single-industry and multi-industry-unintegrated firms. The former are multi-establishment firms that own plants in only one industry. The latter are firms that own establishments in multiple industries, but none of which comprise substantial vertical links as defined above.<sup>19</sup>

The figure reveals that each of the (logged) employment size distributions is unimodal, though they clearly have different central tendencies.<sup>20</sup> Single-industry multi-unit firms are the smallest and have the most symmetric size distribution. Vertically integrated firms are clearly

---

<sup>17</sup> These are of course general patterns across the hundreds of manufacturing industries in our sample. These broad patterns do not imply that the relative importance of these sources of type differences doesn't vary across individual industries. It is quite possible that in certain industries most of the type differences reflect changes that occur when plants become integrated rather than pre-existing type dissimilarities.

<sup>18</sup> Recall that we define vertical ownership at the plant, not firm level. For our purposes here, however, we define a firm as vertically structured if it owns any vertically linked establishments. As a practical matter, most plants in what we call vertically structured firms here are also classified as being in vertical chains by our plant-specific definition.

<sup>19</sup> The distribution of plants across these firm sets is as follows. Over the entire manufacturing sample, multi-unit plants of all types accounted for 19.7 percent of establishments, 71.4 percent of employment, and 87.0 percent of the capital stock. Vertically integrated plants' shares were, respectively, 12.0, 54.1, and 74.8 percent. Multi-unit single-industry plants accounted for 5.7 percent of establishments and 12.8 and 9.1 percent of employment and capital, while multi-industry unintegrated plants comprised 2.0, 4.5, and 3.1 percent.

<sup>20</sup> We only plot the 1997 distributions rather than those pooled across years in order to remove any secular shifts in firm sizes. Checks of other years show similarly shaped distributions. We have also trimmed the extreme tails of each distribution (10 to 15 or so of the smallest and largest firms in each organizational form category) for data confidentiality reasons.

the largest on average, and their distribution is more skewed than the other firm types. (While not plotted, single-establishment firms are smaller than the multi-unit single-industry firms, as one might expect.) Thus not only are vertically integrated plants larger, their firms are as well.

Given that firms with vertical structures tend to be the largest, it's natural to ask whether the differences in plant types seen above simply reflect underlying differences in *firms*. That is, if large firms tend to own systematically larger (and more productive, etc.) plants, this might explain the distinctive type patterns of plants in vertical structures, rather than their vertical ownership linkages per se. That is, perhaps the high types of plants in vertical ownership structures are a function of firm size rather than firm structure.

To see if this is the case, we rerun the plant type regressions above while including controls for firm size. We regress plant type measures on an indicator for vertically integrated plants and industry-year dummies as above, while now adding flexible controls for firm size. These controls are quintics of logged firm employment, the firm's logged number of establishments, and the logged number of industries in which the firm operates. We restrict the sample here to plants owned by multi-unit firms (the same sample whose firm size distributions are discussed above), but few differences are seen if single-establishment firms are also included. In effect, this specification lets us compare plants in the same industry that are in firms of the same size, regardless of the firms' internal structures.

The results of these regressions are in Table 6. Once we control fully for firm size, much of the correlation between a plant's type and its vertical ownership structure goes away. Indeed, if anything, vertical plants have slightly lower TFP and output levels than others in their industry owned by similarly sized firms. This is a particularly striking result given the enormous average plant size difference between vertical and non-vertical plants documented above. Two type "premia" remain after we control for firm size: labor productivity and capital intensity. Each is on the order of four percent, a fraction of the initial 35 and 45 percent differences.

Hence, much of what makes plants in vertical ownership structures different isn't really related to vertical ownership itself, but rather the facts that the largest plants tend to be in the largest firms, and the largest firms tend to own vertically linked plants. Once this is accounted for, their TFP and output leads disappear and only relatively small differences in labor

productivity and capital intensity remain.<sup>21</sup>

## V. Discussion

We documented several facts about plants in vertical ownership structures. First, raising the primary puzzle of our inquiry, we find that most shipments from upstream plants in vertically structured firms do not stay within the firm. The vast majority of such plants ship either nothing or only very small fractions of their output to downstream units in their firm. This result, which is robust to a number of measurement approaches, is not in accordance with the common presumption that vertical ownership is used to facilitate the movement of goods along a production chain within a firm.

Second, we show that a plant that is part of a vertical ownership has higher levels of several plant type measures than others plants in its industry. These differences primarily reflect persistent dissimilarities existing either at plant birth (if born into a vertical structure) or before the plant becomes part of an integrated firm. Changes that happen when an unintegrated plant becomes part of a vertical link within a firm play a smaller role. However, we also find that most of the observed type differences are not related to plants' vertical ownership structure. Rather, they reflect the fact that vertical plants tend to be in large firms, and large firms of all organizational forms have high-type plants.

While the first set of results is not consistent with vertical ownership (typically) being used to facilitate the movement of physical inputs along a production chain, we propose that the second set of results offers insight into an alternative explanation: namely, that vertical ownership allows efficient intra-firm transfers of *intangible* inputs. Managerial oversight and planning strike us as among the most important of these intangibles, but marketing/sales know-how or other information goods can also be readily transferred among integrated establishments in a firm. We discuss this alternative here.

