Vertical Integration and Technology: Theory and Evidence

(Article begins on next page)

Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 342

Vertical Integration and Technology: Theory and Evidence

Daron Acemoglu, Philippe Aghion, Rachel Griffith and Fabrizio Zilibotti

December 2007

Vertical Integration and Technology: Theory and Evidence*

Daron Acemoglu,†
Department of Economics, Massachusetts Institute of Technology

Philippe Aghion,
Department of Economics, Harvard University and IFS

Rachel Griffith,
Institute for Fiscal Studies (IFS) London and UCL

Fabrizio Zilibotti
Institute for International Economic Studies Stockholm and IFS

August 31, 2005

*We thank Richard Blundell, Robert Gertner, Oliver Hart, Bengt Holmstrom, Ali Hortacsu, Paul Joskow, Ariel Pakes, Helen Simpson, Helena Svaleryd, John Van Reenen, and seminar participants at Canadian Institute of Advanced Research, Chicago, Harvard, IIES-Stockholm, MIT and NBER, and the CEPR-IUI Waxholms Conference. We also thank Christina Loennblad for assistance. This work contains statistical data from ONS which is Crown Copyright and reproduced with the permission of the controller of HMSO. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data.

†e-mails: daron@mit.edu, paghion@arrow.fas.harvard.edu, rgriffith@ifs.org.uk, fabrizio.zilibotti@iies.su.se.

Abstract

This paper investigates the determinants of vertical integration. We first derive a number of predictions regarding the relationship between technology intensity and vertical integration from a simple incomplete contracts model. Then, we investigate these predictions using plant-level data for the UK manufacturing sector. Most importantly, and consistent with theory, we find that the technology intensities of downstream (producer) and upstream (supplier) industries have opposite effects on the likelihood of vertical integration. Also consistent with theory, both these effects are stronger when the supplying industry accounts for a large fraction of the producer's costs. These results are generally robust and hold with alternative measures of technology intensity, with alternative estimation strategies, and with or without controlling for a number of firm and industry-level characteristics.

JEL: L22, L23, L24, L60.

Keywords: hold-up, incomplete contracts, internal organization of the firm, investment, residual rights of control, R&D, technology, UK manufacturing, vertical integration.

Abstract

This paper investigates the determinants of vertical integration. We first derive a number of predictions regarding the relationship between technology intensity and vertical integration from a simple incomplete contracts model. Then, we investigate these predictions using plant-level data for the UK manufacturing sector. Most importantly, and consistent with theory, we find that the technology intensities of downstream (producer) and upstream (supplier) industries have opposite effects on the likelihood of vertical integration. Also consistent with theory, both these effects are stronger when the supplying industry accounts for a large fraction of the producer's costs. These results are generally robust and hold with alternative measures of technology intensity, with alternative estimation strategies, and with or without controlling for a number of firm and industry-level characteristics.

JEL: L22, L23, L24, L60.

Keywords: hold-up, incomplete contracts, internal organization of the firm, investment, residual rights of control, R&D, technology, UK manufacturing, vertical integration.

Abstract

This paper investigates the determinants of vertical integration. We first derive a number of predictions regarding the relationship between technology intensity and vertical integration from a simple incomplete contracts model. Then, we investigate these predictions using plant-level data for the UK manufacturing sector. Most importantly, and consistent with theory, we find that the technology intensities of downstream (producer) and upstream (supplier) industries have opposite effects on the likelihood of vertical integration. Also consistent with theory, both these effects are stronger when the supplying industry accounts for a large fraction of the producer's costs. These results are generally robust and hold with alternative measures of technology intensity, with alternative estimation strategies, and with or without controlling for a number of firm and industry-level characteristics.

JEL: L22, L23, L24, L60.

Keywords: hold-up, incomplete contracts, internal organization of the firm, investment, residual rights of control, R&D, technology, UK manufacturing, vertical integration.

1. Introduction

Many experts believe that recent technological developments and globalization are transforming the internal organization of the firm. For example, it is argued, new technologies, especially information technology, are creating a shift from the old integrated firms towards more delayered organizations and outsourcing.[1] It is also often maintained that the greater competitive pressures created by both globalization and advances in information technology favor smaller firms and more flexible organizations that are more conducive to innovation.[2]

Despite the importance of these issues in the public debate and a large literature on vertical integration,[3] the economics profession is still far from a consensus on the empirical determinants of vertical integration in general, and the relationship between technological change and vertical integration in particular. This paper provides a simple incomplete contracts model of vertical integration and, in the light of the predictions of this model, presents detailed empirical evidence on the determinants of vertical integration.

The two leading theories of vertical integration are the "Transaction Cost Economics" (TCE) approach of Williamson (1975, 1985) and the "Property Right Theory" (PRT) approach of Grossman and Hart (1986) and Hart and Moore (1990).[4] Both approaches emphasize the importance of incomplete contracts and ex post opportunistic behavior (hold up) on ex ante relationship-specific investments. The TCE approach views vertical integration as a way of circumventing the potential holdup problems, and thus predicts that vertical integration should be more common when there is greater specificity and holdup is more costly. The PRT approach, on the other hand, focuses on the role of ownership of assets as a way of allocating residual rights of control, and emphasizes both the costs and the benefits of vertical integration in terms of ex ante investment

[1] Breshanan et al. (1999) find that IT use is associated with more decentralized decision-making within firms. Similarly, Acemoglu et al. (2005) show that in a sample of French firms, higher productivity firms are more likely to be decentralized. Helper (1991), on the other hand, documents the increase in outsourcing in the U.S. automobile industry.

[2] See, for instance, Milgrom and Roberts (1990), Athey and Schmutzler (1995) and Marin and Verdier (2002, 2003). See also Feenstra and Hanson (1996, 1999), and Feenstra (1998) on trade and decentralization.

[3] We discuss the empirical literature below. On the theory side, see, among others, Klein, Crawford and Alchian (1978), Williamson (1975, 1985), Grossman and Hart (1986), Hart and Moore (1990), Bolton and Whinston (1993), Aghion and Tirole (1994a,b and 1997) and Legros and Newman (2003), and the surveys in Holmstrom and Tirole (1989) and Hart (1995). See also the models of vertical integration in the context of international trade equilibria or economic growth, such as McLaren (2000), Grossman and Helpman (2002, 2003), Antras (2003) or Acemoglu, Aghion and Zilibotti (2004).

[4] See Whinston (2001) and Joskow (2003) for recent discussions.

incentives. To illustrate the central insight of the PRT, consider a relationship between a supplier (upstream firm) and a (downstream) producer. Moreover, suppose that only two organizational forms are possible: (backward) vertical integration, where the downstream producer buys up the upstream supplier and has residual rights of control, and non-integration (outsourcing), where the producer and supplier are separate firms. In this world, vertical integration does not automatically improve efficiency. Instead, by allocating the residual rights of control to the producer, who has ownership and thus control of the assets if there is a breakup of the relationship, vertical integration increases the producer's bargaining power and encourages its investment. However, by the same mechanism, it also reduces the supplier's ex post bargaining power and, hence, her incentives to invest. Non-integration, on the other hand, gives greater investment incentives to the supplier. Consequently, vertical integration has both costs and benefits in terms of ex ante investments, and its net benefits depend on whether the investments of the producer or those of the supplier are more important for the output and success of the joint venture.

While the key predictions of the TCE approach could be tested by investigating the relationship between measures of specificity and vertical integration, as also emphasized by Whinston (2001), the PRT approach is harder to test because it makes no predictions about the overall level of specificity and vertical integration.

In this paper, we develop a simple methodology to study the forces emphasized by the PRT approach. First, we shift the focus from relationship-specific investments to technology intensity. The presumption is that parties making technology investments, especially in R&D, are subject to holdup, and this will lead to the type of problems highlighted by the TCE and PRT approaches.[5] Second, we consider the relationship between pairs of supplying and producing industries and focus on the prediction that vertical integration should affect the investment incentives of suppliers and producers in *opposite* directions. Our approach therefore exploits cross-industry (cross-product) implications of the PRT.[6]

First, we develop these points using a simple theoretical framework and derive a

[5] This could be for a variety of reasons. First, R&D investments are often made for technologies specific to each firm (or their mix of products). Second, associated with any technological investment, parties are also likely to make specific investments. Finally, market imperfections, for example search frictions, typically turn "technologically general" investments into specific investments (e.g., Acemoglu, 1996, Acemoglu and Pischke, 1999).

[6] Although the specific regressions estimated in this paper are motivated by the PRT approach, the results are informative about, and could be consistent with, other approaches to vertical integration. In the concluding section, we discuss how these results could be reconciled with other theories.

number of predictions that are testable with the data we have available. The framework highlights that backward vertical integration gives greater investment incentives to the producer, while forward vertical integration encourages supplier investment. Non-integration provides intermediate incentives to both parties. This analysis leads to two key predictions:

1. The importance of the technology intensity of the producer and supplier should have opposite effects on the likelihood of vertical integration.

2. Vertical integration should be more responsive to the technology intensities of both the supplier and the producer when the supplier accounts for a larger fraction of the input costs of the producer.

We investigate these predictions, and other determinants of vertical integration, using detailed data on all British manufacturing plants from the UK Census of Production (ARD). To identify the effects of both supplier and producer technology, we look across all manufacturing industries. Using this dataset and the UK Input-Output table, we calculate two measures of vertical integration, defined at the level of firm-industry-pair (more precisely, for firm i producing product j with input from industry k). The first measure is a dummy variable indicating whether the firm owns a plant producing input k necessary for product j. The second measure calculates how much of the inputs from industry k, necessary for the production of j, the firm can produce in-house.[7]

Using the ratio of R&D expenditures to value added (calculated from a sample pre-dating our vertical integration measures), we find the following correlations in the data:

- Consistent with prediction 1 above, technology (R&D) intensities of the producing (downstream) and supplying (upstream) industries have effects with opposite signs on the likelihood of vertical integration.

- Consistent with prediction 2 above, the effect of the technology intensities of both the producing and the supplying industry are substantially larger when the share of costs of the supplying industry in the total costs of the producing industry (for short, "share of costs") is high.

- We also find that technology intensity of the producing industry is associated with more vertical integration, while technology intensity in the supplying industry is

[7] These measures do not distinguish between backward or forward integration, since we do not observe who has residual rights of control.

associated with less integration. In light of our model, this pattern is consistent with the marginal form of vertical integration in the data being backward integration. In addition, we find a positive association between the share of costs and vertical integration, which is also consistent with the view that the marginal form of vertical integration in the data is backward integration.

We subject our basic results to a series of robustness checks. The results are generally robust. First, including a range of firm-level covariates does not change the relationship between R&D intensity and vertical integration. Second, the results are broadly similar when we restrict attention to multiplant firms and control for firm level fixed effects.[8] Third, the results are similar when we proxy for technology intensity by physical investments rather than R&D. Fourth, the results are robust to excluding top or bottom quartiles of firms by size and to using an alternative measure of vertical integration. Finally, the results are also similar when we use a probit model rather than a linear probability model.

We also investigate the relationship between competition (measured as the number of firms in supplying and producing industries) and vertical integration. Our results here are consistent with theory, and indicate that having more firms in the supplying industry reduces the likelihood of vertical integration, while a larger number of firms in the producing industry increases it.

It is useful to emphasize that what we uncover are correlations, not necessarily causal relations. In our regressions, a measure of vertical integration is on the left-hand side, and industry and firm characteristics are on the right-hand side. However, in theory, and most likely in practice, vertical integration also affects technology choices. Moreover, other factors omitted in the regression could influence both vertical integration and technology intensity, and in a cross-industry regression there are many potential omitted variables. Although our fixed effects regressions control for many such omitted characteristics, there is still a concern regarding causality.

