

VERY PRELIMINARY AND INCOMPLETE.
Please do not cite without permission.

Vertical Integration and Production: Some Plant-Level Evidence^{*}

Ali Hortaçsu
University of Chicago
and NBER
hortacsu@uchicago.edu

And

Chad Syverson
University of Chicago
and NBER
syverson@uchicago.edu

July 2007

^{*} Very preliminary and incomplete; please do not cite without contacting the authors. We thank Daron Acemoglu, Austan Goolsbee, and Tom Holmes for helpful discussions and seminar participants at the Chicago Micro Lunch for their comments. Margaret Triyana provided excellent research assistance. Syverson appreciates funding from the NSF (award no. SES-0519062), the John M. Olin Foundation, and the Stigler Center. 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. Both authors can be contacted at the Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637.

Abstract

We use two very large datasets to obtain some of the first systematic micro evidence on how production in vertically integrated firms differs from that of unlinked producers in the same industries. The Economic Census lets us observe virtually every establishment in the private economy, and crucially, the organizational structures of the firms that own them. The Commodity Flow Survey microdata offer a random sample of plants' shipments, offering an unusually comprehensive glimpse into patterns of goods flows within firms. We find that overall levels of vertical integration were relatively stable during our 1977-97 sample period, though they declined somewhat after 1987. Within industries, vertical integration is related to differences in productivity, size, and capital intensity. These differences mostly embody persistent differences in the plants that are launched by or brought into vertically integrated firms, but to some extent also reflect changes that formerly unintegrated plants experience upon integration. Matching of plant "types" appears to be important in vertical integration: the correlation in these plant-level measures is stronger in firms with vertical structures than those that are purely horizontally structured. In terms of goods flows, we find that surprisingly few shipments of upstream integrated establishments go to any other plant in their firm. Across integrated plants, the modal fraction of their value of shipments that is intra-firm is zero; the median is 2.6 percent. We interpret this evidence as suggesting that vertical integration is often about transfers of intangible inputs like managerial oversight rather than physical ones. To the extent that such intangible inputs are complementary to physical inputs, equilibrium assignment entails the allocation of the best intangible inputs across multiple establishments in vertically related product categories.

I. Introduction

A large literature has explored the causes and effects of vertical integration. This literature has pointed to a variety of factors being related to integration, including asset specificity, supply uncertainty, market power, incomplete contracting, transaction costs, and regulation.¹ Despite this, there is little systematic micro-level evidence on the ways production in vertically integrated firms differs from that of unintegrated producers in the same (narrowly defined) industries. This applies not only to differences in hard-to-measure attributes like capital specificity, contracting environments, and transaction costs, but even to more basic features like scale, factor intensity, and productivity.

We hope to begin to fill this gap with this paper. We use data on millions of plants—and the organizational structure of firms that own them—to show how being part of a vertically integrated structure is reflected in plant and firm attributes. We also use data on millions of individual shipments from a random sample of these plants to gain an unusually comprehensive view of the flow of goods within integrated production chains across the economy. This data includes information on the shipments' establishment of origin, their destination ZIP codes, the products being shipped, and their values, weights, and modes of transportation.

We find that the aggregate extent of vertical integration was fairly stable over our 1977-1997 sample period, though it dropped somewhat after 1987. Firms with vertically integrated structures are considerably larger on average than other firms, even those that operate in multiple (but not vertically linked) industries. In one of the strongest patterns observed in the data, we also find that *establishments* in vertically integrated structures are systematically larger.²

When we focus more extensively on the manufacturing sector, which affords a larger amount of plant-level production data, we find that vertical integration is strongly related to other within-industry differences besides size. These include differences in productivity measures and

¹ The size of the literature precludes comprehensive citation. Surveys include Perry (1989), Salop (1998), Joskow (2005), and Lafontaine and Slade (2007). Coase (1937) is often cited as the seminal paper of the literature. 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).

² An establishment is a unique location where production takes place, like a store, warehouse, or factory (indeed, we will often refer to establishments as “plants” in this paper, since much of the empirical work focuses on the manufacturing sector). Note that the fact that integrated *establishments* are larger is not necessarily implied by the finding that *firms* with integrated structures are larger. Firms with vertical structures could in principle be large only via the extensive margin; i.e., vertically integrated establishments could be the same size (or even smaller) than unintegrated ones, but firms with integrated structures could simply own more plants.

capital intensity. We demonstrate that these differences primarily embody persistent differences in the plants that are started by or brought into firms with vertical structures, but to some extent they also reflect changes that formerly unintegrated plants experience upon integration. We also document that the correlations of these plant-level measures, which we interpret as embodying fundamental differences in plant “type,” are stronger in firms with vertical structures than those that are purely horizontally structured.

When we turn our attention to the patterns of goods flows within integrated production chains, we find that surprisingly little output of upstream establishments in vertically integrated structures is shipped to plants in the same firm. (This is true whether we restrict internal destinations to downstream plants in the firm or instead use the broader criterion of calling a shipment internal if it goes to *any* plant in the firm). The modal fraction of these upstream establishments’ shipments that are internal to their firms is zero; one-third them of report *no* shipments to other establishments in their firm. The median fraction of internal shipments across plants is 3.8 percent, if shipments are counted equally; only 2.6 percent are internal in terms of total dollar values or weight. There is some evidence of bimodality: about 2 percent of upstream plants are exclusively dedicated to serving the needs of their firms. However, these internal specialists comprise a share of integrated producers that is only a fraction of the one-third reporting no internal shipments. Further, even the 90th percentile internal shippers are hardly dedicated makers of inputs for their firms’ downstream operations: 42 percent of the value of their shipments is sent outside the firm.

We draw two conclusions about any theoretical explanations for the facts we document. First, because these patterns hold across such a broad array of industries—indeed, throughout the entire manufacturing sector and beyond—they do not result from industry-specific strategic settings, legal environments, technologies, or product markets, but rather from more pervasive factors. Second, the shipments data suggest that vertical integration need not necessarily imply, and in fact usually does not, physical goods being passed along production chains within firms.

We propose the following explanation that we believe concords with all of the presented facts: a firm’s boundary, and consequently its degree of vertical integration, is often drawn up to facilitate intrafirm transfers of intangible inputs. Managerial talent and oversight seem likely to be among the most important of such intangible inputs, although other inputs like marketing and

sales know-how may also be transferred readily across upstream and downstream units.³ That vertical integration is often about transfers of intangible inputs rather than physical ones may seem unusual at first glance. The assumption that firms use vertically integrated structures to produce upstream (physical) products for use as inputs in their downstream units is typically at least implicit, and often explicit, in the literature. 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. Within a given industry, however, we find few differences between vertically integrated plants that ship inside the firm as opposed to those that do not. Thus, it appears that the “make-or-buy” decision (at least referring to physical inputs) can explain only a fraction of the vertically integrated structures we observe in the economy.

How, then, can this theory also explain our first set of findings regarding the attributes of vertically integrated establishments? We take 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 intangible inputs such as managerial talent or marketing know-how are complementary to the physical inputs involved in making a vertically linked set of products, equilibrium assignment typically entails the allocation of higher-type intangible inputs to higher-type plants in each product category. To the extent that plant size is restricted by physical scale constraints, “better” intangible inputs may also be shared across multiple plants within a product category. Further, and more directly to our findings, equilibrium assignment allocates the best intangible inputs to multiple plants in distinct but often *vertically* related product categories. Simply put, higher-quality intangible inputs (e.g., the best managers) are spread across a greater number of production units. Some of these units can be vertically linked, but their vertical linkage need not necessarily imply the transfer of physical goods among them. In this way, vertical expansions can be seen in the same light as horizontal expansions. We shall return to these points below.

The paper 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

³ 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.

integration and shipments internal and external to integrated 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

Throughout this paper 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 by vertical integration status of manufacturing plants that are owned by firms that are in fact integrated, but only into non-manufacturing sectors (say, for example, a firm that owns a manufacturing plant and a retail store where it 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. (See the discussion in Section III below of how this classification is done.) In addition, the Census of Manufactures portion of the EC also contains considerable data on plants’ production activities. This includes information on 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 that will be a primary focus of the empirical work. Details on these measures 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.⁴

Commodity Flow Survey. The Commodity Flow Survey (CFS) collects data on shipments originating from mining, manufacturing, wholesale, and catalog and mail-order retail establishments. 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 in terms of ownership patterns, but we expect this will be small given the modest annual rates at which plants are bought and sold by firms.