The results that vertically integrated plants have different “types”—but that those differences are largely explained by firm size rather than structure—are consistent with theories of the firm as the outcome of an assignment mechanism. These models, like Lucas (1978),

---

<sup>21</sup> This evokes the result in Hortaçsu and Syverson (2007) that vertically integrated ready-mixed concrete plants' productivity and survival advantages don't reflect their vertical structure per se, but rather that these plants tend to be owned by firms with clusters of ready-mixed plants in local markets. (The clusters allow them to harness logistical efficiencies.) Once we compared vertically integrated concrete plants to non-integrated plants that were also in clusters, many of the differences seen between integrated and nonintegrated plants disappeared.

Rosen (1982), and more recently by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007) imply assortative matching in the presence of complementarities among inputs.<sup>22</sup> That is, higher-quality intangible inputs (e.g., better managers) are spread across better and/or a greater number of production units. The highest-quality intangible inputs are allocated to multiple establishments in distinct product categories (each among the highest types within their industry), some of which are vertically linked. The end result is what we document in the data: vertically integrated production chains are found in the largest firms composed of the highest-type plants. This firm matching/sorting implication is also supported by our results that plants that will become parts of vertical ownership structures *already* have considerably higher type measures than other plants in the industries. Firms with high-type plants seek out other high-type plants to bring into the fold. It's also consistent with the fact that (not reported here for space reasons) plants' types within firms are positively correlated; firm's with high-type plants in one industry tend to have high-type plants in their other industries.

Note that if this alternative explanation for vertical ownership is correct, the distinction between "downstream" and "upstream" becomes one of convenience rather than an accurate depiction of intra-firm transfers. Managerial, marketing, or other similar inputs are just as likely to be transferred from a firm's downstream units to its upstream ones as the reverse. The names reflect the flow of the physical production process, which may be nonexistent or otherwise very small; they do not necessarily indicate the flow of inputs within the firm.

And in fact, as we saw in the fact that most plant type differences are explained by , verticality itself may not be an important distinction. Vertical firm expansions may not be altogether different than horizontal expansions. When a firm expands horizontally, it usually begins operating in markets that are new but still near to its current line(s) of business, under the expectation that its current abilities can be carried over into the new markets. Physical goods transfers among the firm's establishments are not automatically assumed in such expansions, though inputs like management and marketing are expected to flow to the new units. Many vertical expansions may operate the same way. The industries immediately upstream and downstream of a firm's current operations are, almost by definition, related lines of business. Firms will occasionally expand into these lines under the expectation that their current capabilities will prove useful in the new markets. Transfers of managerial and other non-

---

<sup>22</sup> These models are in turn built on foundations laid out earlier by Koopmans' (1951) and Becker (1973).

tangible inputs will be made to the new establishments, yet no physical good transfers from upstream to downstream establishments need occur.

#### *A. Additional Evidence the Vertical Structures Facilitate Intangible Input Transfers*

While we were able to check the robustness of our results above in many ways, we do not have the best data for testing our “intangible input” explanation for vertical ownership structures. Ideally, we would have information on the application of managerial or other intangible inputs (like managers’ time-use patterns across the different business units of the firm). This would let us see the flows of information-based inputs within a firm and how they change with firm structure, just as we can now with physical goods in the CFS. Unfortunately, we do not have such data, and it almost surely does not exist for the breadth of industries which we are looking at here. That said, we briefly take a few approaches at garnering evidence in this section, though we stress that it is only suggestive.

We first look at the internal shipment patterns of a very select group of establishments in vertical ownership structures. Specifically, these are newly vertically integrated establishments on the upstream end of a production chain (not just newly integrated, in fact, but newly multi-unit plants—they were single-unit firms in the previous Economic Census). They have also been bought by firms that, concurrent with the purchase, begin owning plants in a vertical production chain for the first time. That is, these are the establishments that *make* these firms vertically structured. We expect that these establishments might provide one of the clearest windows into the reasons why firms expand vertically. Because of the narrow selection definition, the number of these establishments in the CFS is small—a total of 622 establishment-years, reporting 58,622 shipments—but the subsample still offers enough leverage to make a meaningful comparison to the overall patterns discussed above.

The results for this group of establishments indicate the prevalence of internal shipments is even lower for this group than for the entire sample. The median fraction of internal shipments is 0.4 percent, and 46 percent report no internal shipments at all. Because the small sample raises questions of whether these differences are statistically significant, we also conduct regression comparisons that project plants’ intra-firm shipment shares on an indicator for these new-VI establishment/firm units and full set of industry-year dummies. The estimated coefficient on the subsample indicator in the dollar-value-share regressions is -0.035 (s.e. =



0.009). (The coefficient is also significantly negative when shares of shipment counts or weights are used as the dependent variable.) Thus these establishments do in fact have significantly lower internal shipments shares.

These results indicate that even for establishments that expressly bought as part of a firm's move to build a vertically integrated ownership structure, internal sourcing of physical inputs is not widespread practice, also consistent with other motives playing an important, and perhaps dominant, role driving the integration decision.