As an imperfect attempt at dealing with the endogeneity problem, we report results where the technology intensity of each industry is instrumented with the technology intensity of the same industry in the United States. This instrumentation strategy generally yields results similar to, and in fact quantitatively larger than, the ordinary least

[8] In particular, while the main effect of producer R&D intensity is no longer statistically significant, both supplier R&D intensity and the interaction between both supplier and producer R&D intensities and share of costs remain significant. When we control for endogenous selection, the effect of producer R&D intensity is again statistically significant.

squares strategy.[9] Overall, we conclude that there is an interesting pattern in the data, with technology intensity of producing and supplying industries having opposite effects on the likelihood of vertical integration. This pattern should be important in evaluating the predictions of a range of different theories of vertical integration (even though we have motivated our empirical approach from a specific theory based on incomplete contracts).

In addition to the theoretical studies mentioned above, this paper is related to a large empirical literature on vertical integration. In contrast to our approach, most empirical studies of vertical integration are motivated by the TCE approach and focus on a single industry. These include Joskow's (1987) seminal paper on ownership arrangements in electricity generating plants, Stuckey's (1983) study of integration between aluminium refineries and bauxite mines, Monteverde and Teece's (1982) investigation of integration in the automobile industry, Masten's (1984) work on the aerospace industry, Ohanian's (1994) work on the pulp and paper industry, and Klein's (1998) work on the Fisher Body and General Motors relationship. More recently, important papers by Baker and Hubbard (2000, 2002) study the trucking industry, Lerner and Merges (1998) consider the biotech sector, Woodruff (2002) studies integration in the Mexican footwear industry, Chipty (2001) investigates vertical integration and market foreclosure in the cable television industry, and Hortacu and Syverson (2005) study vertical integration in the U.S. cement industry. The only cross-industry evidence relevant to our investigation of which we are aware is due to Caves and Bradburd (1988), who document a positive cross-industry correlation between measures of specificity and vertical integration, and from Antras (2003), who looks at the share of intra-firm imports over total imports for 23 U.S. industries and relates this to capital intensity. We are not aware of any other papers investigating the prediction that the technology intensity of suppliers and producers have opposite effects on vertical integration decisions.

The paper is organized as follows. Section 2 presents the theoretical framework and derives the main testable implications. Section 3 details the construction of our measure of vertical integration, and also discusses data sources and the construction of the other key variables. Section 4 presents the main results. Section 5 discusses robustness checks and additional tests. Section 6 briefly investigates the effect of competition in producing and supplying industries on vertical integration. Finally, Section 7 discusses alternative theoretical approaches that may account for the correlations presented in this paper and

[9] However, when we simultaneously instrument both for the main effects and the interaction terms, the results are imprecisely estimated.

concludes.

2. Theory

The basic model is an extension of Grossman and Hart (1986), henceforth GH. We consider a one period relationship between a producer and a supplier, which are both risk-neutral. Both parties can undertake technological investments to increase the productivity of the relationship. Throughout, we assume these investment decisions have a *specific* component in that greater technology intensity leads to a greater possibility of holdup. Decision rights over these investments cannot be transferred between the two parties, for example, because the investments require tacit knowledge or human capital. This implies that the producer cannot make the supplier's investments, or vice versa.

As is standard in this literature, we assume that the investments and the output of the relationship are non-verifiable. Consequently, neither contracts conditional on investments nor contracts specifying rules for ex post revenue-sharing are possible. However, before investments and production take place, the parties can choose an organizational form and transfers. We denote the amount of ex ante transfer to party i conditional on the organizational form z by $T_i(z)$, where P and S denote the producer and the supplier, respectively. The organizational form can be either backward vertical integration (VIB), where the supplier is employed by the producer, or forward vertical integration (VIF), where the producer is employed by the supplier, or non-integration/outsourcing (NI), where the two parties remain independent.

The timing of events in this relationship is as follows:

1. The producer offers an organizational form (ownership structure) $z \in \{VIB, NI, VIF\}$ and associated transfers, $T_P(z)$ and $T_S(z)$, such that $T_P(z) + T_S(z) = 0$. There are no credit constraints, implying that $T_i(z)$ can be negative.

2. The supplier decides whether to accept or reject the offer. If the offer is not accepted, the game ends with payoffs $\left\{O_P^{NI}, O_S^{NI}\right\}$ defined below. Otherwise, the producer and the supplier simultaneously choose their investments, e_P and e_S.

3. The supplier and the producer bargain over the division of the revenue, according to the Nash bargaining solution given the organizational form z. Output is realized and shared.[10]

[10]In this game form, the assumptions that the producer makes the organizational form offer and that the parties receive their non-integration outside options are without loss of any generality. Moreover,

The production technology of the relationship is:

$$F(x_S, e_P, e_S) = \phi x_S(pe_P + se_S + 1) + (1 - \phi)(pe_P + 1). \quad (2.1)$$

The first term in (2.1) is the output generated by the producer and the supplier conditional on the supplier providing a customized (relationship-specific) input, denoted $x_S = 1$. If $x_S = 0$ and this input is not supplied, these activities generate no revenue. The value of the relationship can be further increased by the producer's and the supplier's investments, e_P and e_S. The parameters p and s designate the relative importance of investments by the producer and the supplier, and $\phi \in (0, 1)$ corresponds to the share of the producer's inputs accounted for by the supplier.[11] Note that ϕ also determines the importance of the supplier's investment, e_S.[12] This production function has also normalized the level of output in the absence of any investments to 1, which is without any loss of generality. The feature that there are no complementarities between the investments of the supplier and the producer is for simplicity, and highlights the fact that, for the results we emphasize, such complementarities are not essential.

To simplify the expressions, we assume that the supplier can provide the basic input x_S at no cost, and also that the costs of investment for both parties are quadratic:

$$\Gamma_P(e_P) = \frac{1}{2}e_P^2 \text{ and } \Gamma_S(e_S) = \frac{1}{2}\phi e_S^2. \quad (2.2)$$

Notice that the investment costs of the supplier are multiplied by ϕ. This ensures that the costs are proportional to the scale of operation and that the socially optimal levels of both e_P and e_S are independent of ϕ.[13]

In the event of disagreement, the two parties receive their outside options, which depend on the organizational form. We denote the outside option of party i under organizational form z by O_i^z.

following other papers in this literature, we are using the Nash bargaining solution (see Binmore, Rubinstein and Wolinsky, 1986, for a potential justification for Nash bargaining and a discussion of alternative bargaining rules), but our qualitative results do not depend on Nash bargaining.

[11] With competitive spot market transactions and without any specific investments, i.e., $e_P = e_S = 0$, ϕ would exactly correspond to the share of costs of the producer accounted for by the supplier in question. Although with positive investments and ex post bargaining, there will be a wedge between the two, we refer to ϕ as the "share of costs" to simplify the terminology.

[12] Symmetrically, we could introduce another parameter, say η, to capture the importance of the producer for the supplier. Comparative statics with respect to η are very similar to those with respect to ϕ. We omit this generalization to reduce notation, and discuss empirical results regarding the effect of a measure related to η in subsection 5.

[13] The socially optimal levels of investment are $e_P = p$ and $e_S = s$. Modifying the supplier's cost function to $\Gamma_S(e_S) = e_S^2/2$ introduces an implicit "scale economies", and an increase in ϕ makes the supplier's investment more profitable (the socially optimal level of investment for the supplier becomes $e_S = s\phi$). Consequently, the comparative static results with respect to ϕ become ambiguous.

With *backward vertical integration* (VIB), the producer owns all assets, and in the event of ex post breakup the supplier simply walks away from the firm without receiving anything. The producer, who has residual control rights, keeps all the assets and the customized input, but lack of cooperation from the supplier causes the loss of a fraction λ of the supplier's investment, so the "effective investment" of the supplier is reduced to $(1-\lambda)e_S$ where $\lambda \in [0,1)$.[14] Therefore, the outside options of the supplier and the producer in this case are:

$$O_S^{VIB}(e_P, e_S) = 0 \text{ and } O_P^{VIB}(e_P, e_S) = F(x_S = 1, e_P, (1-\lambda)e_S).$$

With *non-integration* (NI), the supplier and the producer own their separate firms and assets. In case of disagreement, the producer does not receive the customized input from the supplier ($x_S = 0$), and consequently, generates no output from the part of the operations relying on those inputs. The supplier can sell her input in the market, but suffers in this case some revenue loss because of the specificity of the input to this producer. Therefore, the outside options under non-integration are:

$$O_S^{NI}(e_P, e_S) = \theta\phi(se_S + 1) \text{ and } O_P^{NI}(e_P, e_S) = F(x_S = 0, e_P, e_S) = (1-\phi)(pe_P + 1),$$

where $\theta \in [0,1)$ is an inverse measure of how much the supplier loses if she sells the input outside of the specific relationship.[15] The general equilibrium determination of θ is beyond the scope of our paper. Here it is treated as exogenous and in the empirical section it is proxied by the relative number of producers to suppliers (with more producers, it might be easier for the supplier to find a suitable buyer to her input in the secondary market).

The third organizational form is *forward vertical integration* (VIF), where the supplier owns all the assets. In this case, with a similar reasoning to before, the outside options are:

$$O_S^{VIF}(e_P, e_S) = F(x_S = 1, (1-\lambda')e_P, e_S) \text{ and } O_P^{VIF}(e_P, e_S) = 0,$$

where $\lambda' \in [0,1)$ is the fraction of the producer's investment the supplier loses in case of disagreement.

[14] Alternatively, λ can be interpreted as the fraction of investment which is incurred at the end of the period by the supplier to fine-tune the quality of the input. The supplier would not undertake this investment in the event of disagreement.

[15] It is possible to also allow a secondary market in which the producer can purchase a less suitable input, in which case his outside option would be:

$$O_P^{NI}(e_P, e_S) = (1-\phi)(pe_P + 1) + \rho\phi(pe_P + 1),$$

where $\rho + \theta < 1$. This modification has no effect on the results.

Let y_i^z denote the output accruing to party i under organizational form z. Symmetric Nash bargaining implies that:

$$y_i^z(e_P, e_S) = O_i^z(e_P, e_S) + \frac{1}{2}\left[F(x_S = 1, e_P, e_S) - O_P^z(e_P, e_S) - O_S^z(e_P, e_S)\right], \qquad (2.3)$$

where the term in square brackets is the relationship-specific surplus over which bargaining takes place, and is positive for all $z \in \{VIB, NI, VIF\}$. The important feature is that each party's share of revenue will be increasing in her own outside option, and decreasing in that of the other party. This feature creates a link between outside options and investment incentives, and through this channel, between organizational forms and investment incentives.

Finally, the utility of party $i \in \{P, S\}$ can be expressed as:

$$U_i^z(y_i(e_P, e_S), e_i) = y_i^z(e_P, e_S) - \Gamma_i(e_i) + T_i(z). \qquad (2.4)$$

We can now characterize the unique equilibrium of the game specified in the previous subsection. Unless otherwise specified, we refer to an equilibrium by the on-the-equilibrium-path actions and revenues, $(z^*, T_P^*, T_S^*, e_P^*, e_S^*, y_P^*, y_S^*)$.

It is useful to define the "total surplus" of the relationship as:

$$S^z = U_S^z(y_S^z(e_P^*(z), e_S^*(z)), e_S^*(z)) + U_P^z(y_P^z(e_P^*(z), e_S^*(z)), e_P^*(z)),$$

where $e_i^*(z)$ denotes party i's optimal investment under ownership structure z. Using equations (2.3) and (2.4) and the fact that $T_S(z) + T_P(z) = 0$ gives the total surplus of the relationship as:

$$S^z = F(x_S = 1, e_P^*(z), e_S^*(z)) - \Gamma_P(e_P^*(z)) - \Gamma_S(e_S^*(z)). \qquad (2.5)$$

Since both parties have access to perfect credit markets and ex ante transfers, the subgame perfect equilibrium will always pick the organizational form that maximizes the surplus, S.[16] In other words, $S^{z^*} \geq S^z$ for all $z \in \{VIB, NI, VIF\}$.[17]

[16] With credit constraints, the less constrained party may become the owner even when this structure does not maximize the ex ante social surplus, because the other party does not have the cash to compensate the first party for giving up ownership (see, for example, Aghion and Tirole (1994a), or Legros and Newman (2003)).