III. How We Measure Vertical Integration and Shipments within Firms’ Vertical 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 integrated establishments are used within the firm or sent to external buyers.

The first step in determining which businesses are vertically integrated is to ascertain the

⁴ 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.

industry affiliation of every establishment in the Economic Census microdata. 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.⁵ 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.⁶

The next step is identifying in which industries firms operate, and whether any substantial vertical links are present within it. 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. We next determine if a firm owns establishments in industry pairs that we define to be two ends of a substantial link in a vertical production chain. We define a "substantial link" to exist 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.⁷

⁵ 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. Since the input-output classification system is primarily concerned with goods and services transfers within the production chain, 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."

⁶ 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).

⁷ 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 made several checks of our major findings using a ten-percent cutoff and found few differences (the overall level of integration is of course lower in the ten-percent case, because it is more stringent). We will also

Finally, we find all establishments that the firm owns on both ends of a substantial vertical link and classify them as being vertically integrated. 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.

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 establishment and firm identifiers that are comparable across the two datasets. 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

contrast vertically integrated producers with those in other multi-unit firm structures, which captures sensitivity of our classification to variations under a less stringent criterion—i.e., if the threshold were zero, any firm operating in multiple industries would have a “vertical” structure.

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.) For each upstream vertically integrated establishment, we then use the EC to find the ZIP codes of all the downstream plants owned by that 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.⁸

IV. Empirical Results

1. Aggregate Patterns: Vertical Integrated Firms Comprise a Large Share of Production

We begin by measuring the overall extent of integration in the economy over our 20-year sample period. The results are in Table 1, with panel A showing the level of integration for the entire non-farm private sector (computed from the entire Economic Census), and panel B focusing on the manufacturing sector, since much of the empirical work below will concentrate on manufacturing plants.

Several features of the data are salient in the tables. First, whether measured by the share of the total number of establishments or of total employment, the overall extent of vertical integration has been, roughly speaking, stable over the 20-year sample period. However, the prevalence of integration rose during the first half of the sample, and fell thereafter, giving back the earlier gains.

⁸ 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 (though likely only slightly, since in many industries, it is unlikely that there will be more than one establishment in a particular ZIP code) the fraction of internal shipments. One factor leading to understatement, on the other hand, is that we only observe just over 90 percent of establishments' ZIP codes in the EC data. 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 will be just over 10 percent higher than reported. This understatement is not relative to the true value, because of the other factors that lead to overstatement, but rather what would be obtained holding overstatement constant while observing all ZIP codes. We will further explore below many other possible measurement issues with classifying internal versus external shipments.

Second, establishments in vertically integrated structures are a relatively small share—between roughly 8 and 9.5 percent, depending on the year—of establishments in the economy. However, roughly 75 percent of the establishments in the economy are single-unit businesses; that is, they are the only establishment owned by their firm. Such plants cannot be vertically integrated by our definition, as any integrated structure under our measurement method requires a firm to have multiple establishments.⁹ Once we restrict our attention to only multi-unit establishments, whose establishment counts and employment levels are also shown in the table for comparison purposes, vertically integrated establishments account for roughly 35 to 40 percent of these multi-unit businesses.

Third, despite their modest share of the overall number of establishments, vertically integrated businesses account for a much larger employment share, 25-30 percent, and roughly half of multi-unit employment. Thus vertically integrated establishments are larger on average than are single-unit businesses or non-integrated multi-units. We will see this pattern manifested in several ways below.

Several of the patterns seen in the broader economy are present within the manufacturing sector. Integrated establishments are a small share (though somewhat larger than in the overall economy) of all establishments; they comprise a considerably larger employment share; and the level of integration was fairly stable over the sample, with some decline seen after 1987. New patterns emerge, such as integrated units accounting for an even greater share of multi-unit establishments and employment, as well as vertically integrated establishments' share of total revenues (revenue data are not available outside of manufacturing). The revenue shares echo the overall movements in the extent of integration. What they also make clear is that not only do vertically integrated manufacturing plants have higher average employment levels, they have higher average revenue levels—and indeed, higher average revenue-per-employee levels, since their revenue shares are notably larger than their employment shares, whether restricting the comparison to only multi-unit plants or all manufacturing plants. We will see below that the higher revenue productivity of labor finding is robust to any industry composition differences

⁹ It is of course possible that individual establishments might be internally vertically integrated to different extents. For example, one may bring in intermediate materials at a less refined stage than another making the same output from more refined inputs. Conceptually, one could consider the former to be more vertically integrated. However, our method can make no distinction in the levels of integration of these establishments. In future work, we expect to be able to address this shortcoming by focusing on industries where all establishments use raw materials in the same stage of processing, or by controlling for plants' input use directly using the Census of Manufactures Materials Supplement.

between integrated and unintegrated plants.

2. Plants in Vertically Integrated Firms Are Large

The relationship between vertical integration and establishment size discussed above is a good point at which to begin deepening our analysis.

Figure 1 shows the integration-size relationship within the 1.8 million plant-year observations of our Census of Manufactures data. It plots the share of plants that are vertically integrated as a function of the plants' within-industry size percentiles, where size is measured by logged real output (the construction of this and other plant-level production variables is detailed in the Data Appendix). For comparison purposes, it also shows the likelihood that plants are in two other multi-unit organizational structures: single-industry and multi-industry-unintegrated multi-units. The former are plants in multi-unit firms that own plants in only one industry. These facilitate a comparison of vertically integrated plants, which are necessarily multi-unit plants, to other multi-unit plants, but only those whose firms restrict their scope of operations to one industry. The latter plants are owned by firms that have plants in multiple industries, but none of which comprise substantial vertical links as defined above. These afford comparisons of VI plants to those in other multi-industry firms, but those without vertical structures.¹⁰

Intra-industry differences in integration likelihoods are clearly related to size.¹¹ The size “effect” also has a substantial magnitude. While industries' smallest plants are almost never integrated, the median-sized plant is integrated 7 percent of the time, and a whopping 67 percent of plants in the top percentile of their industry size distribution are integrated.¹² Furthermore, while the likelihood of plants being in one of the other two multi-unit plant types also grows in size (the residual category is single-unit plants), these types do not exhibit the strong nonlinearity seen in vertically integrated plants. Thus even within plants in multi-unit firms, size is most closely linked to vertical integration. This is seen in Figure 2, which restricts attention only to

¹⁰ The prevalence of these plant types 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.

¹¹ No standard error bounds are plotted because they would be virtually imperceptible given our sample size.

¹² The relationship's lack of noise is notable. We stress that the figure is essentially raw data—frequency counts by category. No functional forms have been imposed or parametric approximations made to produce it. The power-law-like relationship is purely a factor of the underlying process driving vertical integration in the data.

multi-unit plants (i.e., the denominator is the number of multi-unit plants rather than all plants, so the fractions necessarily add to one). The fraction of multi-unit plants in each of the organizational forms is roughly constant up to about the 60th percentile (though of course rising as a share of all plants, as seen in Figure 1), with vertically integrated plants accounting for just over half of multi-unit plants. Above that relative size, vertically integrated plants become more prevalent, while multi-industry-unintegrated and especially single-industry organizational forms become less common.¹³

These patterns indicate the largest plants tend to be in vertical structures. This naturally leads to the question of whether firms with vertical structures are themselves larger. In other words, does the extensive margin (how many plants a firm owns) augment or counteract the intensive margin? Figure 3 plots the densities of firm size (measured as logged total employment) for the three types of multi-unit firms discussed above. Before discussing the distributions, a few distinctions between this plot and the plant-level results we just presented should be noted. Here, size is measured by total firm employment rather than real revenue as above. This is necessary because we need to know the size of firms' establishments outside the manufacturing sector, where revenue data is unavailable. Further, 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. Additionally, we must drop the plant-specific definition of vertical integration in order to classify every firm into a unique category. Thus firms with any vertically integrated plants will be labeled integrated, even if they own plants outside of vertical production chains. Though as a practical matter, most plants in what we call vertically integrated firms here are vertically integrated by our plant-specific definition. Finally, for data confidentiality reasons, the plots have had their extreme tails trimmed, dropping 10 to 15 or so of the smallest and largest firms in each organizational form

¹³ We have checked that these results are not artifacts of some peculiar aggregation phenomenon by constructing analogous plots for three individual industries: pumps and compressors, inorganic and organic chemicals, and miscellaneous plastics. We chose these industries only because they happen to sit at (respectively) the 75th, 90th, and 99th percentiles of numbers of industry establishments. (We need industries with many establishments for such plots because they are only meaningful if there are hundreds of plant-year observations for the particular industry.) Similar patterns were observed in each industry. The *level* of vertical integration varied across industries, presumably reflecting the influence of industry-specific technological or strategic considerations that make integration more advantageous. However, *changes* in integration likelihoods across different-sized plants were similar across industries. Furthermore, we also checked our results for robustness to "substantial link" thresholds other than five percent and found little qualitative difference, besides the obvious difference that levels of vertical integration were lower under more restrictive definitions.

category.