Our second test digs deeper into the changes seen in plants that become vertically integrated, as with those observed in panel D of Table 5. We saw there that the only significant changes in type measures observed for such plants were in labor productivity and capital intensity. (These are also the only two significant differences that remain between integrated and nonintegrated plants in the cross section once we fully control for firm size.) It's not surprising that these two measures are positively correlated; higher capital intensity implies more output per unit labor in any production technology where capital and labor are complements.

We decompose these changes into changes in their respective components by repeating the exercises in Table 2, panel D, but this time running the specifications separately for plants' capital stocks and labor stocks. So we can exactly decompose these changes, we restrict the sample to sets of plants for which we observe each of the production measures. This way, the changes in the ratios' (logged) components will add to the change in ratios. Furthermore, for reasons that will become clear momentarily, we look at the individual changes in two types of labor inputs: production and nonproduction workers. The results are shown in Table 7.

The 4.8 percent labor productivity growth in this sample is driven both by a 2.4 percent increase in output (unlike the sample as a whole, which saw no significant change) and by an equally sized decline in hours. The increase in capital intensity is sourced in both a dose of investment—capital stocks in newly vertical establishments grow 6.7 percent faster than they do in plants in the same industries that remained unintegrated—and through the same 2.4 percent decrease in labor inputs.

What is most interesting about this change in labor inputs is the change in labor composition that accompanies it. Nonproduction worker counts fall more than do production workers; in fact, the change in the latter is not statistically significant. If we also look at the share of nonproduction workers in the plants' total employment, this also falls significantly when

the plant becomes vertically integrated.

These changes in capital intensity and labor composition are consistent with an intangible inputs motive for vertical ownership. Capital intensity would rise upon a plant becoming part of a vertical link if skilled managerial or other intangible inputs have stronger complementarities with capital than labor, for example. This would be expected in the allocation mechanism we discuss above.<sup>23</sup> As for the labor composition effect, the relative decline in nonproduction workers upon integration is consistent with some of the plant's former management, marketing, sales, or any staff associated with providing intangible inputs, being replaced with the new intangible inputs of the vertically integrated structure. Fewer workers are needed to provide these new inputs in the integrated structure because of centralization and scale returns and/or higher efficacy.

The results in this section, as we discussed above, are only suggestive. We cannot observe workers' positions within the firm at any finer level than the production/nonproduction worker dichotomy, and would need much more detailed information on managerial or other intangible inputs to test the theory convincingly. Still, we find the results an intriguing starting point for continued work.

## VI. Conclusion

We have used data on millions of plants, the organizational structure of firms that own them, and their shipments, to explore production behavior in vertical ownership structures. We find that the common view of vertical ownership supporting efficient intra-firm transfers of goods along a production chain may not be its primary purpose. Firms' upstream plants ship only a fraction (and often none) of their output to downstream units inside the firm. This finding is robust to a number of measurement methods. Thus, outside of some exceptional plants that we find are clearly dedicated to internal production, most vertical ownership appears to have a different motive.

Motivated by patterns we document in plants' "types" within and across firms, we propose an alternative explanation for vertical integration. Namely, that it facilitates efficient

---

<sup>23</sup> Firms with vertical ownership structures might also face lower effective capital costs, which would shift their optimal factor allocation toward a more capital-intensive orientation. Since we know vertical firms are larger on average, and there is evidence that larger firms might be less credit constrained (e.g., Fazzari, Hubbard, and Petersen (1988) and Eisfeldt and Rampini (forthcoming)), this is a plausible alternative.

transfers of *intangible* inputs (e.g., managerial oversight) within firms. It is plausible that the market would have a more difficult time mediating transfers of knowledge inputs than of physical goods. Unfortunately, the data only allow enough leverage to gather suggestive, rather than conclusive, evidence on this point.

Note that if this explanation is correct, there may not be anything particular about vertical structure within firms; intangible inputs can flow in any direction across a firm's production units. Vertical firm structures and expansions may not be fundamentally different from horizontal structures and expansions. Instead, a more generalized view of firm organization, like the firm as an outcome of an assignment mechanism that matches heterogeneous tangible and intangible inputs, may be warranted, and is consistent with some of the other patterns we document in the data.

One interesting point of comparison between our findings and the existing literature is with regard to international trade flows. For countries where such data is available, intra-firm trade accounts for roughly one-third of international goods shipments (see, e.g., Bernard, Jensen, Redding, and Schott (2007)). This is clearly substantially larger than the modest, domestic within-firm shipment volume we document here. A possible explanation might be that multinational firms are more likely to be comprised of the types of plants at the right tail of our Figure 1: large, dedicated producers to the firms downstream plants. Why multinationals would choose to structure themselves in a way so different than domestic shippers is less clear; we see this question as a good launching point for further research.

## Data Appendix

We describe here details on the construction of our production variables.

*Output.* Plant output is its inventory-adjusted total value of shipments, deflated to 1987 dollars using industry-specific price indexes from the NBER Productivity Database.