[17] Suppose that the equilibrium involves z^*, but $S^{z^*} < S^{z'}$. Then the producer, which has the bargaining power in the first stage of the game, can propose z' together with a compensating transfer to the supplier, and increase its payoff. Namely, she can offer

$$T_S(z') = T_S(z^*) + y_S^{z^*} - y_S^{z'} - \Gamma_S(e_S^*(z^*)) + \Gamma_S(e_S^*(z')) + \varepsilon$$

with $\varepsilon > 0$, which would be at least as attractive for the supplier, and for $\varepsilon < S^{z'} - S^{z^*}$, also profitable for the producer, thus yielding a contradiction.

We now characterize the equilibrium by calculating the levels of social surplus under backward integration (S^{VIB}), non-integration (S^{NI}) and forward integration (S^{VIF}). The equilibrium organizational form is then given by $z^* = \arg\max_{z\in\{VIB,NI,VIF\}} S^z$.

Equilibrium investments are determined as the Nash equilibrium of a game where each party chooses its investment so as to maximize utility, given the other party's investment and the ownership structure. More formally, the equilibrium conditional on the ownership structure z is given by the pair $\{e_S^*(z), e_P^*(z)\}$ such that:

$$e_P^*(z) = \max_{e_P}\{y_P^z(e_P, e_S^*(z)) - \Gamma_P(e_P)\} \text{ and } e_S^*(z) = \max_{e_S}\{y_S^z(e_P^*(z), e_S) - \Gamma_S(e_S)\},$$

where the expressions for $y_i^z(.)$ are given in (2.3), and those for Γ_P and Γ_S are given in (2.2). The Nash equilibrium investment levels under each of the three ownership structures can be calculated as:

$$e_P^*(VIB) = p \text{ and } e_S^*(VIB) = \frac{\lambda}{2}s \tag{2.6}$$

$$e_P^*(NI) = \left(1 - \frac{\phi}{2}\right)p \text{ and } e_S^*(NI) = \frac{1+\theta}{2}s \tag{2.7}$$

$$e_P^*(VIF) = \frac{\lambda'}{2}p \text{ and } e_S^*(VIF) = s. \tag{2.8}$$

These expressions highlight the effect of the different ownership structures on investment incentives. The investment made by the producer is highest under backward vertical integration (i.e., $e_P^*(VIB) > e_P^*(NI) > e_P^*(VIF)$), while that made by the supplier is highest under forward vertical integration (i.e., $e_S^*(VIF) > e_S^*(NI) > e_S^*(VIB)$). Furthermore, most relevant for our empirical analysis, backward vertical integration increases the investment of the producer and reduces the investment of the supplier relative to non-integration. This is a fundamental result in this class of models: (backward) vertical integration reduces the supplier's outside option, and increases the share of the surplus accruing to the producer. It therefore discourages supplier investment and encourages producer investment. Another important feature is that with non-integration, the investment level of the producer is decreasing in ϕ, since a greater share of costs increases the scope for holdup by the supplier. Also with non-integration, the investment of the supplier is increasing in θ because a greater θ provides her with a better outside market (the outside market is irrelevant for the other organizational forms, since one of the parties has residual rights of control over the input and the assets).

Finally, substituting for $e_S^*(z)$ and $e_P^*(z)$ in (2.5), we obtain the total surplus under the three ownership structures, S^{VIB}, S^{NI}, and S^{VIF}, and the comparison of the surpluses gives the following proposition (the relevant expressions and the proof are in Appendix A):

Proposition 1. *There exist $\underline{r}$, $\overline{r}$, and $\widehat{r}$ such that the unique subgame perfect equilibrium ownership structure, z^*, is given as follows:*

- *If $\underline{r} < \overline{r}$, then $z^* = VIB$ for $p/s > \overline{r}$, $z^* = NI$ for $p/s \in (\underline{r}, \overline{r})$, and $z^* = VIF$ for $p/s < \underline{r}$. Moreover,*

$$\frac{\partial \overline{r}}{\partial \phi} < 0,\ \frac{\partial \underline{r}}{\partial \phi} > 0,\ \frac{\partial \overline{r}}{\partial \theta} > 0 \text{ and } \frac{\partial \underline{r}}{\partial \theta} < 0.$$

- *If $\underline{r} \geq \overline{r}$, then $z^* = VIB$ for $p/s > \widehat{r}$, and $z^* = VIF$ for $p/s < \widehat{r}$. Moreover,*

$$\frac{\partial \widehat{r}}{\partial \phi} > 0 \text{ and } \frac{\partial \widehat{r}}{\partial \theta} = 0.$$

This proposition gives the most important comparative static results that will be empirically investigated in the second part of the paper. First, the proposition shows that, given the other parameters, the choice of organizational form depends on the ratio of p to s. When this ratio is high, backward integration is the equilibrium organizational form; for intermediate values, non-integration may emerge; and when this ratio is small, forward integration results in equilibrium. Intuitively, backward integration becomes more likely when p is large because, in this case, the tasks in which the producer specializes are highly "technology intensive" (i.e., the investment of the producer is more important), so increasing the producer's investment is the first priority. Backward vertical integration achieves this by increasing the producer's outside option and reducing that of the supplier. In contrast, when s is large, backward integration becomes less likely, since the investment of the supplier is now more important, and by reducing the supplier's outside option backward integration discourages her investment. The opposite comparative static results apply for forward integration.

Second, as long as we are in the first case with $\underline{r} < \overline{r}$, non-integration is a possibility, which is clearly the empirically relevant case. In this case, an increase in ϕ makes backward integration more likely relative to non-integration, and also non-integration more likely relative to forward integration. A greater share of costs (of the supplier's inputs in the producer's total costs) increases the degree to which the producer will be held up by the supplier. Backward vertical integration becomes more preferable because it avoids this problem. In addition, this result also implies that there are important interaction effects: the effect of p/s on vertical integration is amplified by

ϕ. To see this let us focus of the comparison between non-integration and backward integration, and denote the surplus difference between these two organizational forms by $\Delta^B S \equiv S^{VIB} - S^{NI}$. Then we have that:

$$\frac{\partial^2 \Delta^B S}{\partial \phi \partial p} > 0 \text{ and } \frac{\partial^2 \Delta^B S}{\partial \phi \partial s} < 0.$$

This prediction is also quite intuitive. It suggests that when the relationship between the producer and the supplier is less important, their respective technology intensities should have less effect on integration decisions.

Finally, a greater θ makes non-integration more likely relative to backward integration; with a greater θ, the supplier invests more under non-integration because she has a better outside option, and this makes non-integration a more desirable organizational form. If we interpret θ as the degree of competition in the market, this result would imply that, consistent with some of the claims made in the popular press, greater competition encourages non-integration. However, a more appropriate interpretation might be that θ is a function of the ratio of producers to suppliers in the market since, with a larger number of producers, the supplier is more likely to find a suitable match in the secondary market after a breakup.

In summary, the most important empirical prediction of this framework is that the technology intensity of the producer and that of the supplier should have opposite effects on the likelihood of vertical integration. In addition, there should be interaction effects between the producers' and suppliers' technology intensity on the one hand and the share of costs on the other, such that a greater share of costs should increase the magnitude of both effects. The rest of the predictions depend on whether the relevant margin in the data is backward or forward integration. In the case of backward integration, the results suggest that greater technology intensity of producers should be associated with greater vertical integration, greater technology intensity of suppliers should be associated with less vertical integration, and a greater share of costs should encourage vertical integration. Finally, we may also expect the number of producing firms relative to supplying firms to encourage non-integration.

3. Data and Measurement

3.1. Vertical Integration

Central to our empirical strategy is a measure of vertical integration, which we define at the firm line of business level for each potential supplying industry. Namely, for each firm $i = 1, 2, .., N$, our first measure is a dummy for whether, for each product (industry) $j = 1, 2, .., J$ it is producing, the firm owns a plant in industry $k = 1, 2, .., K$ supplying product j:

$$vi_{ijk} = \begin{cases} 0 & \text{if the firm does not own a plant in industry } k \text{ supplying industry } j \\ 1 & \text{if the firm owns at least one plant in industry } k \text{ supplying industry } j \end{cases}. \tag{3.1}$$

This measure provides a direct answer to the question of whether the firm can supply some of its own input k necessary for the production of product j. However, it does not use any information on how much the firm produces in its own plants.

We also construct an alternative (continuous) measure using this information. Let c_{ij} denote the total cost (including intermediate, capital and labor costs) of firm i in producing j, and w_{jk} denote the proportion of total costs of producing j that are made up of input k, which is obtained from the UK Input-Output table. We can think of $c_{ij}w_{jk}$ as the firm's demand for input k for product j (to obtain the firm's total demand for k we sum over j). Let y_{ik} denote the amount of k that firm i produces. The alternative measure of the degree of vertical integration of firm i in the industry pair jk is calculated as:[18]

$$\overline{vi}_{ijk} = \min\left\{\frac{y_{ik}}{c_{ij}w_{jk}}, 1\right\}. \tag{3.2}$$

When a firm produces several different products that demand input k, and where the total demand is greater than what can be supplied by the firm itself, we assume that it allocates the input across plants proportionately to their demand, so that the measure becomes

$$\overline{vi}_{ijk} = \min\left\{\frac{y_{ik}}{\sum_j c_{ij}w_{jk}}, 1\right\}. \tag{3.3}$$

In practice, there is little difference between vi_{ijk} and $\overline{vi}_{ijk}$, because when a firm owns a plant in a supplying industry, it is typically sufficient to cover all of its input requirements from that industry. So for most of our analysis, we focus on the vi_{ijk} measure.

Our main source of data is the Annual Respondents Database (ARD).[19] This is collected by the UK Office of National Statistics (ONS) and firms have a legal obligation to reply. These data provide us with information on input costs and output for each

[18] Davies and Morris (1995) construct a related index with more aggregate data, while Fan and Lang (2000) measure corporate relatedness using a similar measure.

[19] This dataset is constructed using the data from the Annual Business Inquiry from 1998 and onward. Before that the name of the Inquiry was The Annual Census of Production (ACOP). See Griffith (1999) and Barnes and Martin (2002) for a description of these data.

production plant located in the UK at the 4-digit industry level and on the ownership structure of these plants.[20] These data do not, however, tell us directly whether a plant purchases inputs from a related plant in the same firm. Data on the demand for intermediate inputs are available at the 2/3-digit industry level from the Input-Output Domestic Use Table. The Input-Output table contains information on domestic input flows between 77 manufacturing industries, giving 5,929 pairs of producing-supplying industries, 3,840 of which have positive flows. Appendix Table A.1 lists all 77 (supplying) industries together with their largest purchaser and other information.

Because of the level of aggregation of the Input-Output table, one difficulty arises when we look at industry pairs where the input and output are in the same 2/3-digit industry. In this case, we consider a firm to be vertically integrated only if it has plants in more than one of the 4-digit industries within that 2/3-digit industry.

Further details on the construction of these measures are provided in Appendix B.

3.2. Technology Intensity and the Share of Costs

Our main measure of technology intensity is R&D intensity, but we also report robustness results using investment intensity. Both these measures are constructed at the industry level rather than at the firm level for two reasons: first, our methodology focuses on the technology of an industry, not on whether a specific firm is more R&D or investment intensive; second, we need measures that apply both to integrated and non-integrated relationships, which naturally takes us to the industry level.

R&D intensity is calculated as R&D expenditure divided by total value added.[21] We use R&D data pre-dating the vertical integration sample (1994-1995). The total value added in the denominator includes all firms in the industry (both those performing R&D and those that do not).