Figure 3 reveals that each of the (logged) employment size distributions is unimodal, though they clearly have different central tendencies. Single-industry multi-unit firms are the smallest and have the most symmetric size distribution. Vertically integrated firms are clearly the largest on average, and their distribution is more skewed than the other firm types. (While not plotted, single-unit firms/establishments 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.

The connection between establishment scale and integration status has, to our knowledge, not yet been theoretically addressed or documented empirically. Yet as we will show, it is part of a robust pattern of within-industry heterogeneity in production measures being tied to the likelihood that a plant is in an integrated structure within its firm.

3. Vertically Integrated Establishments are High-“Type” Plants

We broaden the integration-scale connection explored above by generalizing it to other plant-level measures of “type.” We conceptualize plants’ types as being combinations of idiosyncratic demand and supply fundamentals that affect plant profitability.¹⁴ In industry equilibrium, these are tied to plant observables like size, productivity, and (in some cases) factor intensities.

We use four such measures in our empirical work. They are not independent, but do differ enough in construction to allow us to gauge the consistency (or lack thereof) of our results. One, seen above already, is the plant’s logged output. This is its reported value of sales deflated by the appropriate price index for the plant’s four-digit SIC industry.¹⁵ Two are productivity measures that differ in their measure of inputs: output per worker-hour and total factor productivity. Both are expressed as the log of the plant’s output-input(s) ratio. The other is a plant’s logged capital-labor ratio (capital stock per worker-hour). Further details on the construction of these measures are in the Data Appendix.

These measures have been shown in various empirical studies to be positively related to

¹⁴ Foster, Haltiwanger, and Syverson (2005) 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.

¹⁵ Industry-specific price deflators are from the NBER productivity database.

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 four-digit SIC 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 integrated and unintegrated plants. By including fixed effects we are identifying type differences across integrated and unintegrated plants in the same industry-year, avoiding confounding differences in productivity, scale, or factor intensity levels across industries and time. As a point of comparison, we use a separate specification that also includes an indicator for any plant that is part of a multi-industry firm. VI plants are a subset of these plants; the remainder are the multi-industry/unintegrated plants discussed above. This second specification allows us to see if differences between integrated and others in their industry are separable from differences between any plants in multi-industry firms and the remainder of the industry. We estimate these specifications separately for each of the four plant type proxies and report the results in Table 2, panel A.¹⁶

It is clear that vertically integrated plants have higher types. Besides being larger, integrated producers have significantly higher productivity levels and are more capital intensive. Their labor productivity levels are on average 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. The result in the output specification essentially maps the pattern in Figure 1 into a single coefficient. Capital intensities are

¹⁶ Sample sizes differ across the specifications because not all the necessary variables for construction of each is 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.

substantially higher in integrated plants as well, explaining why the integrated plants' labor productivity advantages are so much larger than the TFP differences. Furthermore, all of these gaps in types are preserved, though all become smaller, when we control for being in a multi-industry firm, as seen in the results from specification 2.

These patterns can be most succinctly summarized in plots similar to those discussed above for plant size (output). These show the likelihood that plants are of a given type as a function of their location in their within-industry type distribution. Figure 4 shows likelihoods as a fraction of all plants (corresponding to Figure 1 for size), and Figure 5 shows that for multi-unit plants only (corresponding to Figure 2). The basic patterns are the same as seen with plant size; plants at the high end of their within-industry distributions of type, be it productivity or capital intensity, are more likely to be vertically integrated. The only departures from the size results is that the magnitude of the relationships are weaker—never are more than 40 percent of the highest-percentile productivity or capital intensity plants vertically integrated—and for TFP, the propensity to be vertically integrated is no greater than to be part of any other multi-unit organizational form as one moves up the within-industry distribution.

A natural question that follows from these results is the causal nature of vertically integrated plants' type differences. There are three possibilities, and they are not mutually exclusive. The gaps could be due to the fact that newly built integrated plants are different than newly built plants in unintegrated 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 vertically integrated structure might be associated with a change in an existing plant's type.

We are able to separately investigate these possibilities in the data. To see if new integrated plants are different than newly built plants in unintegrated ownership structures, we regress all new plants' type measures on a dummy for their integration 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.¹⁷ New plants are an important part of the formation of vertically integrated structures in the economy: entering

¹⁷ 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 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 2 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 between new integrated and unintegrated plants are still quite large, though somewhat less pronounced than the gap seen in the overall sample. Thus many of the differences observed between integrated and unintegrated producers' type measures reflect persistent differences present even at the time of the plants' entry.¹⁸

The second possible source of integrated plants' higher type measures is that firms comprised of high-type, vertically integrated 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-integrated plants compare *before integration* to other plants in their industry that do not become integrated over the same five-year period.

The results, which are in panel C of Table 2, 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, but still of a similar order of magnitude as are the gaps measured above.

Finally, we investigate if becoming integrated is associated with unusually high growth in

¹⁸ As with the broader comparisons above, we have also checked whether these differences among new plants remain after controlling for being part of a multi-industry firm. The results were similar to those in panel A: new vertically integrated plants have significantly higher type measures even compared to other new plants in multi-industry firms, though the residual differences were of smaller magnitude.

productivity, scale, or factor intensities by comparing 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 unintegrated in the prior EC.

The results are shown in panel D of Table 2. Becoming vertically integrated 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. Whether capital deepening occurs because the plant experiences growth in capital, declines in labor, or both, is an open question at this point -- but one which we will explore further below. 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 integrated and unintegrated plants 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.¹⁹

¹⁹ 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.

4. Assortative Matching of Types within Firms

The fact pointed to in Table 2, panel C—that plants that are to become vertically integrated already have higher types than their industry competitors—indicates that firms with vertically integrated structures, when choosing to expand through merger or acquisition, do so by incorporating existing plants that are also high-type. This is true even if the matched plants are not vertically integrated before being incorporated into the new vertical structure. This is precisely the pattern implied by assortative matching equilibria where firms are constructed by matching plants of differing types. As in Koopmans' (1951) assignment model and Becker's (1973) matching model, complementarities may predict this sort of assortative matching in equilibrium. If firms' vertical structures are driven by complementarities, this result has implications for the correlation of plant types within firms. Because the gain to matching with a high-type producer is higher for other high-type operations, firms with high-type plants have a greater willingness to pay to match with other high-type plants and will outbid low-type plants to do so. The equilibrium outcome has the highest-type producers matching with highest-type suppliers, slightly lower-type producers match with other similar-type suppliers, and so on down the type distribution. The upshot of the matching process is that plants in vertical production chains should have positively correlated types.

We investigate this prediction using our four plant type proxies. Unlike the comparisons above that are completely among plants in the same industries, we must now compare type measures across a firm's several plants, where each plant operates in a different industry. Inter-industry technology differences make direct comparison of type measures very difficult; for example, a certain output level may be large for a plant in one particular industry but considered quite small in another. Therefore we use plant type measures that have been normalized by their four-digit SIC industry medians in the year of observation. Since all our measures are in logs, these normalized values correspond to proportional deviations relative to the median, meaning we are correlating the relative positions of a firm's plants *within* their respective industries.

This is operationalized by first computing for each plant the weighted average (the weights are deflated revenues) of these proportional deviations across all *other* plants in its firm. (Obviously, only multi-plant firms can be included in the sample.) We next regress the plant's own within-industry productivity, output, or capital intensity deviation on their respective weighted averages across the other plants in the firm. A positive relationship means that plants

with relatively high type measures relative to their industry median are in firms with other plants that are themselves above their industry's median type, and that low-type plants are in firms with other low-type plants. To see if this correlation is systematically different for firms with vertically structures, we interact the weighted-average-type-deviations with an indicator for such firms, and also include this indicator separately in the regression.