*Labor Hours.* Production worker hours are reported directly in the CM microdata. To get total plant hours, we multiply this value by the plant's ratio of total salaries and wages to production worker wages. This, in essence, imputes the hours of non-production workers by assuming that average non-production worker hours equal average production worker hours within plants.

*Labor Productivity.* We measure labor productivity in terms of plant output per worker-hour, where output and total hours are measured as described above.

*Total Factor Productivity.* We measure productivity using a standard total factor productivity index. Plant TFP is its logged output minus a weighted sum of its logged labor, capital, materials, and energy inputs. That is,

$$TFP_{it} = y_{it} - \alpha_l l_{it} - \alpha_k k_{it} - \alpha_m m_{it} - \alpha_e e_{it},$$

where the weights  $\alpha_j$  are the input elasticities of input  $j \in \{l, k, m, e\}$ . Output is the plant's inventory-adjusted total value of shipments deflated to 1987 dollars. While inputs are plant-specific, we use industry-level input cost shares to measure the input elasticities. These cost shares are computed using reported industry-level labor, materials, and energy expenditures from the NBER Productivity Database (which is itself constructed from the CM). Capital expenditures are constructed as the reported industry equipment and building stocks multiplied by their respective BLS capital rental rates in the corresponding two-digit industry.

*Real Materials and Energy Use.* Materials and energy inputs are plants' reported expenditures on each divided by their respective industry-level deflators from the National Bureau of Economic Research Productivity Database.

*Capital-Labor Ratio.* Equipment and building capital stocks are plants' reported book values of each capital type deflated by the book-to-real value ratio for the corresponding three-digit industry. (These industry-level equipment and structures stocks are from published Bureau of Economic Analysis data.) Any reported machinery or building rentals by the plant are inflated to stocks by dividing by a type-specific rental rate.<sup>24</sup> The total productive capital stock  $k_{it}$  is the sum of the equipment and structures stocks. This is divided by the plants' number of labor hours to obtain the capital-intensity measure used in the empirical tests.

---

<sup>24</sup> Capital rental rates are from unpublished data constructed by the Bureau of Labor Statistics for use in computing their Multifactor Productivity series. Formulas, related methodology, and data sources are described in U.S. Bureau of Labor Statistics (1983) and Harper, Berndt, and Wood (1989).

*Nonproduction Worker Ratio.* Plants directly report both their number of production and nonproduction employees. Nonproduction workers are defined by the Census Bureau as those engaged in “supervision above line-supervisor level, sales (including a driver salesperson), sales delivery (truck drivers and helpers), advertising, credit, collection, installation, and servicing of own products, clerical and routine office functions, executive, purchasing, finance, legal, personnel (including cafeteria, etc.), professional and technical [employees]. Exclude proprietors and partners.” The nonproduction worker ratio is simply such employees’ share of total plant employment.

## References

- Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti. "Vertical Integration and Technology: Theory and Evidence." NBER Working Paper 10997, 2004.
- Acemoglu, Daron, Simon Johnson, and Todd Mitton. "Integration: Finance, Contracts, and Regulation." NBER Working Paper 11424, 2005.
- Antras, Pol. "Firms, Contracts, and Trade Structure." *Quarterly Journal of Economics*, 118(4), 2003, 1375-1418.
- Arrow, Kenneth J. "Vertical Integration and Communication." *Bell Journal of Economics*, 6(1), 1975, 173-183.
- Bartelsman, Eric J. and Mark Doms. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*, 38(3), 2000, 569-95.
- Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. "Intra-Firm Trade." Mimeo, 2007.
- Carlton, Dennis W. "Vertical Integration in Competitive Markets Under Uncertainty." *Journal of Industrial Economics*, 27(3), 1979, 189-209.
- Carlton, Dennis W. and Jeffrey M. Perloff. *Modern Industrial Organization, Fourth Edition*. Boston: Pearson Addison Wesley, 2005.
- Coase, R. H. "The Nature of the Firm." *Economica*, 16(4), 1937, 386-405.
- Doms, Mark, Timothy Dunne, and Mark J. Roberts. "The Role of Technology Use in the Survival and Growth of Manufacturing Plants." *International Journal of Industrial Organization*, 13(4), 1995, 523-542.
- Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. "The Growth and Failure of U. S. Manufacturing Plants." *Quarterly Journal of Economics*, 104(4), 1989, 671-698.
- Eisfeldt, Andrea and Adriano Rampini. "Leasing, Ability to Repossess, and Debt Capacity." *Review of Financial Studies*, forthcoming.
- Ericson, Richard and Ariel Pakes. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62(1), 1995, 53-82.
- Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen. "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity*, 1, 1988, 141-95.
- Foster, Lucia, John Haltiwanger, and Chad Syverson. "Reallocation, Firm Turnover, and

- Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1), 2008, 394-425.
- Garicano, Luis and Thomas N. Hubbard. "The Return to Knowledge Hierarchies." NBER Working Paper 12815, 2007.
- Garicano, Luis and Esteban Rossi-Hansberg. "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics*, 121(4), 2006, 1383-1435.
- Gollop, Frank M. and James L. Monahan. "A Generalized Index of Diversification: Trends in U.S. Manufacturing." *Review of Economics and Statistics*, 73(2), 1991, 318-30.
- Grossman, Sanford J. and Oliver D. Hart. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94(4), 1986, 691-719.
- Harper, Michael, Ernst R. Berndt, and David O. Wood. "Rates of Return and Capital Aggregation using Alternative Rental Prices." in *Technology and Capital Formation*, Dale W. Jorgenson and Ralph Landau (eds.), MIT Press, 1989.
- Hillberry, Russell and David Hummels. "Intranational Home Bias: Some Explanations." *Review of Economics and Statistics*, 85(4), 2003, 1089-92.
- Hillberry, Russell and David Hummels. "Trade Responses to Geographic Frictions: A Decomposition Using MicroData." *European Economic Review*, 52(2), 2008, 527-550.
- Hopenhayn, Hugo A. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5), 1992, 1127-50.
- Hortaçsu, Ali and Chad Syverson. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy*, 115(2), 2007, 250-301.
- Joskow, Paul. "Vertical Integration." in Claude Menard and Mary Shirley, eds., *Handbook of New Institutional Economics*. Berlin: Springer/Kluwer, 2005.
- Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica*, 50(3), 1982, 649-70.
- Lafontaine, Francine and Margaret Slade. "Vertical Integration and Firm Boundaries: The Evidence." *Journal of Economic Literature*, 45(3), 2007, 629-85.
- Lucas, Robert E. "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2), 1979, 508-23.
- Melitz, Marc J. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6), 2003, 1695-1725.

- Perry, Martin K. "Vertical Integration: Determinants and Effects." in Richard Schmalensee and Robert D. Willig, eds., *Handbook of Industrial Organization*. Amsterdam: North-Holland, 1989.
- Rosen, Sherwin. "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13(2), 1982, 311-323.
- Salop, Steven C. "Vertical Mergers and Monopoly Leverage." in Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*, London: Macmillan, 1998.
- Stigler, George. "The Division of Labor Is Limited by the Extent of the Market." *Journal of Political Economy*, 59(1), 185-193.
- Syverson, Chad. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, 86(2), 2004, 534-550.
- Teece, David J. "Towards an Economic Theory of the Multiproduct Firm." *Journal of Economic Behavior and Organization*, 3(1), 1982, 39-63.
- Tirole, Jean. *The Theory of Industrial Organization*. Cambridge, MA: MIT Press, 1988.
- U.S. Bureau of Labor Statistics. "Trends in Multifactor Productivity: 1948-81." Bulletin 2178. Washington D.C., Government Printing Office, 1983.
- U.S. Census Bureau. "Instructions for Completing the Commodity Flow Survey." Washington D.C., Government Printing Office, 1997.
- Williamson, Oliver E. "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics*, 22(2), 1979, 233-261.



Table 1. Plant-Level Shares of Internal Shipments

A. Benchmark

Internal share of:	Percentile					Fraction = 0	Fraction = 1
	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>		
Plant shipment counts	0%	0%	3.8%	18.8%	51.8%	33.6%	2.1%
Plant dollar value of shipments	0	0	2.6	19.2	58.0	33.6	2.1
Plant total weight of shipments	0	0	2.6	20.0	58.5	33.6	2.1

Note: These tables report shares upstream plants' shipments that are internal to their firm. The sample consists of 29,931 plant-years aggregated from 2,826,296 shipments. For data confidentiality reasons, the reported percentiles are averages of immediately surrounding percentiles; e.g., the median = 0.5\*(49<sup>th</sup> percentile + 51<sup>st</sup> percentile).

B. Robustness Checks (Share of Dollar Value Shown)

Specification/Sample	Percentile					Frac. = 0	Frac. = 1
	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>		
1. At least median number of shipments	0%	0%	3.1%	17.4%	48.4%	29.4%	0.7%
2. No exporters	0	0	2.7	21.0	64.9	35.7	2.9
3. No wholesalers	0	<0.1	5.4	26.6	69.0	24.0	2.7
4. Shipments to any plant in firm are internal	0	0.4	6.5	27.7	67.2	19.7	2.5
5. County, not ZIP, determines internal	0	2.6	28.0	68.3	96.3	17.1	6.3
6. 25 least differentiated industries	0	0.6	6.1	16.5	39.3	17.2	1.4
7. 10% cutoff definition for VI	0	0	0.8	13.4	46.2	42.7	1.6
8. 50% cutoff definition for VI	0	0	0	11.1	43.1	53.2	1.5

Notes: Each row shows for a different subsample the distributions of the shares (by dollar value) of upstream integrated establishments' shipments that are internal to the firm. The criteria for inclusion in and size of each subsample is discussed in the text. For data confidentiality reasons, the reported percentiles are averages of immediately surrounding percentiles; e.g., the median = 0.5\*(49<sup>th</sup> percentile + 51<sup>st</sup> percentile).