R&D intensity is our preferred measure since it is directly related to investment in new technologies. A possible concern is that the distribution of R&D across industries is rather skewed. Another concern might be that R&D could be spuriously correlated

[20] Data on employment is available for all plants. Data on other inputs and output are available at the establishment level. An establishment is often a single plant, but can also be a group of plants owned by the same firm that operate in the same 4-digit industry. We have input and output data on all establishments with over one hundred employees, data from smaller establishments are collected from a random stratified sample and values for non-sampled plants are imputed. Throughout, we exclude single plant firms with fewer than twenty employees.

[21] Differently from the U.S. data, in the UK data we observe firm-level R&D separated both by the industry of the firm conducting the R&D and the product category for which the R&D is intended. This enables us to have a more accurate measurement of R&D by producing and supplying industries.

with vertical integration; for example, because it is better reported in industries with many large firms and large firms are more likely to be vertically integrated (though, in many specifications, we also control for firm size).

For these reasons, we consider an alternative indicator of industry technology by looking at physical investment intensity. This information is reported at the level of the firm's line of business and can be directly linked to the producing or supplying part of vertically integrated firms. It is also more widely reported and less skewed both within and between industries. As with R&D data, we use data pre-dating the measure of vertical integration, 1992-95, and aggregate the data from the firm's line of business to the industry level. The disadvantage of this measure relative to R&D intensity is that physical investment intensity may be less related to technology and may also have a more limited firm-specific component, which is important for the model we have used to motivate our empirical analysis.

The share of costs between each industry pair jk, sc_{jk}, is calculated from the Input-Output table as the share of inputs from industry k in the total cost of industry j (£ of input k necessary to produce £1 of product j).

3.3. Descriptive Statistics

Table 1 gives descriptive statistics for the whole sample, and also for subsamples separated according to whether the producer (supplier) has a high or low R&D intensity. There are 3,840 industry pairs where the Input-Output table indicates that transactions occur. There are 46,392 manufacturing firms with twenty or more employees operating in the UK at some time over the period 1996-2001. Since individual firms seldom change their organization structure over this short time period we collapse the data into a single cross-section. An observation in our data represents a firm i producing product j which uses input k; this gives us 2,973,008 observations at the firm-industry pair level.

The first row of Table 1 gives the mean and standard deviation of the continuous measure of vertical integration, $\overline{vi}_{ijk}$. The mean is 0.008 with a standard deviation of 0.087, which shows that there is substantial variation across firms and at the sub-firm level. There is also substantial variation within industry pairs. To illustrate this, we calculate the average within-industry-pair standard deviation of $\overline{vi}_{ijk}$, which is 0.086 (not shown in the table). This indicates that, even within a relatively narrow industry-pair, there is as much variation in the extent of vertical integration as in the whole sample.

The low mean of this variable is driven by the large number of zeros. The mean of $\overline{vi}_{ijk}$ conditional on $\overline{vi}_{ijk} > 0$ (not shown) is 0.93. This indicates that if a firm can produce some of its inputs k in-house, it can typically produce *all* that input (k) necessary for production.[22] This motivates our focus on the simpler dummy variable vi_{ijk}, which indicates whether the firm owns a plant producing input k which it needs in the production of the product j (see equation (3.1)). Not surprisingly, the second row shows that the mean of this variable, 0.009, is very similar to that of $\overline{vi}_{ijk}$.

The other columns illustrate the differences in the extent of vertical integration when we separate firm-industry pairs by producer R&D intensity and supplier R&D intensity. These differences, which will be investigated in greater detail in the regression analysis below, indicate that vertical integration is higher when the R&D intensity of the producing industry is high. Interestingly, the descriptive statistics do not show any difference between vertical integration when we cut the sample by whether the R&D intensity of the supplying industry is high or low.[23] The regression analysis below will show a negative effect of supplier R&D intensity as well as supplier investment intensity on vertical integration, but due to nonlinearities in this relationship (see also Appendix Table A.3), the high-low cut does not show this result.

R&D intensity is positively correlated with investment intensity, although the correlation is quite low (0.251). The relatively weak correlation between these measures means that each measure is an imperfect proxy for the overall technology intensity of the sector, and consequently, there might be some attenuation bias in our estimates of the relationship between technology intensity and vertical integration. It also suggests that these measures capture different dimensions of technology intensity, so that it is useful to study the relationship between each of them and vertical integration separately.

The table also shows the means and standard deviations of the other main covariates, defined in Appendix B.

[22] Naturally this does not imply that if a firm is vertically integrated for one of its inputs, it is also vertically integrated for its other inputs. In fact, the mean of $\sum_k w_{jk}\overline{vi}_{ijk}$ conditional on $\overline{vi}_{ijk'} > 0$ for some k' is 0.053, so on average, across firms that are vertically integrated in any one input, firms are vertically integrated in around 5% of their total inputs demanded.

[23] When we group firms on the basis of investment intensity, we see that greater supplier investment intensity is associated with lower vertical integration. This cut of the data is not shown in Table 1 to save space.

4. Results

4.1. Benchmark Specification

Table 2 reports the main results. It reports estimates from the following linear probability model:

$$vi_{ijk} = \alpha sc_{jk} + \beta_P RD_j^P + \beta_S RD_k^S + X'_{ijk}\eta + \varepsilon_{ijk}, \tag{4.1}$$

where sc_{jk} is the share of costs, RD_j^P is R&D intensity in the producing industry j, RD_k^S is R&D intensity in the supplying industry k, X_{ijk} is a vector including the constant term and firm and industry characteristics (firm size and age, average firm size and age in producing and supplying industries). The main coefficients of interest are α, β_P and β_S. The regressions are at the firm industry-pair level, while some of the main regressors are at the (producing or supplying) industry level. For this reason, throughout all standard errors are corrected for clustering at the industry pair level.[24]

The first two columns of Table 2 consider the bivariate relationship between R&D intensity in the producing and supplying industries and vertical integration. Column 1 shows a positive and highly statistically significant relationship between R&D intensity in the producing industry and vertical integration. The estimate of β_P is 0.038 with a standard error of 0.006. Column 2 shows a negative and highly statistically significant relationship between R&D intensity in the supplying industry and vertical integration; the estimate of β_S is -0.010 (standard error of 0.002). These relationships are robust to the inclusion of other covariates in the rest of the table.

The third column includes both R&D intensity variables and the share of costs. The R&D intensity variables continue to be highly statistically significant, with coefficients close to those in columns 1 and 2 (0.040 and -0.007), while the share of costs is positive and also statistically significant. The pattern of opposite signs on R&D intensity of producing and supplying industries is consistent with the theoretical prediction derived above.

In addition, the directions of the effects of R&D intensities and the share of costs are consistent with the theory, as long as the relevant margin in the data is backward inte-

[24]There is also a potential correlation between observations for the same firm in different industry pairs. Unfortunately, we were unable to estimate a variance-covariance matrix with multiple random effects or multiple levels of clustering, due the large size of the dataset. Nevertheless, we believe that the downward bias in the standard errors should be small in our case, since, as noted in footnote 22, the probability of a firm that is vertically integrated for one of her inputs also being integrated for other inputs is relatively small. In any case, in Table 4, we estimate these models including a full set of firm fixed effects (for those firms that operate in more than one industry), which directly removes any potential correlation across firm observations.

gration (recall that greater technology intensity of producers should increase backward, but not forward, vertical integration). Since we find the same pattern in practically all of our specifications, from now on we take the relevant comparison to be the one between backward vertical integration and non-integration, which is also consistent with the greater prevalence of backward integration in the manufacturing sector documented in previous studies (see, for example, Joskow (1987) and Stuckey (1983)).

The theoretical model above also suggests the possibility of interaction effects between the share of costs and R&D intensity. To investigate this issue, we modify our estimating equation to

$$vi_{ijk} = \alpha sc_{jk} + (\beta_P + \gamma_P sc_{jk})\, RD_j^P + (\beta_S + \gamma_S sc_{jk})\, RD_k^S + X'_{ijk}\eta + \varepsilon_{ijk}, \qquad (4.2)$$

with γ_P and γ_S as the additional coefficients of interest. Theory suggests that γ_P should have the same sign as β_P, and that γ_S should have the same sign as β_P, so that the effects of R&D intensity in producing and supplying industries should be amplified when there is a greater share of costs. Throughout, when including interaction terms, we report the main effects evaluated at the sample mean, so that these estimates are comparable to those in the models without interaction effects.

The estimates in column 4 are consistent with the theoretical predictions. The main effects are close to those in the previous columns, and the interaction effects are large and statistically significant: γ_P is positive (1.112 with a standard error of 0.402), while γ_S is negative (-0.909 with a standard error of 0.353).

Columns 5 and 6 add a number of characteristics at the firm-line of business and industry level, namely firm size and age (in that line of business), and average firm size and average firm age in producing and supplying industries. All five coefficients of interest are robust, and remain close to their baseline values (the only minor exception is β_P, which declines from 0.040 in column 3 to 0.030 in column 6). The coefficients on the controls are also interesting. They indicate, for example, that larger and older firms are more likely to be vertically integrated, which is plausible. Furthermore, greater average firm size in the producing industry makes vertical integration more likely, while average firm size in the supplying industry appears to reduce the probability of integration. This opposite pattern of coefficients, with firm size in the producing industry having a positive effect, is also consistent with our conjecture that the relevant margin in the data is backward integration.

Overall, the results in Table 2 show an interesting pattern of opposite-signed effects from technology intensity in producing and supplying industries. They also show that these effects are magnified when the share of costs accounted for by the supplying industry in the total costs of the producing industry is large. These results are consistent with the predictions of the PRT approach discussed above. In addition, the direction of the effects is also consistent with the theory as long as the relevant margin in the data is backward integration.

Before further investigating the robustness of our findings, it is useful to discuss the economic magnitudes of the estimates in Table 2. The implied magnitude of the main effects is very small. For example, the coefficient of -0.013 in column 4 of Table 2 suggests that a one standard deviation (0.107) increase in the R&D intensity of supplying industry reduces the probability of vertical integration by slightly more than 0.1% ($-0.013\times0.107\approx-0.001$). However, this small effect applies at the mean of the distribution of share of costs, which is 0.010. If, instead, we evaluate the effect at the 90th percentile of the share of cost distribution, which is about 0.20, then the overall effect is much larger; again using the numbers from column 4, a one standard deviation increase in supplier R&D intensity leads to almost a 2% decrease in the probability of vertical integration ($-0.001+(-0.909\times0.19\times0.107)\approx0.02$). Similarly, the impact of producer R&D intensity is small when evaluated at the mean, about 0.2%, ($0.037\times0.055\approx0.002$), but sizable, 1.4%, when evaluated at higher levels of share of costs ($0.002+(1.104\times0.19\times0.055)\approx0.014$).

This pattern is in fact quite sensible. For the vast majority of industry pairs, the scale of the relationship between the producer and the supplier is so small that it would be surprising if technology intensity were of any great importance for integration decisions. Our theory should instead be relevant for industry pairs where the scale of the relationship (as measured by the share of costs) is large, and this is exactly where we see a significant positive effect of the R&D intensity of producing industries and a significant negative effect of the R&D intensity of supplying industries. This discussion also implies that the interaction effects are as important as are the main effects for the relevance of the pattern documented here.

4.2. Within Firm Variation

A more demanding test of the relationship between technology intensity and vertical integration is to investigate whether a particular firm is more likely to be vertically integrated in producing industries that are more technologically intensive and with supplying industries that are less technology intensive. This is done by estimating a model including firm fixed effects.

Naturally, this can only be investigate using multiplant firms, i.e., those that produce in more than one industry, which introduces a potential selection bias. At some level, this is mechanical; vertically integrated firms have to be multiplant firms. More generally, producer and supplier technology intensity may affect the likelihood of being a multiplant firm differentially, and if so, regressions on the subsample of multiplant firms may lead to biased estimates of the effect of technology (R&D) intensity on vertical integration.