Panel A of Table 2 shows the results of this exercise. Columns correspond to different plant type measures. (The vertical firm indicator main effect is not shown for the sake of parsimony.) The within-firm correlation of each plant type measure is significantly stronger in vertically integrated firms than in other multi-unit firms. This is true both in statistical terms and in terms of economic magnitudes. For example, the coefficients relating plants' labor productivities to those of other plants in their firm are roughly 55 percent higher (e.g., 0.359 for exclusively horizontally integrated multi-unit firms and 0.554 for firms with vertically integrated plants). Even for TFP, which exhibits the smallest difference in the correlations, vertically integrated firms have a coefficient over 25 percent higher than that for exclusively horizontal firms. Plants with high productivity (or output or capital intensity) levels tend to be in firms with other high-productivity (or high-output or high-capital-intensity) plants, while low-type plants join together in their own firms.

Vertically integrated plants are therefore owned by firms with other plants that are similarly located in their industries' type distributions. These patterns are consistent with assortative matching being tied to vertical integration. It is also possible that the correlations partially reflect common firm-level fixed factors that impact the type measures of all plants in the firm. However, common fixed factors are unlikely to be the only explanation for the observed patterns, since they do not explain the fact we point to at the beginning of this section: soon-to-be-integrated plants are already of higher-than-average type for their industry.

5. Vertically Integrated Establishments' Shipments

Our look at the shipments of vertically integrated establishments begins by finding those establishments in the Commodity Flow Surveys that we know are in vertically integrated structures and that are an upstream link within their firm's vertical production chain(s). We then compare the destination ZIP codes of these plants' shipments to those of the firm's downstream plants. If the destination ZIP matches the location of any of the downstream plants, we consider

the shipment to be internal to the firm. The details of this process are discussed in Section III above.

The combined 1993 and 1997 CFS yield a total of 29,931 upstream VI plant-year observations that, among them, report a total of 2,826,296 shipments. Panel A of Table 4 shows the prevalence of internal shipments within this sample. It reports quantiles of the distribution of the shares of plant's shipments that are internal, measured as the fraction of the total number, dollar value, and weight of the establishment's shipments.²⁰

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 90th percentile plant sells over 40 percent of its output outside the firm. The exceptions to this general pattern are a small set of establishments dedicated to serving only the needs of their firm: 2.1 percent of the sample report only internal shipments. This specialization is even more apparent from the histogram of plants' internal shipment shares in panel A of Figure 6. The overall pattern is one of very little internal transfers, as seen in the quantiles. The number of establishments is declining monotonically in the fraction of internal shipments. The clear exception, however, are the cluster of internally dedicated establishments that create to the bimodality. That said, the number of internal specialists is still only a small fraction of the number reporting no internal sales at all.

These results imply that the traditional view that firms own plants in upstream industries to control supply of an input may not be the primary driver of vertical ownership patterns in the economy. Certainly other motivations must apply for those plants that do not make any internal shipments. Even for those that do serve their own firm, though, the small share of internal shipments suggests that this role may not be their primary one.²¹

²⁰ For data confidentiality reasons, the reported quantiles are actually averages of the immediately surrounding percentiles; e.g., the median is the average of the 49th and 51st percentiles, the 75th percentile is the average of the 74th and 76th percentiles, and so on.

²¹ Of course, it is possible in some production chains for an upstream plant to serve its firm's downstream needs completely with only a small fraction of its output. This is more likely in production chains where upstream industry establishments tend to be large and downstream ones small. In the aggregate, however, these types of production chains will tend to be balanced out by those with small upstream and large downstream plants, which would tend to raise internal shares for such production chains. The fact that such a large fraction of upstream integrated establishments delivers so small a share of their output to their firm makes it unlikely that the result is

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

First, it seems 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 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.” Furthermore, the Census Bureau audits responses by comparing the establishment’s implied annual value of shipments from the CFS with that from other sources, such as the Economic Census (which is completed at a different time and, often, by another individual). If the disparity is beyond statistical variance, the respondent is contacted by the Bureau and the responses are reviewed for accuracy. If integrated establishments were systematically underreporting internal shipments because of confusion or by not following directions, the auditing process would presumably catch this. Moreover, most establishments do report *some* internal shipments, and a few nothing but internal shipments, indicating that they have not interpreted the definition of shipments as precluding intra-firm transfers.

We also conduct a series of robustness checks. The corresponding distributions of establishments’ internal shipments are shown by row for each separate check in panel B of Table 4. 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 that report 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 and just under 1.9 million shipments (note that this is greater than half the number of establishment-years in the entire sample because several observations report exactly the median number of shipments). The results for this sample are shown in the first row of panel B. Extreme values are in fact rarer in this sample: 29.4

driven by the peculiarities of a small number of production chains. We will conduct several robustness checks below to test the resiliency of the result.

percent report making no internal shipments, down from 34 percent in the full sample, and 0.7 percent report nothing but 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 90th 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 shipments for export. In the CFS, the destination ZIP code of shipments for export corresponds to the port of exit, with a separate note made regarding 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 same ZIP code as the port. 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.

The next check drops all wholesale establishments from the sample. Many industries have the wholesale sector as both an upstream and downstream link. Therefore, it is possible that we are misclassifying a wholesaler as an upstream link in a vertical structure in a firm, when it may in fact be downstream, more likely to be shipping to final demanders like consumers, which are obviously external to the firm. Moreover, because of the nature of their business, wholesalers comprise a significant portion of the CFS sample, so the impact of any such classification problem would be amplified. 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. The median fraction of internal shipments is 5.4 percent, and the 90th 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 as internal shipments destined for the ZIP code of *any* plant in the firm, not just those in downstream links of vertical chains. It is possible that some vertical production may actually occur outside those chains we identify using the Input-Output tables. Some may in fact 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 what shipments might represent the intra-firm transfer of physical goods along a production chain. As seen in row 4 of panel B, all quantiles involve internal shipment fractions higher than the benchmark numbers, as they must. Even so, the fractions of internal shipments still strike us as modest. The median is 6.5 percent, and the 90th percentile 67.2 percent. About 20 percent of establishments 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 6. While internal shares are of course higher, it has the same shape as the benchmark distribution in panel A.

Finally, we make the very 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 small 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. The median shipment in the dataset travels only about 75 miles, so a large fraction of shipments never even leave their county of origin. The generous criterion is much more likely to yield internal specialists or near-specialists: the 90th 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 share is 28 percent.

V. Discussion

We have documented a number of facts relating observables of plants' production behavior and whether or not they are owned within a vertically integrated structure. Being integrated is correlated with higher levels of several plant type measures (relative to others in the plant's industry): size, productivity, and capital intensity. These differences primarily reflect persistent dissimilarities existing either at plant birth (if it is born into a vertically integrated structure) or before it is brought into the integrated firm, but also embody changes that happen when an unintegrated plant becomes integrated. Further, there is evidence that matching on plant types is particularly strong in firms with vertical structures.

When we look at the patterns of shipments from upstream vertically integrated plants, we find that a large fraction of them do not ship their outputs internally, or if they do, the portion of their output which is accounted for by intra-firm transfers is small.

Is there an explanation that is consistent with all of these patterns, and that is likely to apply across industries with very diverse technologies, product markets, and strategic environments? We propose one here: namely, that vertical integration is used to facilitate efficient transfers of inputs, but not necessarily (or even usually) physical inputs along a production chain, as is commonly assumed. We expect that managerial oversight and planning is among the most important of these, although marketing and sales know-how might also be readily transferred among integrated establishments in a firm. Note that in such cases, the distinction between “downstream” and “upstream” becomes one of convenience rather than an accurate depiction of the flow 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.

In this way, vertical expansion by a firm may not be altogether different than horizontal expansion. 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. Because of the special structure of production chains, one might expect vertically integrated establishments to make physical transfers in a way that wouldn’t be expected from horizontally related establishments. However, the evidence above indicates that many vertical expansions could instead operate similarly to horizontal ones. 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-tangible inputs will be made to the new establishments. Yet no physical good transfers from upstream to downstream establishments need occur.