Table 2. Internal Shipments as Share of Smaller of Upstream Shipments or Downstream Usage

Downstream usage measure	Value share of shipments percentiles				
	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>
Firm's total downstream manufacturing materials purchases	0%	0%	2.7%	24.1%	73.8%
Firm's downstream use of 2-digit product	0	0	1.7	44.8	125.5
Firm's downstream use of 4-digit product	0	0	3.5	72.8	267.5

Note: These tables report shares upstream plants' shipments that are internal to their firm, as a fraction of the smaller of a) the total shipments of a firm's upstream plants or b) the firm's downstream use of a product. Sample construction and sizes are detailed in the text. For data confidentiality reasons, the reported percentiles are averages of immediately surrounding percentiles; e.g., the median = 0.5\*(49<sup>th</sup> percentile + 51<sup>st</sup> percentile).

Table 3. Relative Material Use Intensity Ranks between Plants in Vertical Ownership Structures and Other Plants

		Material's intensity rank in non-vertically linked plants				
		1	2	3	4	5
Material's intensity rank in vertically linked plants	1	56.9	14.4	7.2	4.1	3.4
	2	14.5	29.1	16.5	10.6	6.1
	3	7.7	16.3	20.1	14.4	10.2
	4	4.6	10.7	13.1	15.1	11.3
	5	3.7	6.9	6.3	13.8	13.2
	6	2.6	4.0	6.3	8.6	11.0
	7	2.2	4.5	5.3	5.8	9.6

Notes: This table shows, for a sample of 8716 industry-material-year cells, the ranks of materials intensity use (by share of materials purchases) for the five most intensively used materials in plants we define as not in vertical ownership structures. The entries in the table correspond to the fraction of cells where vertical and non-vertical plants in the same industry share a particular pair of materials intensity rankings. For example, across all industry-years in the sample, the most intensively used (rank 1) material by non-vertical plants in an industry-year is the most intensively used material by the industry-year's vertical plants 56.9 percent of the time. Non-vertical plants' rank 1 material is vertical plants' second most intensively used material 14.5 percent of the time, and so on.

Table 4. Changes in Shipments upon Vertical Ownership

A. Shipment Metrics in 1997

Shipment Metric	Mean	<i>I</i> [1997]	<i>I</i> [vertical]
ln(total \$ of shipments)	17.46	0.415* (0.014)	-0.224* (0.026)
ln(avg. \$ per ship)	7.689	0.096* (0.014)	0.049 (0.029)
ln(avg. lbs. per ship)	7.235	0.082* (0.017)	-0.007 (0.034)
ln(avg. miles per ship)	5.283	-0.045* (0.010)	0.027 (0.020)
ln(value-per-pound)	0.454	0.015 (0.011)	0.055* (0.020)
destination ZIP codes ÷ number of shipments	0.541	0.007* (0.003)	0.005 (0.005)

Notes: This table shows outcomes of regressing various shipment metrics on indicators for 1997 and for plants in vertical ownership structures. The sample consists of plants in both the 1993 and 1997 CFS and that were not in vertical ownership structures in 1993;  $n = 12,496$  plant-years. Thus the vertical indicator can only be equal to one in 1997; it shows whether differences in shipment metrics between 1993 and 1997 are systematically different for plants that entered vertical structures.

B. Shipments to “Internal” ZIP Codes, 1993 and 1997

	Percentile					Frac. = 0	Frac. = 1
	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>		
Share of \$ value, 1993	0%	0%	0%	5.0%	22.7%	57.7%	0.7%
Share of \$ value, 1997	0	0	0.6	10.1	40.8	44.5	1.5

Notes: This table shows, for the sample of plants in both the 1993 and 1997 CFS that were not part of vertical ownership structures in 1993 but were by 1997, quantiles of the distribution of these plants’ “internal” shipments. These are defined in 1993 as those to ZIP codes that will be internal in 1997—i.e., ZIP codes where the plant’s firm will own a downstream plant in 1997). In 1997 these are the standard measure of shipments to ZIP codes where the firm currently owns a downstream plant.

Table 5. Plant Attributes by Vertical Ownership Structure

	Output per hour	TFP	Output	Capital-labor ratio
A. Within-industry differences				
N	1,048,887	739,366	1,073,978	787,283
Indicator for vertical plants	0.348* (0.002)	0.017* (0.001)	1.509* (0.004)	0.443* (0.003)
B. Differences among new plants				
N	236,228	152,618	246,464	166,272
Indicator for vertical plants	0.285* (0.004)	0.032* (0.003)	1.251* (0.010)	0.368* (0.007)
C. Comparing unintegrated plants: to-be-vertical vs. remaining non-vertical				
N	648,449	578,634	667,727	621,131
Indicator for to-be-vertical plants	0.215* (0.004)	0.027* (0.003)	1.416* (0.009)	0.245* (0.006)
D. Changes upon entering vertical ownership				
N	481,777	303,775	496,688	331,854
Newly vertical indicator	0.045* (0.005)	-0.011 (0.006)	-0.012 (0.007)	0.089* (0.010)

Notes: This table shows plant “type” comparisons between plants in (or to-be-in) vertical ownership structures and their non-vertical counterparts. Panel A compares across all plants for which type measures are available. Panel B compares new plants. Panel C compares *prior period* types among non-vertical plants that will become part of vertical ownership structures by next period to those remaining non-vertical. Panel D compares changes in type for plants that become part of vertical ownership structures to changes for unintegrated plants that remain so. All regressions include industry-year fixed effects. Samples are comprised of non-AR manufacturing plants. See text and data appendix on construction of type measures and additional details. An asterisk denotes significance at a five percent level.