In Table 3, we investigate both the robustness of our results to the inclusion of firm fixed effects and potential selection issues.

Column 1 reports our basic specification (without fixed effects) on multiplant firms only. Comparing this to column 4 of Table 2, a number of features are noteworthy. First, the number of observations is now 891,942 rather than 2,973,008 as in Table 2. Second, despite changes in coefficient estimates, the overall pattern is quite similar. In particular, there is a positive effect of producer R&D intensity and a negative effect of supplier R&D intensity on vertical integration. Both these effects are larger than those in Table 2. The interaction effects also have the expected signs and are about twice the size of the interaction effects in Table 2. This pattern of results is reassuring, since it shows that our main results in Table 2 were not driven by the contrast of single to multiplant firms. Within multiplant firms, those with greater producer R&D intensity are also more likely to be vertically integrated, while those with greater supplier R&D intensity are less likely to be vertically integrated.

Column 2 adds a full set of firm fixed effects to the specification in column 1. This has surprisingly little effect on the results. The coefficients on the share of cost, the interaction between producer R&D intensity and share of cost, supplier R&D intensity and the interaction between share of costs and supplier R&D intensity are essentially identical to those in column 1. The only change is in the main effect of producer R&D intensity, which falls from 0.073 to 0.024 and is no longer statistically significant.

However, it is important to emphasize that even though producer R&D intensity is insignificant when evaluated at the mean, the overall pattern is not very different; due to the significant interaction effect, producing industries with substantial R&D intensity are much more likely to vertically integrate activities that are important for their products. In fact, since the interaction effect is larger than in Table 2, the effect of a one standard deviation increase in producer R&D intensity on the probability of vertical integration for a pair at the 90th percentile of the share of costs distribution is now greater than in Table 2 at 3% ($0.03\times0.055+(2.741\times0.19\times0.055)\approx0.030$), rather than 1.4% when using the estimates from column 5 of Table 2. Similarly, now a one standard deviation increase in supplier R&D intensity has a larger effect, 5% (-0.016×0.107+(-

$2.482 \times 0.19 \times 0.107) \approx 0.052$), rather than 2% when using the estimate from column 5 of Table 2.

Column 3 repeats the model of column 2 including the full set of covariates. The results are similar to those in column 2, except that the coefficients on the level of producer and supplier R&D decline.

These specifications do not deal with the potential selection problem. In columns 4 and 5 we estimate a standard Heckman selection model. Column 4 shows estimates from the probit model of the probability of a firm being a multiproduct firm as a function of the full set of variables used to explain vertical integration. Among the main variables of interest, only the R&D intensity of the producing industry turns out to have an effect on the probability of a firm of being multiproduct, whereas the share of costs, R&D intensity of the supplying industry and the interaction of producer and supplier R&D intensities with the share of costs appear to have no influence on multiplant status. This implies that the R&D intensity of the producing industry is most likely to suffer from a bias due to endogenous selection (interestingly, this is the only coefficient for which the estimate changed substantially after the introduction of firm fixed effects).

Column 5 shows the second stage of whether a firm is vertically integrated, conditional on being a multiplant firm. The second-stage equation excludes the firm and industry characteristics.[25] Given the pattern in column 4, it is not surprising that the most remarkable change occurs in the main effect of the producer R&D intensity which is now larger (0.030) and statistically significant (standard error 0.012). The main effect of the supplier R&D becomes smaller (in absolute value), but remains highly significant.

In summary, Table 3 subjects our basic specification to a more demanding test by including a full set of firm fixed effects. These fixed effects control for various unobserved firm characteristics which may be potentially correlated with the propensity to become integrated. These specifications still show a negative effect of supplier R&D, and a positive (sometimes significant and sometimes insignificant) effect of producer R&D. Most importantly, all of these specifications show large, statistically significant and opposite-signed interaction effects between producer and supplier R&D intensities and the share of costs.

[25] We do not have any natural exclusion restrictions to impose on this model. We attempted to estimate the selection model using only functional form restrictions, however, as is commonly the case in such models, we were not able to obtain sensible estimates due to the high degree of collinearity between the variables.

4.3. An Instrumental Variables Strategy

So far, the results point to statistically significant associations between vertical integration and the technology intensity in the producing and supplying industries. However, these associations do not necessarily correspond to the causal effects of the technology intensity variables on vertical integration decisions. First, as highlighted by the above theory, vertical integration also affects investment in technology, so that there is scope for reverse causality. Second, there may be other variables that are omitted from the regressions, which have a causal effect on both technology intensity and vertical integration. This will mean that the error term is correlated with the regressors and will lead to biased estimates of the coefficients of interest. To the extent that the omitted variables are at the firm level, this is controlled for by the within-firm estimates shown above, but there may still be omitted variables at the firm-industry level affecting the estimates.

A more satisfactory approach would be to use an instrumental variable strategy, with instruments that affect technology intensity, without influencing vertical integration through other channels (i.e., they should be orthogonal to the error term, ε_{ijk}, in equations (4.1) and (4.2)). Although we do not have such perfect instruments, measures of technology intensity in the same industry in the U.S. are potential candidates. These instruments are useful in avoiding the potential reverse causality problems and in removing the effect of UK-specific omitted variables, although this procedure would *not* help with omitted industry-specific variables that are common across the U.S. and the UK. Therefore, these results should not necessarily be interpreted as causal estimates, but as estimates investigating a different source of variation in the data.

Since we do not have U.S. data on R&D intensity at the same level of disaggregation, we use the investment intensity of the corresponding supplying and producing sector in the United States (at the same 2/3-digit industry) as instrument for R&D intensity. The first-stage equations for the model in (4.1) are:

$$\begin{aligned} RD_j^P &= \pi_1^P sc_{jk} + Z'\pi_2^P + X'_{ijk}\eta^P + u_{ijk}^P \\ RD_k^S &= \pi_1^S sc_{jk} + Z'\pi_2^S + X'_{ijk}\eta^S + u_{ijk}^S, \end{aligned} \tag{4.3}$$

where the Z_j's is the vector of instruments for technology intensity in the supplying and producing industries (in other words, investment intensity in the supplying and producing industries in the U.S.), while π_2^P and π_2^S are vectors of parameters.

In column 1 of Table 4, we start with instrumental variables (IV) estimates of equation (4.1). The top panel shows the second-stage results, while the bottom panels report the first-stage coefficients from (4.3) as well as the R^2 and the p-value of the F-statistics for the significance of the instruments in the first stage. The first-stage relationships are highly significant, and show a very appealing pattern: producer investment technology intensity in the U.S. is correlated with producer R&D intensity in the UK, but not with the supplier technology intensity in the UK. The pattern is similar (i.e., the coefficients are reversed) for supplier technology intensity in the U.S.. This gives us confidence that the IV estimates are capturing a useful source of variation.

The second stage estimates in column 1 are interesting: the producer R&D intensity is positive and supplier R&D intensity is negative. Both estimates are highly significant and much larger than the OLS estimates. The larger IV estimates for the main effects of the technology intensity variables likely reflects the fact that the IV procedure is reducing the attenuation bias in the OLS estimates resulting from classical measurement error. This type of attenuation bias might be quite important here, since our measures of technology (R&D) intensity are very imperfect proxies for the importance of relationship-specific technology investments. Another possible interpretation is that, consistent with the significant interactions between technology intensity and the share of costs, which show heterogeneous effects conditional on observables, there are also heterogeneous effects conditional on unobservables. In that case, because OLS and IV have different weighting functions, it is natural that they will lead to different estimates (see Angrist and Imbens, 1995). In any case, we interpret these results as supportive of the main prediction of our theory.

For the model in (4.2), we have four first-stage equations, which are

$$\begin{aligned}
RD_j^P &= \pi_1^P sc_{jk} + Z'\pi_2^P + (Z'sc_{jk})\,\pi_3^P + X'_{ijk}\eta^P + u_{ijk}^P \\
RD_j^P sc_{jk} &= \pi_1^{PC} sc_{jk} + Z'\pi_2^{PC} + (Z'sc_{jk})\,\pi_3^{PC} + X'_{ijk}\eta^{PC} + u_{ijk}^{PC} \\
RD_j^S &= \pi_1^S sc_{jk} + Z'\pi_2^S + (Z'sc_{jk})\,\pi_3^S + X'_{ijk}\eta^S + u_{ijk}^S \\
RD_j^S sc_{jk} &= \pi_1^{SC} sc_{jk} + Z'\pi_2^{SC} + (Z'sc_{jk})\,\pi_3^S + X'_{ijk}\eta^{SC} + u_{ijk}^{SC},
\end{aligned}$$

where, for example, $RD_j^P sc_{jk}$ is the interaction between producer R&D intensity and share of cost, and $Z'sc_{jk}$ denotes the interaction between the vector of instruments and the share of cost.

In columns (2) to (4) of Table 4 we report IV estimates based on these (to save space, in the bottom panel we report only the two first-stage equations for the main effects).[26] While the instrumental variable strategy works reasonably well for the main effects, instrumenting for the interaction effects is more challenging, since instruments

[26]Given the size of the data set, all IV estimates are implemented using the "control function"

constructed by interacting the original instruments with the share of costs are highly correlated with the instruments for the main effects. In column 2, we report estimates where both main effects and interaction terms are instrumented. The main effect estimates are very similar to those in column 1, but the interaction estimates are very large and imprecisely estimated (though still statistically significant). This reflects the difficulty of simultaneously instrumenting for main effects and interaction effects that are mechanically correlated.[27] In column 3, we add our usual set of covariates to this specification. Now the multicolinearity issue becomes worse and the main effect of producer R&D intensity turns negative, while the interaction terms are even larger and more imprecisely estimated. Finally, in column 4, we add fixed effects. Interestingly, in this case, the estimates are more similar to those in column 2.

Overall, the IV estimates using the U.S. values as instruments for UK producer and supplier R&D intensity are mixed. When we only instrument the main effects, the results confirm the overall picture emerging from the OLS regressions. When we also instrument the interaction terms, there is too much multicolinearity to learn much from the estimates.

5. Robustness

In this section we consider a number of robustness checks. First, we investigate whether the use of a linear probability models versus a probit model matters for the results. Second, we report results using our alternative measure of technology intensity, which uses information on physical investment. Finally, we consider a number of other robustness checks. In the Appendix, we also check the robustness of our main results in various subsamples and investigate the role of nonlinearities.[28]

approach, which involves including the residuals from the first stages as additional regressors in the second stage (see, for example, Wooldridge, 2002, Chapter 18).

[27] As an additional experiment, we used only the control functions from the two levels equations, which essentially amounts to only instrumenting for the main effects. This yields results that are much more favorable to our hypothesis and similar to the OLS. For example, the coeffiecient estimates for the main effects and interactions of producer and supplier R&D in the equivalent specification to column 2 are, respectively, 0.509 (s.e=0.061), 1.126 (s.e=0.401), -0.164 (s.e=0.015), and -0.958 (s.e=0.352).

[28] We also investigated the stability of the basic relationship between technology intensity and vertical integration over years, and the results are stable across years (details available upon request). In addition, we estimated specifications controlling for the share of the output of the supplying industry going to the producing industry in question. When entered by itself, this variable is significant with the expected sign, but when entered together with the share of cost, it is no longer significant and typically has the opposite of the sign predicted by our model (the results are available upon request, and see also footnote 12 above).

5.1. Probit Estimates versus Linear Probability Models

Tables 2-4 use linear probability models. These have a number of attractive features, including being easier to interpret and estimate (for example, with large samples and individual fixed effects). Nevertheless, it is important to investigate whether alternative estimation strategies lead to similar results. This issue is addressed in Table 5.

Specifically, Table 5 compares estimates from the linear probability model and the probit model. Column 1 repeats column 3 from Table 2. Column 2 reports marginal effects from a probit model evaluated at the mean value of all right-hand side variables. Column 3 reports the mean, minimum and maximum values of the marginal effects (from the same probit model as in column 2). Columns 4 to 6 repeat this for the specification shown in column 4 of Table 2 and columns 7 to 9 for the specification shown in column 6 of Table 2. The marginal effects of the interaction terms are calculated using the formulae given by Ai and Norton (2003).