The results regarding the scale, productivity, and capital intensity differences seen in integrated plants is consistent with a world where intangible inputs like management or marketing are complements in production with physical inputs, and they are assigned in a market process to different production units. Complementarities imply assortative matching in equilibrium; 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 that vertically integrated production chains are found in the largest firms that are composed of the highest-type plants.²²

We take several approaches to testing this possibility more directly in the data. We first look at the internal shipment patterns of a very select group of upstream vertically integrated establishments. 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) that have 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 integrated—they were brought into the firm to be part of the firm’s first internal vertical structure. 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 678 establishment-years, reporting 122,560 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 are in the final row of Table 4, panel B. 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 (several hundred of the 678 observations are in the same industry-year, offering enough

²² This explanation shares the spirit of other models in the equilibrium assignment view of firm organization, like Lucas (1978), Rosen (1982), and more recently by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007)

variation to separately identify the subsample dummy). The estimated coefficient on the subsample dummy in the dollar-value-share regressions is -0.035 (s.e. = 0.009), and 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, suggesting that other motives might be driving the integration decision.

We next see if the amount that an establishment ships internally is correlated with its type within its industry. We know from above of course that there are type differences between vertically integrated and unintegrated plants; here we ask if within integrated plants, there is a correlation between type and the plant's propensity to make products for downstream units in its firm. We simply regress plants' internal shipments shares (again we focus on dollar-value shares) separately on each of our four type proxies and a full set of industry-year fixed effects. We also estimate specifications that include controls for the typical attributes of the plant's shipments: the logs of the median mileage, the median weight, and the average dollar value of its shipments.²³ The sample includes only those upstream vertically integrated establishments for which we observe type measures; this essentially limits us to manufacturing plants. And because we cannot measure capital in 1997, the sample is considerably smaller for the TFP and capital-to-labor ratio correlations. The results are in Table 5.

The correlations between type and intra-firm shipment shares are sometimes statistically significant. However, they have inconsistent signs, are sensitive to the inclusion of controls, and have small economic magnitudes. For example, moving to a plant one standard deviation higher in the industry TFP distribution implies a drop in the internal share of 0.8 percentage points in the specification without controls; with controls, it implies a 0.5 percentage point drop. To move from the 25th percentile internal shipments plant to the median, a labor productivity shift of 2 standard deviations is implied; to move from the median to the 75th percentile, a 13-standard-deviation change is implied (and these numbers use the specification without controls, where the

²³ We have made extensive comparisons of the attributes of internal versus external shipments that are not presented here for space reasons. To summarize, the most robust result is that internal shipments tend to travel shorter distances. This is true even focusing exclusively on within-plant variation. Internal shipments also tend to be larger, whether measured by dollar value or weight, but the weight gap is larger and as such they have a lower dollar value per pound. These value and weight findings are much stronger across plants than they are within plants, however.

correlation is much larger). Furthermore, we have also investigated, though do not report for space reasons, how the prevalence of internal shipments is related to plant types among vertically integrated entrants and changes in types for those that become integrated. We found no significant correlations between internal shipment shares and any of the type measures.

As with the results for the newly integrated establishment/firm subsample above, these results suggest that internal transfers of physical goods might not be the primary driver of the formation of vertically linked firms. Intra-firm shipment prevalence does not seem to be related to firms' vertical expansions, even when we focus on establishments that are explicitly tied to a firm adopting a vertical structure, and upstream integrated plants' internal shipment shares are not closely related to their types within their industries.

The final test we conduct digs deeper into the changes seen in plants that become vertically integrated, as with those observed in panel D of Table 2. We pay particular attention to the increases observed in capital intensity, because these are consistent with a world where skilled managerial or other intangible inputs have stronger complementarities with capital than labor. In such a case, the observed shifts in capital intensity would be expected in the allocation mechanism we discuss above as a source of our findings.²⁴

These changes in the capital-labor ratio could come from the acquiring firm making investments in the newly integrated plant, but it could also be driven by decreases in labor inputs. To distinguish between the two possibilities, we repeat the exercises in Table 2, panel D, but this time include separate specifications for plants' capital stocks and labor stocks. So we can exactly decompose the change in capital intensity into its sources, we restrict the sample to sets of plants for which we observe each of the production measures. This also lets us decompose the sources of labor productivity growth, since we observe output and labor changes separately for the same plants. Furthermore, we look at changes in two types of labor inputs separately: production and nonproduction workers. The results are shown in Table 6.

The increase in capital intensity is sourced in both a dose of investment—capital stocks in newly integrated establishments grow 6.7 percent faster than they do in plants in the same industries that remained unintegrated—and through a decrease in labor inputs. (All variables in

²⁴ Firms with vertically integrated 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.

the table except the nonproduction worker share enter the empirical specifications in logs.) Total hours decline 2.4 percent. This decline in hours also pairs with equal-sized output growth to raise labor productivity in newly integrated plant 4.8 percent faster. What is most interesting about this change in labor inputs from the perspective of our explanation for the facts documented above is in the changes in labor composition that accompany this decline. Nonproduction worker counts fall more than do production workers; in fact, the change in the latter is not statistically significant. If we look at the share of nonproduction workers in the plants' total employment, we see that this falls significantly when the plant becomes vertically integrated. While we cannot observe workers' positions within the firm at any finer level than the production/nonproduction worker dichotomy, 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.

In this section, we have proposed an explanation that is consistent with our findings from Section IV and pushed a little deeper into the data to see if other patterns are consistent with it. There does appear to be corroborating evidence of the notion that integration is often about facilitating intra-firm transfers of intangible inputs rather than tangible ones, and that the fact that these inputs are complements with other inputs into production leads to a market assignment that puts more (sometimes vertical) production capability in the hands of those that are the most skilled at providing these inputs.

VII. Future Work and Conclusion

Our analysis thus far indicates that firm organization along the vertical production chain has notable correlations with the size and performance of production plants. Since vertical integration/expansion decisions are endogenous, it is difficult to draw causal inferences, except perhaps from the panel data evidence from newly integrated plants. To supplement this analysis, we are currently investigating various exogenous drivers of vertical merger decisions, following our strategy of utilizing changes in antitrust attitudes in our previous analysis of the ready-mix concrete and cement industries (Hortaçsu and Syverson (2007)).

Another potential concern one may have with the analysis until now is that our definition of vertical integration, which is based on the presence of substantial materials flows between the *industries* that plants belong to, may be too generous. Some plants within a firm may be deemed “vertical partners” by our industry-based definition, even though a more careful look at these plants may indicate that, within their industries, they are engaged in making products that are vertically unrelated. We are attempting to address this measurement issue by looking at narrow vertical product pairs (such as cement and ready-mixed concrete) that we know to be vertically related, and checking whether the internal shipments patterns of upstream plants in such industries to their downstream partners within the firm resemble the empirical patterns we reported in Section IV.5. Preliminary results on this analysis will be made available subject to disclosure review.

We are also working on constructing a cohesive theoretical framework to elucidate the ideas discussed above. It will spell out precisely the contracting problems associated with the transfer of intangible “knowledge” inputs and how the boundaries of firms along the vertical dimension are obtained as the equilibrium outcome of an assignment model. In many ways, the modeling framework of Garicano and Rossi-Hansberg (2006) paves the way for this effort, but the implications of their model in the vertical integration and expansion framework have yet to be studied.

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 α_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.²⁵ 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.

²⁵ 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.
- 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?" NBER Working Paper 11555, 2005.

- 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.
- 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.
- 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*, forthcoming.
- 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 Earning," *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: Macmillian, 1998.
- 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. Aggregate Patterns of Vertical Integration, 1977-1997

A. Non-farm Private Economy

	Year:	1977	1982	1987	1992	1997
VI establishments (thousands)		384.3	421.7	546.7	519.8	549.3
Multi-unit establishments (thousands)		1033.7	1167.0	1336.8	1476.6	1605.6
Total establishments (thousands)		4862.2	5049.8	5855.5	6253.2	6831.1
VI establishment share (percent)		7.9	8.4	9.4	8.3	8.0
VI employment (millions)		20.4	21.5	26.9	26.5	28.3
Multi-unit employment (millions)		38.2	42.7	48.3	53.9	60.7
Total employment (millions)		68.1	75.7	87.7	93.6	106.1
VI employment share (percent)		29.8	28.4	30.7	28.3	26.7

Notes: Computed from the Longitudinal Business Database; vertically integrated establishments are defined as described in the text.