Table 6. Plant Type Differences Controlling for Firm Size

	Output per hour	TFP	Output	Capital- labor ratio
<u>Multi-unit firm dummy</u>				
N	1,048,887	739,366	1,073,978	787,283
VI indicator	0.257* (0.003)	0.014* (0.002)	0.886* (0.007)	0.326* (0.005)
Multi-industry indicator	0.101* (0.003)	0.003 (0.002)	0.693* (0.007)	0.131* (0.005)
<u>Flexible controls for firm size</u>				
N	261,982	198,096	266,981	205,462
VI indicator	0.043* (0.005)	-0.011* (0.003)	-0.020* (0.009)	0.037* (0.007)

Notes: This table shows the results from regressing plant-level type measures on an indicator for vertically integrated plants, a set of industry-year fixed effects, and controls for firm size. The firm size controls are a dummy for single-industry firms and quintics of several measures of the plant's owning-firm size: employment, the number of establishments, and number of industries.

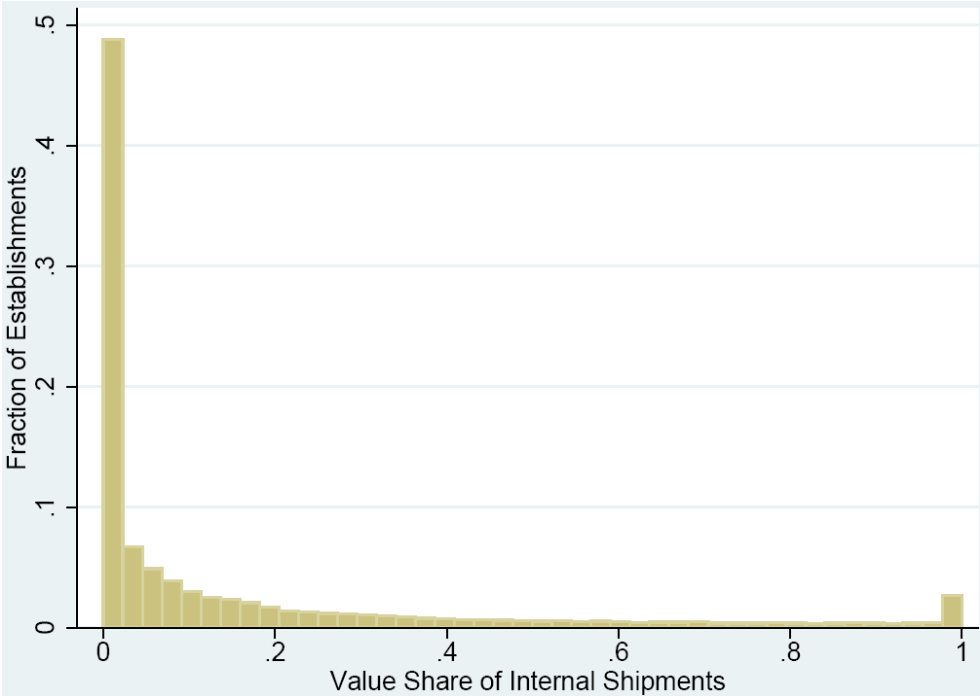
Table 7. Changes in Plant Attributes Upon Integration

	Change upon VI
Output per hour	0.048* (0.006)
Output	0.024* (0.007)
Capital-labor ratio	0.091* (0.010)
Capital	0.067* (0.010)
Hours	-0.024* (0.007)
Production workers	-0.011 (0.007)
Nonproduction workers	-0.038* (0.008)
Nonproduction worker share	-0.005* (0.002)

Note: The table shows specifications repeating the exercises in panel D of Table 2, but with additional plant production measures included. Further, the sample consists of only those 282,240 newly integrated plants that have nonmissing data for all production measures. See text for details of the construction of the variables. All regressions include industry-year fixed effects. An asterisk denotes significance at a five percent level.

Figure 1. Shares of Intra-firm Shipments by Upstream Vertically Integrated Establishments