These results show that, on the whole, our main results are not sensitive to the choice of functional form. With probit, the main effect for producer R&D intensity is somewhat smaller, but similar and statistically significant (e.g., compare column 2 to column 1). For the interaction effects, the probit results are substantially smaller when the marginal effects are evaluated at the mean. However, more interesting is the mean of the marginal effects, reported in columns 6 and 9; here, the interaction effects are larger, though still smaller than the linear probability models. These columns also show that the range of estimates includes much larger interaction effects. Therefore, our interpretation of the somewhat smaller interaction effects in the probit estimation is that the nonlinearity in the probit specification is giving us information about a different part of the distribution of heterogeneous effects.

Overall, despite some differences in magnitudes, these results are generally supportive of the patterns shown in Tables 2-4.

5.2. Results with Investment Intensity

Our alternative measure of technology intensity uses information on physical investments instead of R&D. Even though greater physical investment need not be associated with more specific investments, we expect firms making more investments overall to also undertake greater investments in technology and relationship-specific assets. In this light, it would be reassuring if the results were similar when using the investment intensity measure of technology.

Table 6 repeats the benchmark regressions of Table 2 using our alternative measure of technology intensity. Overall, the results point to a similar pattern: investment intensity in the producing industry is positively associated with vertical integration, while investment intensity in the supplying industry is negatively associated with integration. For example, column 3 shows a coefficient of 0.030 (standard error =0.006) on producer technology intensity, while the coefficient on supplier technology intensity is -0.046 (standard error =0.004).

The exception to this pattern is for producer investment intensity when the full set of additional covariates are included. In this case producer technology intensity is no longer statistically significant. Nevertheless, there continues to be a positive, significant, and large effect of the interaction between producer investment intensity and the share of costs. Consequently, producer technology intensity has no effect on vertical integration when evaluated at the mean share of cost, but has a substantial effect when the share of cost is large (e.g. at the 90th percentile). Supplier technology intensity continues to be highly significant, both at the mean and for large shares of costs.[29]

5.3. Other Robustness Checks

The rest of the robustness checks are presented in the Appendix.

Appendix Table A.2 reports estimates from the specification including all the covariates as in column 6 of Table 2 for two subsamples for (we do not report the coefficients on the covariates to save space). In the first column, we show the results excluding the bottom quartile of firms by size, in the second column we exclude the top quartile by size. These models are useful to check whether our results are driven by the comparison of large to small firms. The results hold up in these subsamples.

In column 3, we use our alternative measure of vertical integration, $\overline{vi}_{ijk}$. The results are very similar to those using the vertical integration dummy, vi_{ijk}. In particular, producer R&D intensity has a positive and significant effect on vertical integration while supplier R&D intensity has a negative effect. The interaction effects are significant and have opposite signs as in our baseline estimates.

Finally, Appendix Table A.3 considers potential nonlinearities. It reports results with dummies for share of cost, and a producing (or supplying) industry being at the

[29] We also repeated our other robustness checks using investment intensity. The results are generally similar to those with R&D intensity. For example with firm fixed effects, producer investment intensity loses significance when we include fixed effects, but once we control for selection as in column 5 of Table 3, we recover an estimate very similar to that in Table 6. Similarly, instrumental variable estimates show a similar pattern to those with R&D intensity. These results are available upon request.

second, third or fourth quartile of the corresponding distribution (with the first quartile as the omitted group). The results show that there is generally a monotonic pattern, consistent with the linear regressions reported in previous tables, with the exception of the effect of R&D intensity in the supplying industry. Here the second quartile has the largest negative effect, while the third quartile has a small, and sometimes insignificant, sometimes positive effect. This nonlinear pattern, for which we do not have a good explanation, is the reason why the difference in vertical integration by the R&D intensity of the suppliers was not visible in the descriptive statistics in Table 1.

6. Outside Options and Competition

Finally, as discussed in Section 2, our model points out potential links between competition and vertical integration. In Table 7, we briefly investigate this relationship using the number of firms in producing and supplying industries as our main indicator of competition. The table shows that a greater number of firms in the producing industry is associated with lower vertical integration, while a greater number of firms in the supplying industry leads to more vertical integration. When we include fixed effects (columns 2 and 4), the effect of the number of firms in the producing industry is no longer significant, but the effect of the number of firms in the supplying industry remains significant and of a similar magnitude to the estimate without fixed effects.

It is also noteworthy that the coefficient on the number of firms in the supplying industry is about four times the magnitude of the coefficient for the number of firms in the producing industry. Ignoring this difference in magnitude, for which we do not have any good explanation, these results are consistent with the theory, where we showed that, as long as the relevant margin is backward integration, a greater θ, which increases the outside option of the supplier, should make vertical integration less likely. A greater number of firms in the producing industry is likely to increase the supplier's outside option, while more firms in the supplying industry should reduce it. This is the pattern we find in the data.[30]

Interestingly, if we think of an increase in overall competition as corresponding to a proportional increase in the number of producing and supplying firms, since the coefficient on the number of supplying firms is larger, our estimates suggest that there should

[30] We also experimented with Hirfindahl indices for producing and supplying industries. Although the Hirfindahl indices were sometimes significant, the results were not robust. The addition of the Hirfindahl indices did not change the effects of the number of firms in the producing and supplying industries on vertical integration, however. Also, the results are robust to using physical investments instead of R&D as a measure of technology intensity.

be an increase in vertical integration. Although this result is not our main focus, it sheds some doubt on the popular claims that greater (global or national) competition is necessarily leading to less integrated firms.

7. Summary and Conclusions

Despite a number of well-established theories and a prominent public debate on the effect of technology and technical change on the internal organization of the firm, there is little evidence on the determinants of vertical integration. This paper proposes a new approach to investigate the predictions of the PRT approach of Grossman and Hart (1986) and Hart and Moore (1990). This approach relies on comparing vertical integration patterns across pairs of industries (products). We use data from the entire population of UK manufacturing plants, and document a number of empirical regularities in the relationship between technology intensity and vertical integration.

Our results show that vertical integration in a pair of industries is less likely when the supplying industry is more technology intensive and the producing industry is less technology intensive. Moreover, both these effects are larger when inputs from the supplying industry constitute a large fraction of the total costs of the producing industry. This pattern of opposite effects of technology intensity of producing and supplying industries is consistent with the PRT approach. In addition, the direction of these effects, for example, that vertical integration is more likely when the producing industry is more technology intensive, and the other patterns we document, are consistent with the theory, provided that the relevant margin in the data is the choice between backward vertical integration and non-integration.

We report similar results controlling for firm fixed effects and instrumenting UK technology intensity measures with U.S. measures. We also show the robustness of our results to a variety of other specifications.

Finally, we find that vertical integration is more likely when the average number of producing firms is greater relative to the average number of supplying firms, which is also consistent with the theoretical predictions we derived from a simple incomplete contracts model (and not entirely consistent with the claims made in the popular press about the effect of competition on the structure of firms).

The results in this paper provide a number of empirical patterns that may be useful for theories of vertical integration to confront. Although our empirical investigation is motivated by a specific theoretical approach, the PRT, the empirical patterns we document should be of more general interest and may be consistent with various alternative theories. Nevertheless, the current versions of the most popular alternative approaches are not easily reconciled with the findings. For example, the emphasis of Williamson's TCE approach that vertical integration circumvents holdup problems would be consistent with the positive association between vertical integration and producer technology intensity, but not with the negative effect of supplier technology intensity. Theories based on supply assurance (e.g., Green, 1986, Bolton and Whinston, 1993) could also account for part of the results if more technology-intensive firms require more assurance. But these theories do not provide an explanation for why greater technology intensity of suppliers is associated with less vertical integration. Theories based on securing intellectual property rights (e.g., Aghion and Tirole, 1994b) could account for both main effects of producer and supplier technology intensity, for example, because creating a vertically-integrated structure may provide better protection of intellectual property rights. These theories would not explain why these effects become stronger when the share of costs is high. A theoretical investigation of various alternative explanations for these patterns, as well as further empirical analysis of the robustness of these results with data from other countries, appear to be interesting areas for future research.

8. Appendix A: Proof of Proposition 1

Substituting the optimal investments given by (2.6), (2.7) and (2.8) into (2.5), social surplus under the three organizational forms is obtained as:

$$\begin{aligned} S^{VIB} &= 1+\frac{1}{2}p^2+\frac{\phi}{2}\lambda\left(1-\frac{\lambda}{4}\right)s^2, \\ S^{NI} &= 1+\left(1-\frac{2-\phi}{4}\right)\left(1-\frac{\phi}{2}\right)p^2+\frac{\phi}{2}\left(1+\theta\right)\left(1-\frac{1+\theta}{4}\right)s^2, \\ S^{VIF} &= 1+\frac{1}{2}\lambda'\left(1-\frac{\lambda'}{4}\right)p^2+\frac{\phi}{2}s^2. \end{aligned} \tag{A1}$$

Let

$$\Delta^B \equiv S^{VIB}-S^{NI} = \phi^2p^2/8-\left(3-\theta-\lambda\right)\left(1+\theta-\lambda\right)\phi s^2/8.$$

It is straightforward to verify that Δ^B is increasing in p, decreasing in s, and $\Delta^B=0$ if and only if

$$\frac{p}{s}=\sqrt{\left(3-\theta-\lambda\right)\left(1+\theta-\lambda\right)/\phi}\equiv\overline{r}>0.$$

When $p/s>\overline{r}$, backward integration is preferred to non-integration. When $p/s<\overline{r}$, it is dominated by non-integration. Differentiation establishes that $\frac{\partial\overline{r}}{\partial\phi}<0$ and $\frac{\partial\overline{r}}{\partial\theta}>0$.

Similarly, let

$$\Delta^F \equiv S^{VIF}-S^{NI} = -\left(\left(2-\lambda'\right)^2-\phi^2\right)p^2/8+\phi\left(1-\theta\right)^2s^2/8.$$

Δ^F is decreasing in p and increasing in s. $\Delta^F=0$ if and only if

$$\frac{p}{s}=\sqrt{\frac{\phi\left(1-\theta\right)^2}{\left(2-\lambda'\right)^2-\phi^2}}\equiv\underline{r}>0.$$

When $p/s<\underline{r}$, forward integration is preferred to non-integration. In contrast, when $p/s>\underline{r}$, non-integration is preferred. Again, differentiation establishes that $\frac{\partial\underline{r}}{\partial\phi}>0$ and $\frac{\partial\underline{r}}{\partial\theta}<0$.

First, suppose that $\underline{r}<\overline{r}$. Then, the analysis so far establishes that the equilibrium organizational form is given by

$$z^*=\begin{cases} VIB & \text{if } \frac{p}{s}\geq\overline{r} \\ NI & \text{if } \frac{p}{s}\in\left(\underline{r},\overline{r}\right) \\ VIF & \text{if } \frac{p}{s}\leq\underline{r} \end{cases}.$$

The set of parameters such that $\underline{r}<\overline{r}$ is non-empty. For instance, as $\theta\to1$ we have that $\underline{r}\to0$, whereas $\overline{r}\to\left(2-\lambda\right)/\sqrt{\phi}>0$.

Next, suppose that $\underline{r} \geq \overline{r}$. Then NI is always dominated by either VIF and VIB. Let

$$\Delta^{BF} \equiv S^{VIB} - S^{VIF} = (2 - \lambda')^2\, p^2/8 - \phi\, (2 - \lambda)^2\, s^2/8.$$

Δ^{BF} is increasing in p and decreasing in s, and also $\Delta^{BF} = 0$ if and only if

$$\frac{p}{s} > \frac{2 - \lambda}{2 - \lambda'} \sqrt{\phi} \equiv \widehat{r}.$$

When $p/s > \widehat{r}$, backward integration is preferred to forward integration, and when $p/s < \widehat{r}$, forward integration is preferred. Differentiation establishes that $\frac{\partial \widehat{r}}{\partial \phi} > 0$ and$\frac{\partial \widehat{r}}{\partial \theta} = 0$. This completes the proof of the proposition.