B. Manufacturing

	Year:	1977	1982	1987	1992	1997
Total establishments (thousands)		350.6	348.4	358.9	371	375.9
VI establishment share (percent)		12.5	12.5	12.6	11.4	11.1
Multi-unit establishment share (percent)		20.5	20.7	19.8	18.9	18.8
Total employment (millions)		18.5	17.8	17.7	17.0	17.5
VI employment share (percent)		57.0	52.9	57.0	53.2	50.3
Multi-unit employment share (percent)		75.0	73.1	71.2	69.7	67.8
Total revenue (\$ billions, nominal)		1358.3	1960.2	2475.9	3005.0	3961.8
VI revenue share (percent)		70.9	68.7	72.5	70.8	69.4
Multi-unit revenue share (percent)		84.8	84.7	83.2	83.0	82.6

Notes: Computed from the Longitudinal Business Database and Census of Manufactures; vertically integrated establishments are defined as described in the text.

Table 2. Plant Attributes and Vertically Integrated Plants

A. Within-Industry Differences

	Output per hour	TFP	Output	Capital-labor ratio
N	1,048,887	739,366	1,073,978	787,283
<u>Specification 1</u>				
VI indicator	0.348* (0.002)	0.017* (0.001)	1.509* (0.004)	0.443* (0.003)
<u>Specification 2</u>				
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)

Notes: This table shows plant “type” comparisons between vertically integrated plants and their unintegrated counterparts. Specification 1 regresses plant type on an indicator for vertically integrated plants. Specification 2 also includes an indicator for any plants owned by multi-industry firms (VI plants are a subset of these). 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 2. Plant Attributes and Vertically Integrated Plants, Cont.

	Output per hour	TFP	Output	Capital-labor ratio
B. Differences among New Plants				
N	236,228	152,618	246,464	166,272
VI indicator	0.285* (0.004)	0.032* (0.003)	1.251* (0.010)	0.368* (0.007)
C. Comparing Unintegrated Plants: To-Be-Integrated vs. Remaining Unintegrated				
N	648,449	578,634	667,727	621,131
To-be-VI indicator	0.215* (0.004)	0.027* (0.003)	1.416* (0.009)	0.245* (0.006)
D. Changes upon Integration				
N	481,777	303,775	496,688	331,854
Newly VI 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 vertically integrated (or to-be-integrated) plants and their non-integrated counterparts. Panel B compares new integrated and unintegrated plants, and panel C compares *prior period* types among unintegrated plants that will become integrated by next period to those remaining unintegrated. Panel D compares changes in type for unintegrated plants that become integrated 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 3. Within-Firm Correlations of Plant “Type” Measures

	Output per hour	TFP	Output	Capital-labor ratio
N	295,625	219,517	302,418	229,523
Firm’s weighted mean	0.359* (0.007)	0.344* (0.014)	0.479* (0.007)	0.408* (0.009)
Firm’s weighted mean X VI firm	0.195* (0.010)	0.097* (0.018)	0.257* (0.009)	0.188* (0.012)

Notes: This table reports coefficients from a regression of plant “type” measures (productivity, output, or capital intensity) on the output-weighted average of the same type measure among other plants owned by the same firm, both by itself and interacted with an indicator denoting firms with vertical structures. (The indicator is also included as a main effect but is not reported here.) All type measures are log deviations from industry-year medians. The sample includes all non-AR manufacturing plants in multi-plant firms. See text and data appendix on construction of type measures and additional details. Standard errors are clustered by firm. An asterisk denotes significance at a five percent level.

Table 4. Plant-Level Shares of Internal Shipments

A. Benchmark

Internal share of:	Percentile					Fraction = 0	Fraction = 1
	10 th	25 th	50 th	75 th	90 th		
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*(49th percentile + 51st percentile).

B. Robustness Checks (Share of Dollar Value Shown)

Specification/Sample	Percentile					Frac. = 0	Frac. = 1
	10 th	25 th	50 th	75 th	90 th		
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. Newly VI plants in newly VI firms	0	0	0.4	11.9	47.6	45.7	1.8

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*(49th percentile + 51st percentile).

Table 5. Correlations between Plant-Level Shares of Internal Shipments and Type Measures

Type Measure	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Output per hour	0.022* (0.004)	0.002 (0.004)						
TFP			-0.021* (0.009)	-0.013 (0.009)				
Output					0.000 (0.002)	-0.007* (0.002)		
Capital-labor ratio							0.021* (0.003)	0.012* (0.003)
Controls?	No	Yes	No	Yes	No	Yes	No	Yes
N	8926	8926	14976	14976	15039	15039	9053	9053
R ²	0.161	0.253	0.166	0.263	0.163	0.264	0.164	0.254

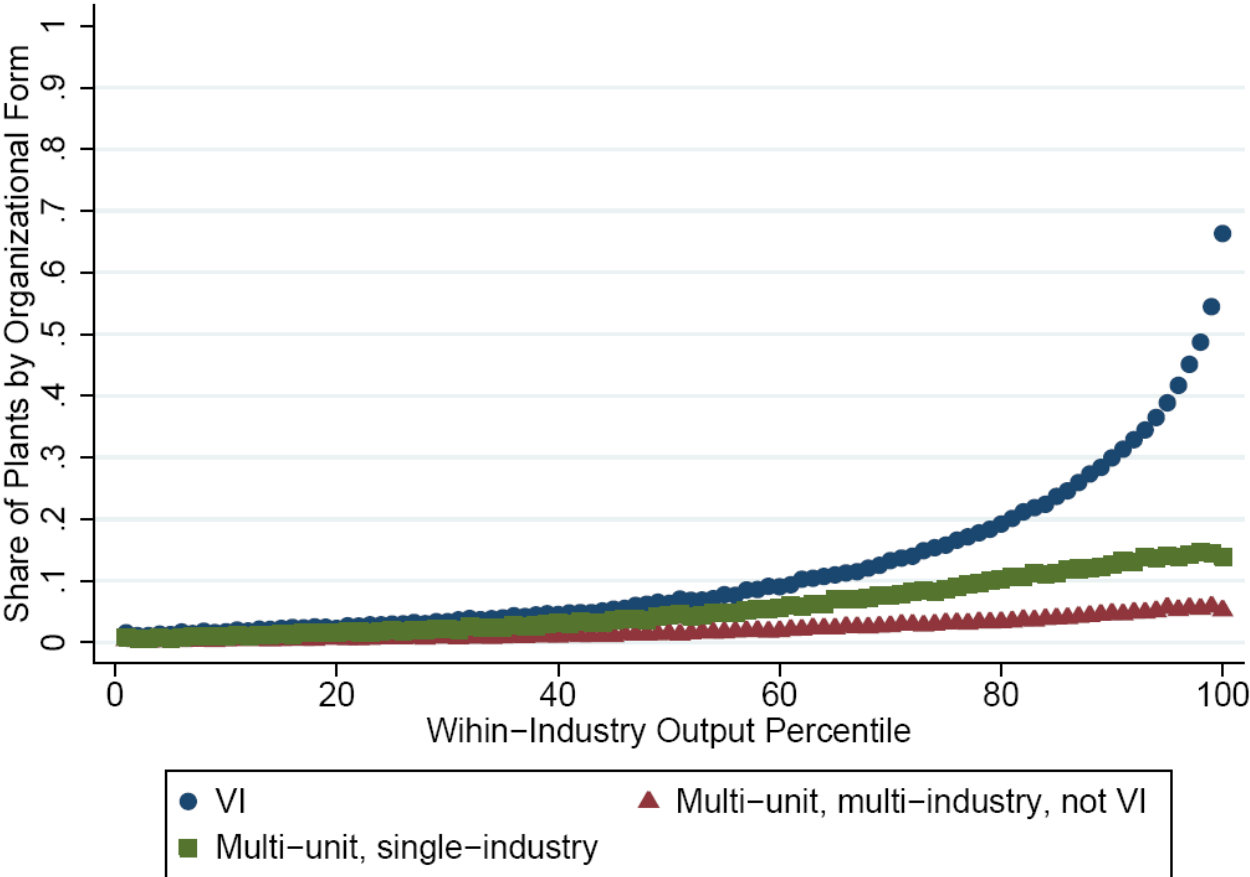
Note: This table shows the results from regressing a plant's share (by dollar value) of internal shipments on its type measures. All regressions include industry-year fixed effects. The sample includes our upstream vertically integrated plants for which type measures are available—i.e., those in the manufacturing sector. Controls include the logs of the median mileage, the median weight, and the average dollar value of plant shipments. An asterisk denotes significance at a five percent level.