A. Internal Shipments Defined as Those to ZIP codes of Firm's Downstream Establishments



B. Internal Shipments Defined as Those to ZIP codes of All Firm's Establishments

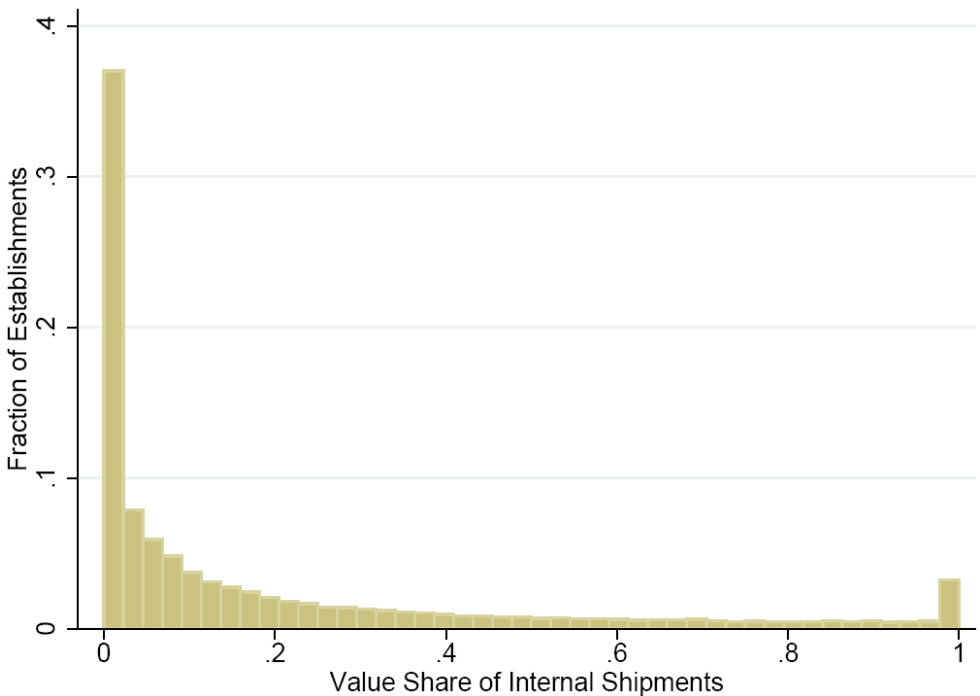
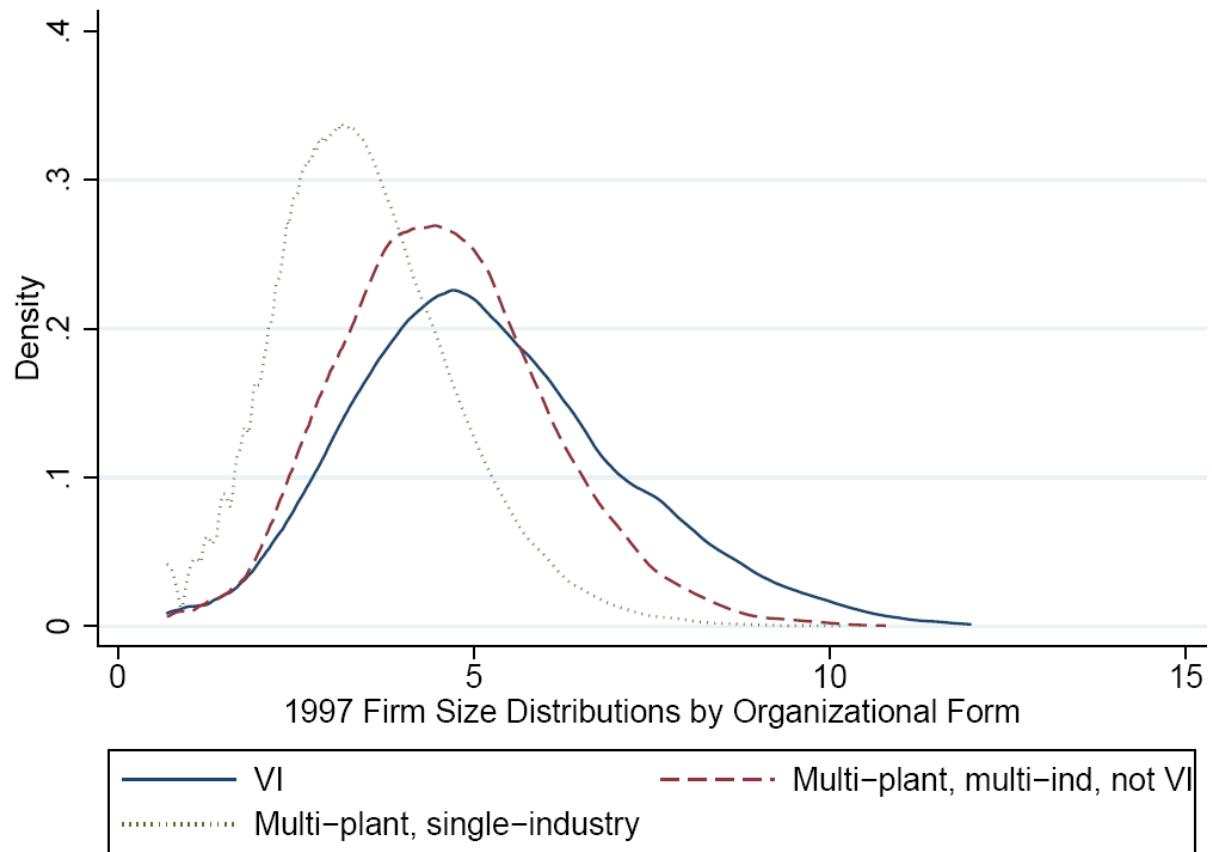




Figure 2. Firm Size Distributions by Organizational Structure



Notes: This figure shows density plots of the firm size distributions (measured by logged total employees) for the three types of multi-establishment firms. See text for details.