9. Appendix B: Data Sources and Construction

Our main source of data is the plant level production data underlying the UK Census of Production (ARD). This is collected by the UK Office of National Statistics (ONS) and firms have a legal obligation to reply. We use the data on all manufacturing plants from 1996-2001, along with information from the Input-Output Domestic Use Table for 1995, to measure vertical integration and other firm characteristics. We use data from the ARD from 1992-1995 to measure a number of other industry characteristics and data from the annual Business Enterprise Researcher and Development (BERD) survey from 1994-1995 to measure R&D expenditure at the industry level. US variables are measured using the US Census data at the 4-digit level (available on the NBER web site). The UK and US data are matched based on a mapping of UK SIC92 to US SIC87 and then aggregated up to input-output industry level. See Acemoglu, Aghion, Griffith and Zilibotti (2004) for further discussion of the data.

9.1. The Plant-Level Production Data

The ARD contains information on all production activity located in the UK. The basic unit for which information on inputs and output is reported is a reporting unit. A reporting unit can be a single plant or a group of plants owned by the same firm operating in the same 4-digit industry. Information on the location and number of employees is available on all plants (called local units) within each reporting unit. There are over 150,000 reporting units in manufacturing industries with non-zero employment in the ARD each year 1996-2000. Detailed data is collected from a random stratified sample.[31] Data on value-added and costs for non-sampled reporting units are imputed.

Single plant firms are identified as those reporting units which represent only one plant and which have no sibling, parent or child plants. Single plants with fewer than 20 employees are dropped from the analysis, resulting in between 100,000 - 130,000 reporting units being dropped per year. In addition 1,000 - 2,000 reporting units per year, which are owned by foreign firms are dropped, as we do not observe their foreign activities.

Plants in the ARD in these years are classified by their major product according to the 1992 revision of the 4-digit standard industrial classification (SIC code). Input-output (IO) tables are reported at the 2/3-digit level. Where more than one reporting unit exists

[31] The sampling probabilities vary over time, with industry and with reporting unit size. Reporting units with 100 or more employees are always sampled. Below that the sampling probabilities range from 1 in 5 to 1 in 2.

within an IO industry these are aggregated so that there is only one observation per firm in each IO industry. The total number of firms used (after dropping the small and single and foreign owned firms and averaging over years) is 46,392. We measure firm age in each producing industry as the number of years since the first plant in that industry was established. We measure firm size in each industry by the number of employees it has in that industry. The average number of firms in an industry is measured from the ARD. Table 1 (in the main text) shows means for these variables.

9.2. The Input-Output Table

We use the Input-Output table for 1995.[32] The Input-Output table contains information on 77 manufacturing industries (supplying and producing). There are 5,929 pairs of producing-supplying industries, for which 3,840 the input-output table indicates positive trade flows. For each industry pair we calculate the proportion of total costs (including intermediate, labour and capital) of producing j that are made up of input k, denoted w_{jk}. In 2,766, or just under half of industry pairs, at least one firm is vertically integrated to some extent. Table A.1 contains descriptive statistics on the share of output from each supplying industry that is sold for intermediate consumption, to all industries and to manufacturing industries, and shows the largest purchasing industry along with the share of sales this purchaser represents (which ranges from a half of a percent to over fifty percent and averages 3.7%) and the share of the purchaser's total costs this input represents (which ranges from zero to 37% and averages 2.7%).

9.3. Technology Indicators

Our measures of technology intensity are all at the industry level. R&D intensity is measured using the micro data underlying the annual Business Enterprise Research and Development (BERD) matched to the ARD. The micro data is aggregated to the industry level using the product code for which the R&D was targeted. This is scaled by total value-added in firms producing in the industry (including both R&D and non-R&D doing firms).

The ratio of physical investment (capital expenditure on machinery, buildings, land and vehicles) to value-added is constructed in a similar manner from the ARD data at the industry level and averaged over the years 1992-1995.

[32] This is available at www.statistics.gov.uk.

References

Acemoglu, Daron (1996) “A Microfoundation For Social Increasing Returns in Human Capital Accumulation”, *Quarterly Journal of Economics*, 111, 779-804.

Acemoglu, Aghion, Griffith and Zilibotti (2004) “Vertical integration and technology: theory and evidence” IFS Working Papers, W04/34

Acemoglu, Daron, Aghion, Philippe, and Fabrizio Zilibotti (2004): “Vertical Integration and Distance to Frontier”, *Journal of the European Economic Association, Papers and Proceedings.*

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen and Fabrizio Zilibotti (2005) “Technology, Information, and the Decentralization of the Firm.” Mimeo IIES, LSE, CREST, Harvard and MIT.

Acemoglu, Daron and Steve Pischke (1999) “The Structure of Wages and Investment in General Training” *Journal of Political Economy*, 107

Aghion, Philippe, and Jean Tirole (1994a) “On the Management of Innovation”, *Quarterly Journal of Economics,* 102, 1185-1207.

Aghion, Philippe, and Jean Tirole (1994b) “Opening the Black Box of Innovation”, *European Economic Review, Papers and Proceedings,* 38, 701-710.

Aghion, Philippe, and Jean Tirole (1997): “Formal and Real Authority in Organizations”, *Journal of Political Economy,* 105, 129.

Ai, Chunrong and Edward Norton (2003) “Interaction terms in logit and probit models” *Economics Letters*, 80 (2003), 123-129

Angrist, Joshua and Guido Imbens (1995) “Two-Stage Least Squares Estimates of Average Causal Effects in Models with Variable Treatment Intensity”, *Journal of the American Statistical Association,* 90, 431-442.

Antras, Pol (2003) “Firms, Contract and Trade Structure”, *Quarterly Journal of Economics,* 118, 1375-1418.

Athey, Susan and Armin Schmutzler (1995) “Product and Process Flexibility in an Innovative Environment” *The RAND Journal of Economics*, 26, 557-574.

Baker, George, and Thomas Hubbard (2000) “Contractibility and Asset Ownership: On-Board Computers and Governance in US Trucking”, NBER Working Paper 7634.

Baker, George, and Thomas Hubbard (2002) “Make Versus Buy in Trucking: Asset Ownership, Job Design and Information”, mimeo Harvard.

Barnes, Matthew and Ralf Martin (2002) “Business Data Linking: An Introduction.” *Economic Trends*, 581, April, pp.34-41.

Binmore, Kenneth, Rubinstein, Ariel, and Asher Wolinski (1986),”The Nash Bargaining Solution in Economic Modelling”, *Rand Journal of Economics,* 17, 176-188.

Bolton, Patrick and Michael Whinston (1993): ”Incomplete Contracts, Vertical Integration and Supply Assurance", *Review of Economic Studies,* 60, 121-148.

Bresnahan Timothy, Brynjolfsson, Eric, and Lorin Hitt (1999) "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence", NBER Working Papers no. 7136.

Caves, Richard, and Ralph Bradburd (1988): "The Empirical Determinants of Vertical Integration", *Journal of Economic Behavior and Organizations*, 9, 265-279.

Chipty, Tasneem (2001): "Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry" *American Economic Review,* 91, 428-53.

Davies, Steve, and Clair Morris (1995) "A New Index of Vertical Integration: Some Estimates for UK Manufacturing" *International Journal of Industrial Organization* 13, 151-177.

Fan, Joseph, and Larry Lang (2000) " The Measurement of Relatedness: An Application to Corporate Diversification", *Journal of Business*, 73: 4, 629-660.

Feenstra Robert (1998) "Integration of Trade and Disintegration of Production", *Journal of Economic Perspective,* 12, 31-50.

Feenstra Robert and Gordon Hanson (1996) "Globalization, Outsourcing and Wage Inequality", *American Economic Review*, 86, 24-245.

Feenstra Robert and Gordon Hanson (1999) "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the US, 1979-1990Globalization, Outsourcing and Wage Inequality", *Quarterly Journal of Economics*, 114, 907-940.

Green, Jerry (1986): "Vertical Integration and Assurance of Markets", in *New Developments in the Analysis of Market Structure,* J.E Stiglitz and G.F Mathewson (eds), MIT Press.

Griffith, Rachel (1999) "Using the ARD Establishment Level Data: An Application to Estimating Production Functions" Economic Journal, 109, F416-F442.

Grossman, Sandford and Oliver Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy,* 94, 691-719.

Grossman, Gene and Elhanan Helpman (2002): "Integration versus Outsourcing in Industry Equilibrium", *Quarterly Journal of Economics,* 117, 85-120.

Grossman, Gene and Elhanan Helpman (2003): "Outsourcing in the Global Economy", *Review of Economic Studies,* forthcoming.

Hart, Oliver (1995): *Firms, Contracts, and Financial Structure,* Oxford University Press.

Hart, Oliver and John Moore (1990) "Property Rights and the Nature of the Firm", *Journal of Political Economy,* 98, 1119-1158.

Helper, Susan (1991) "How Much Has Really Changed Between US Automakers and Their Suppliers?" *Sloan Management Review*, 32, 15-28.

Hortacu, Ali and Chad Syverson (2005) “Vertical Integration, Foreclosure, Productivity and Prices” University of Chicago, mimeo.

Holmstrom, Bengt and Jean Tirole (1989): ”The Theory of the Firm,”Chapter 2 in *The Handbook of Industrial Economics,* Vol I, Schmalensee, Richard and Robert Willig, eds, Elsevier.

Joskow, Paul (1987) “Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets”, *American Economic Review,* 77, 168-185.

Joskow, Paul (2003) “Vertical Integration” forthcoming in *Handbook of New Institutional Economics.*

Klein, Benjamin, (1998): “Vertical Integration As Organized Ownership: the Fisher Body-General Motors Relationship Revisited” *Journal of Law, Economics and Organization,* 4, 199-213.

Klein, Benjamin, Crawford, Robert, and Armen Alchian (1978): ”Vertical Integration, Appropriable Rents, and the Competitive Contracting Process,” *Journal of Law and Economics,* 21, 297-326

Legros, Patrick and Andrew Newman (2003): ”Competing for Ownership”, CEPR Working Paper #2573

Lerner, Josh, and Robert Merges (1998) “The Control of Technology Alliances: An Empirical Analysis of the Biotechnology Industry”, *Journal of Industrial Economics,* 2, 125-156.

Masten, Scott (1984) “The Organizational Production: Evidence from the Aerospace Industry” *Journal of Law and Economics,* 27, 403-17.

McLaren, John (2000): ” Globalization and Vertical Structure,” *American Economic Review,* 90, 1239-1254.

Marin, Dalia, and Thierry Verdier (2002) “Globalization and the Empowerement of Human Capital”, Mimeo, University of Munich and Delta.

Marin, Dalia, and Thierry Verdier (2003) “Globalization and the New Enterprise”, *Journal of the European Economic Association, Papers and Proceedings.*

Milgrom, Paul, and John Roberts (1990) “The Economics of Modern Manufacturing: Technology, Strategy, and Organization” *American Economic Review*, 80, 511-528.

Monteverde, Kirk and David Teece (1982): ”Supplier Switching Costs and Vertical Integration in the Automobile Industry”, *Bell Journal of Economics*, 13, pp 206-213.

Ohanian, Nancy Kane (1994): “Vertical Integration in the U.S. Pulp and Paper Industry, 1900-1940”, *Review of Economics and Statistics*, 76, 202-07.

Stuckey, James (1983): *Vertical Integration and Joint Ventures in the Aluminium Industry,* Harvard University Press.

Whinston, Michael (2001): ”Assessing the Property Rights and Transaction-Cost Theories of Firm Scope”, *American Economic Review*, *Papers and Proceedings*, 91, pp 184-

188.