Table 6. 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)

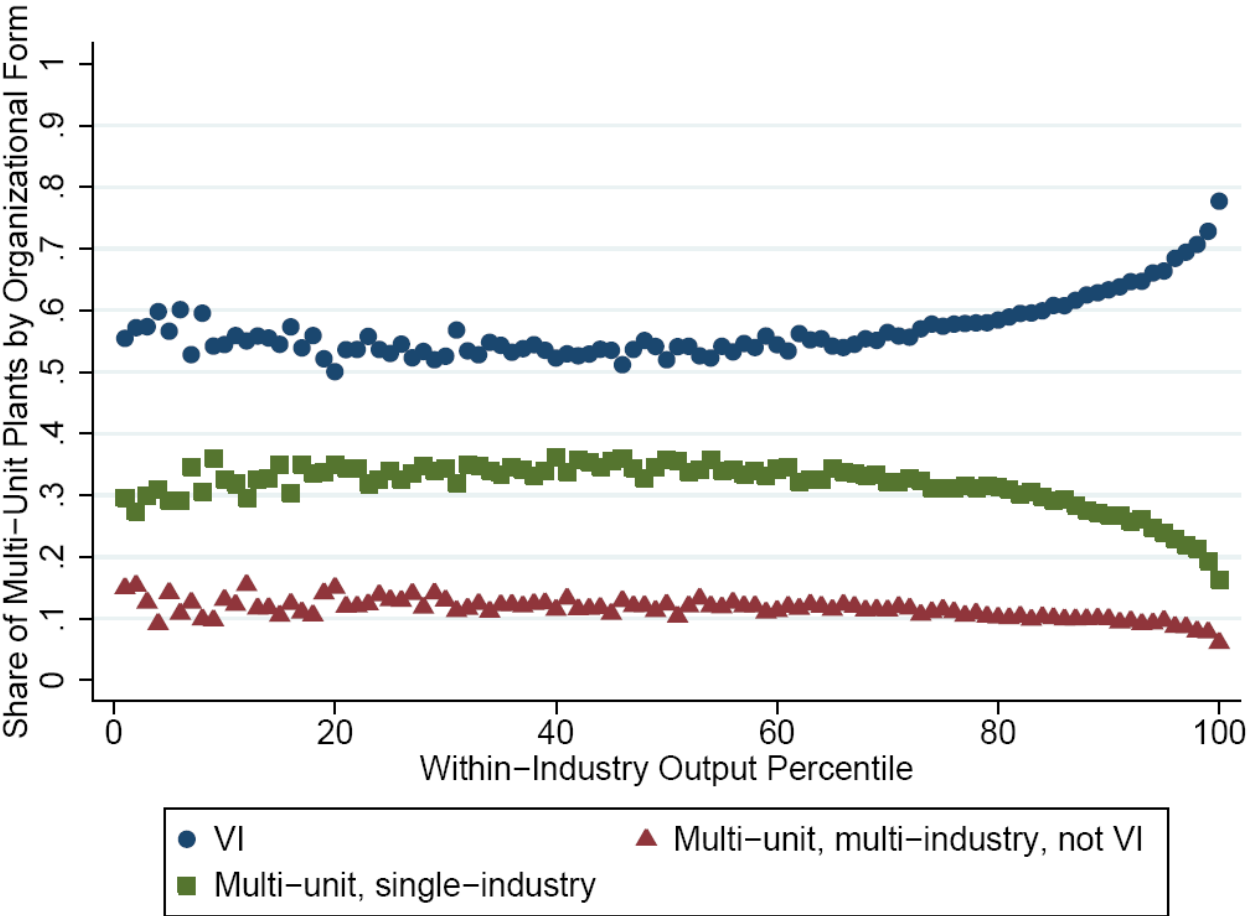
Note: The table shows specifications repeating the exercises in panel D of Table 2, but with additional plant production measures included. Further, the sample consists of only those 282,240 newly integrated plants that have nonmissing data for all production measures. See text for details of the construction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

Figure 1. Scale, Vertical Integration, and Other Organizational Forms



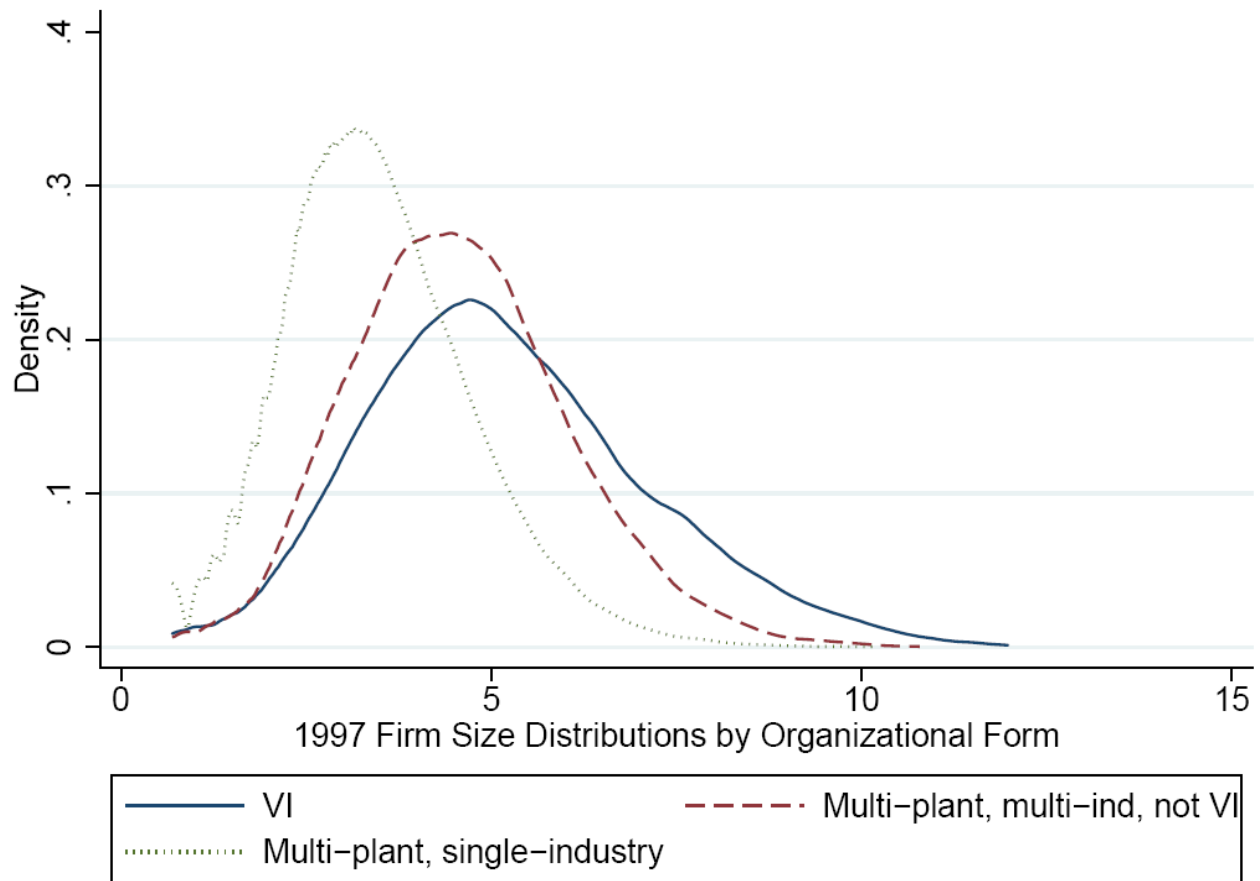
Notes: This figure plots the fraction of manufacturing plants in various multi-unit firm organizational structures by plant scale (measured by real revenue) percentile within industry-year. Approx. sample size is 1.8 million plant-year observations. See text for details.

Figure 2. Scale and Multi-Unit Plants' Organizational Forms



Notes: This figure plots the fraction of *multi-unit* manufacturing plants in various multi-unit firm organizational structures by plant scale (measured by real revenue) percentile within industry-year. Approx. sample size is 356,000 plant-year observations. See text for details.

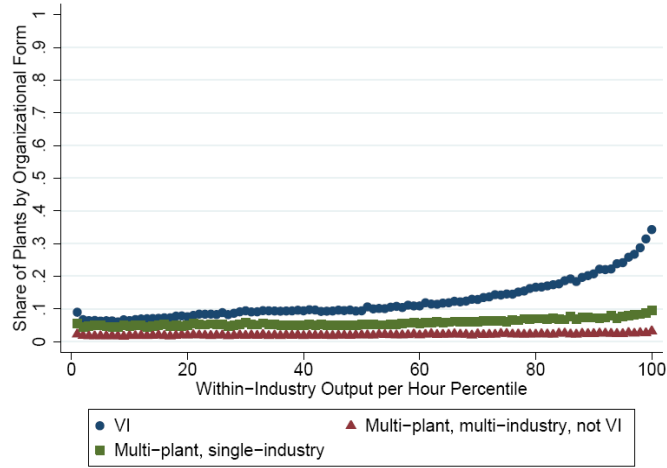
Figure 3. Firm Size Distributions for Multi-Unit Firms, 1997



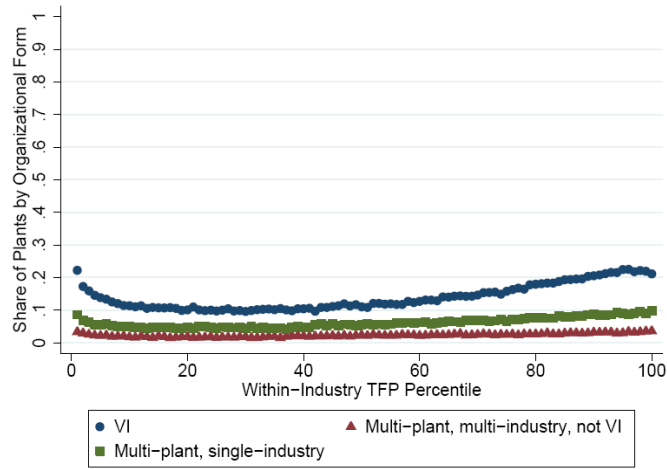
Notes: This figure shows, by organizational form, density estimates of firm sizes (measured by logged total deployment, as shown on the horizontal axis) for multi-unit firms in 1997. See text for details.

Figure 4. Plant Type Measures and Organizational Forms—Share of All Plants

A. Output per Hour



B. TFP



C. Capital-to-Labor Ratio

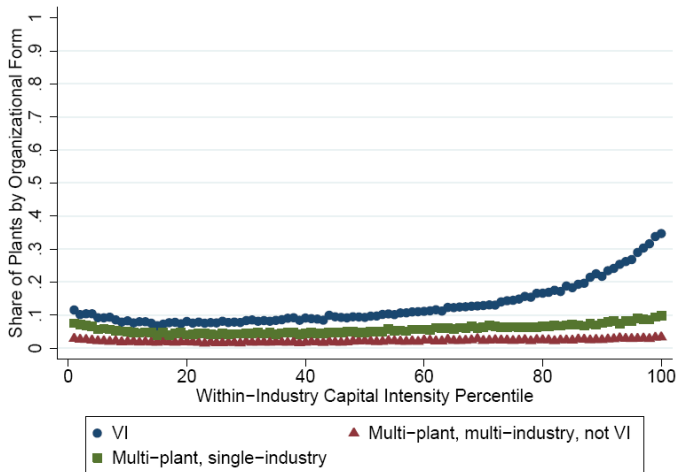
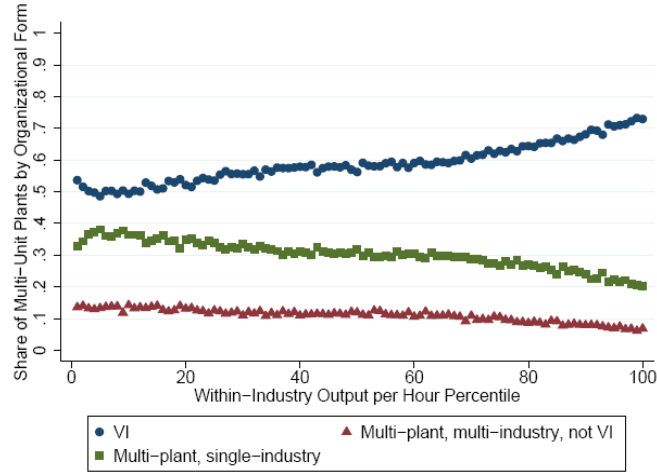
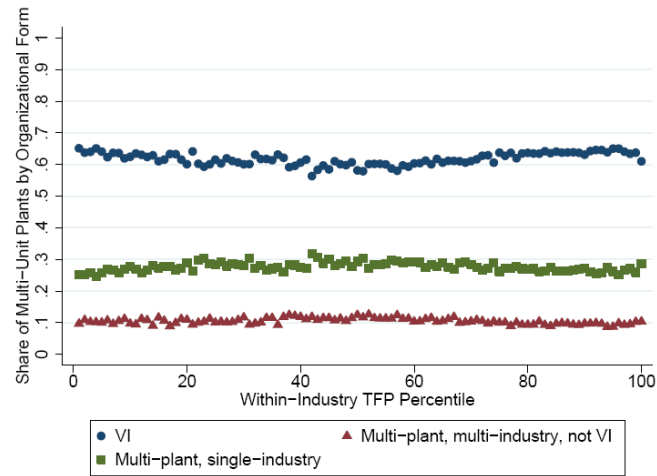


Figure 5. Plant Type Measures and Organizational Forms—Share of Multi-Unit Plants

A. Output per Hour



B. TFP



C. Capital-to-Labor Ratio

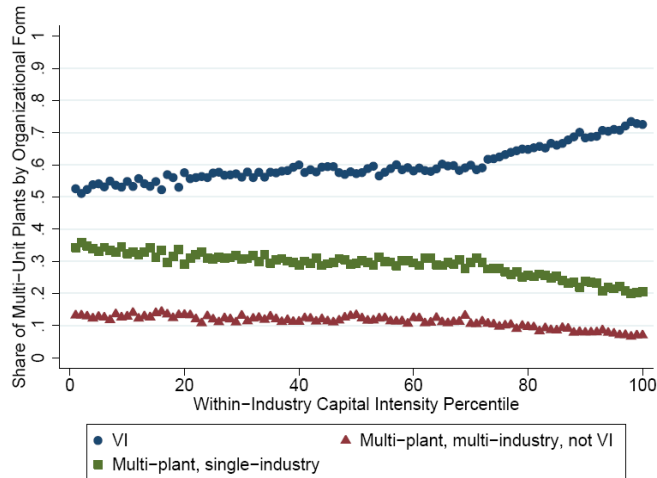
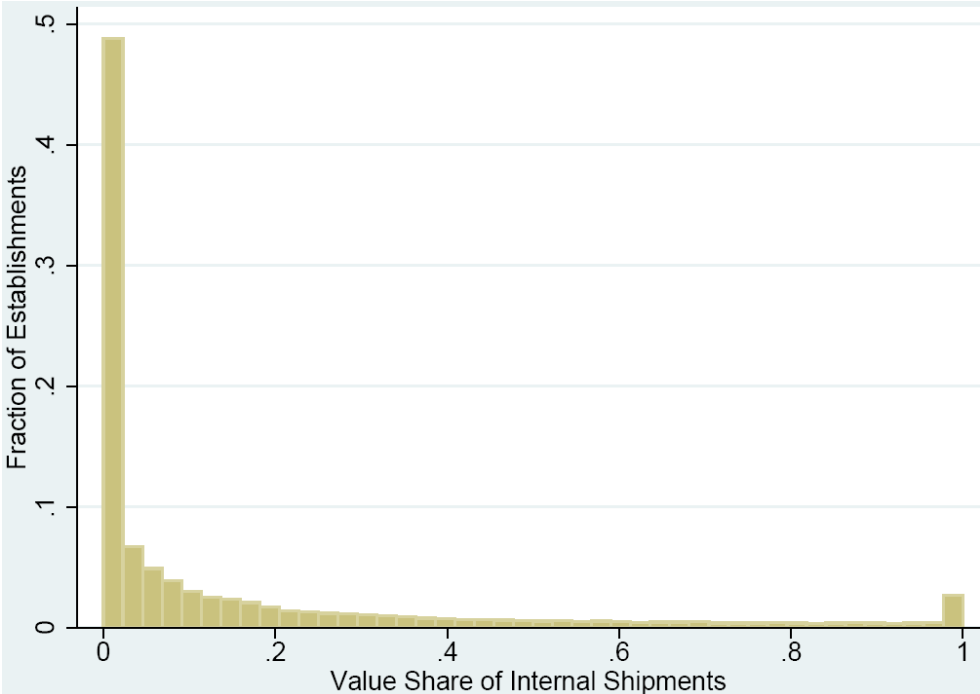


Figure 6. Shares of Intra-firm Shipments by Upstream Vertically Integrated Establishments

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