Williamson, Oliver (1975): *Markets and Hierarchies: Analysis and Antitrust Implications,* Free Press, New-York.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism,* Free Press, New-York.

Woodruff, Christopher (2002) "Non-contractible Investments and Vertical integration in the Mexican Footwear Industry", *International Journal of Industrial Organization,* 20, 1197-1224.

Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT University Press Cambridge Massachusetts.

Table 1: Descriptive statistics

Variable	Mean (s.d.)	Producer R&D		Supplier R&D	
		low	high	low	high
Mean $\overline{vi}_{ijk}$	0.008	0.007	0.010	0.008	0.009
	(0.087)	(0.078)	(0.096)	(0.084)	(0.089)
Mean of vi_{ijk}	0.009	0.009	0.013	0.010	0.011
	(0.091)	(0.093)	(0.114)	(0.101)	(0.104)
Firm age	10	10	10	10	10
	(7)	(7)	(7)	(7)	(7)
Firm employment	111	99	125	109	112
	(455)	(346)	(559)	(444)	(465)
Share of producer costs (jk)	0.010	0.010	0.010	0.012	0.009
	(0.034)	(0.038)	(0.029)	(0.040)	(0.028)
Producing industry					
R&D over value-added	0.027	0.004	0.055	0.026	0.028
	(0.055)	(0.002)	(0.072)	(0.055)	(0.055)
Investment over value-added	0.101	0.095	0.109	0.102	0.101
	(0.041)	(0.031)	(0.049)	(0.040)	(0.041)
Mean number of firms in industry	5757	8267	2763	5755	5759
	(6585)	(7978)	(1635)	(6525)	(6636)
Supplying industry					
R&D over value-added	0.046	0.044	0.050	0.005	0.082
	(0.107)	(0.103)	(0.113)	(0.003)	(0.137)
Investment over value-added	0.122	0.123	0.122	0.106	0.136
	(0.057)	(0.057)	(0.057)	(0.038)	(0.067)
Mean number of firms in industry	2316	2320	2309	3347	1433
	(3730)	(3727)	(3733)	(5065)	(1471)

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: $\overline{vi}_{ijk}$ is a continuous measure of the share of the producers demand can be met by its own supply. vi_{ijk} is a dummy for whether a firm owns plants in both producing and supplying industries. Share of producer costs (jk) is the share of producers in industry j total costs (including labour and capital) that is on input k (from the Input-Output Table). The sample contains 2,973,008 observations on 46,392 firms. Numbers reported are means (standard deviations). The first column reports on the whole sample. Subsequent columns split the sample by median producer R&D and supplier R&D intensities.

Table 2: Main results – R&D intensity

dependent variable: vi_{ijk}	(1)	(2)	(3)	(4)	(5)	(6)
Share of costs (jk)			0.204 (0.029)	0.187 (0.028)	0.187 (0.028)	0.182 (0.027)
R&D intensity, producing (j)	0.038 (0.006)		0.040 (0.005)	0.044 (0.005)	0.037 (0.005)	0.030 (0.005)
x Share of costs				1.112 (0.402)	1.104 (0.397)	1.067 (0.374)
R&D intensity, supplying (k)		-0.010 (0.002)	-0.007 (0.001)	-0.013 (0.003)	-0.013 (0.003)	-0.007 (0.003)
x Share of costs				-0.909 (0.353)	-0.914 (0.351)	-0.871 (0.324)
ln Firm size (ij)					0.0053 (0.0002)	0.0052 (0.0002)
ln Firm age (ij)					0.0010 (0.0001)	0.0009 (0.0001)
ln Average firm size, producing (j)						0.0011 (0.0005)
ln Average firm size, supplying (k)						-0.0036 (0.0004)
ln Average firm age, producing (j)						0.012 (0.003)
ln Average firm age, supplying (k)						0.004 (0.002)

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.
Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. There are 2,973,008 observations at the firm-industry pair level. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. The RHS variables firm size and firm age are measured at the firm-industry pair level in 1996 (age is equal to 1 if the firm enters that industry pair after 1996). R&D intensity is R&D carried out in the UK divided by value-added produced in the UK, taken from plant level R&D data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. Average firm size and age are calculated from the firm-industry pair data and average over the years 1996-2001. In regression with interactions, all main effects are evaluated at sample means.

Table 3: Within firm variation

	(1)	(2)	(3)	(4)	(5)
dependent variable:	vi_{ijk}	vi_{ijk}	vi_{ijk}	= 1 if multi product	vi_{ijk}
Share of costs (jk)	0.434	0.434	0.414	0.025	0.475
	(0.054)	(0.053)	(0.050)	(0.094)	(0.058)
R&D intensity, producing (j)	0.073	0.024	0.014	0.333	0.030
	(0.012)	(0.016)	(0.017)	(0.025)	(0.012)
x Share of costs	2.725	2.786	2.607	-0.058	2.741
	(0.917)	(0.880)	(0.816)	(1.048)	(0.915)
R&D intensity, supplying (k)	-0.041	-0.042	-0.021	0.019	-0.016
	(0.009)	(0.009)	(0.008)	(0.024)	(0.005)
x Share of costs	-2.434	-2.543	-2.303	0.914	-2.482
	(0.897)	(0.865)	(0.786)	(0.990)	(0.891)
ln Firm size (ij)			0.0024 (0.0004)	0.0044 (0.0008)	
ln Firm age (ij)			00004 (0.0005)	0.0076 (0.0007)	
ln Average firm size, producing (j)			0.003 (0.002)	0.014 (0.005)	
ln Average firm size, supplying (k)			-0.012 (0.001)	-0.0005 (0.0042)	
ln Average firm age, producing (j)			-0.006 (0.010)	0.352 (0.021)	
ln Average firm age, supplying (k)			0.017 (0.005)	-0.0019 (0.0224)	
Observations	891,942	891,942	891,942	2,973,008	891,942
Fixed effects	no	6,713 firms	6,713 firms	no	no

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: In columns (1)-(3) and (5) the dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. In column (4) the dependent variable is whether or not the observations is for a multi-plant firm. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. R&D intensity is R&D carried out in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. In regression with interactions, all main effects are evaluated at sample means.

Table 4: Instrumental variables

dependent variable: vi_{ijk}	(1)	(2)	(3)	(4)
Share of costs (jk)	0. 183	0.054	-0.085	0.210
	(0.028)	(0.069)	(0.081)	(0.134)
R&D intensity,	0. 492	0.627	-1.025	0.418
producing (j)	(0.061)	(0.077)	(0.239)	(0.246)
x Share of costs		9.832	29.763	22.288
		(4.207)	(6.773)	(7.977)
R&D intensity,	-0.158	-0.202	-0.497	-1.158
supplying (k)	(0.014)	(0.031)	(0.089)	(0.198)
x Share of costs		-7.519	-16.872	-8.561
		(3.733)	(4.884)	(6.042)
First stage producing industry R&D intensity				
US producing industry	0.172	0.177	-0.059	-0.158
investment intensity	(0.029)	(0.029)	(0.027)	(0.017)
x Share of costs		0.786	0.506	0.333
		(0.814)	(0.580)	(0.429)
US supplying industry	0.011	0.009	-0.007	-0.007
investment intensity	(0.022)	(0.022)	(0.021)	(0.013)
x Share of costs		-1.121	-0.576	-0.363
		(0.515)	(0.365)	(0.261)
F-stat P-value	0.000	0.000	0.091	0.000
R^2	0.007	0.007	0.175	0.487
First stage supplying industry R&D intensity				
US producing industry	-0.028	-0.010	-0.033	-0.040
investment intensity	(0.082)	(0.081)	(0.084)	(0.083)
x Share of costs		4.272	3.812	4.207
		(1.436)	(1.034)	(1.521)
US supplying industry	0.366	0.356	0.093	0.063
investment intensity	(0.060)	(0.60)	(0.056)	(0.056)
x Share of costs		-3.635	-3.268	-3.254
		(1.136)	(1.046)	(1.188)
F-stat (P-value)	0.000	0.000	0.015	0.029
R^2	0.024	0.236	0.069	0.065
Observations	2,973,008	2,973,008	2,973,008	891,942
Fixed effects	no	no	no	yes
covariates	no	no	yes	yes

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. Covariates included in columns (3) and (4) are: producing firm size, age, mean firm size and mean firm age in producing and supplying industries. R&D intensity is R&D and investment carried out in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. In regression with interactions, all main effects evaluated at sample means. Instruments are investment intensity in the same industry in the US.

Table 5: LPM and probit

dependent variable: vi_{ijk}	(1) LPM	(2) PROBIT	(3) PROBIT	(4) LPM	(5) PROBIT	(6) PROBIT	(7) LPM	(8) PROBIT	(9) PROBIT
		marginal effects[a]	mean of marginal effects[b]		marginal effects[a]	mean of marginal effects[b]		marginal effects[a]	mean of marginal effects[b]
Share of costs (jk)	0.204	0.090	0.094	0.187	0.088	0.092	0.182	0.072	0.087
	(0.029)	(0.007)	min 0.044	(0.028)	(0.007)	min 0.014	(0.027)	(0.006)	min 0.002
			max 1.087			max 1.059			max 1.306
R&D intensity, producing (j)	0.040	0.027	0.029	0.044	0.024	0.026	0.030	0.010	0.013
	(0.005)	(0.002)	min 0.014	(0.005)	(0.003)	min 0.004	(0.005)	(0.002)	min 0.0003
			max 0.331			max 0.293			max 0.188
x Share of costs				1.112	0.293	0.572	1.067	0.171	0.337
				(0.402)	(0.126)	min 0.146	(0.374)	(0.080)	min 0.007
						max 14.104			max 8.230
R&D intensity, supplying (k)	-0.007	-0.009	-0.010	-0.013	-0.008	-0.009	-0.007	-0.001	-0.001
	(0.001)	(0.002)	min -0.114	(0.003)	(0.002)	min -0.101	(0.003)	(0.002)	min -0.011
			max -0.005			max -0.001			max -0.00001
x Share of costs				-0.909	-0.271	-0.384	-0.871	-0.140	-0.199
				(0.353)	(0.149)	min -7.189	(0.324)	(0.099)	min -4.298
						max 0.027			max -0.004
Covariates	no	no	no	no	no	no	yes	yes	yes

[a] Evaluated at mean value of explanatory variables [b] Calculated for each individual observation using Ai and Norton (2004) formula.

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. Covariates included in columns (7)-(9) are: producing firm size, age, mean firm size and mean firm age in producing and supplying industries. R&D intensity is R&D carried out in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level.

Table 6: Alternative technology measure – Investment intensity

dependent variable: vi_{ijk}	(1)	(2)	(3)	(4)	(5)	(6)
Share of costs (jk)			0.203 (0.028)	0.191 (0.023)	0.191 (0.023)	0.187 (0.024)
Investment intensity, producing (j)	0.027 (0.007)		0.030 (0.006)	0.038 (0.006)	0.021 (0.006)	-0.002 (0.008)
x Share of costs				1.488 (0.456)	1.471 (0.460)	1.402 (0.453)
Investment intensity, supplying (k)		-0.050 (0.004)	-0.046 (0.004)	-0.055 (0.005)	-0.055 (0.005)	-0.041 (0.005)
x Share of costs				-1.681 (0.490)	-1.666 (0.487)	-1.562 (0.489)
ln Firm size (ij)					0.0054 (0.0002)	0.0052 (0.0002)
ln Firm age (ij)					0.0009 (0.0001)	0.0008 (0.0001)
ln Average firm size, producing (j)						0.0023 (0.0006)
ln Average firm size, supplying (k)						-0.0024 (0.0004)
ln Average firm age, producing (j)						0.011 (0.003)
ln Average firm age, supplying (k)						0.004 (0.002)

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. There are 2,973,008 observations at the firm-industry pair level. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. Investment intensity is investment carried out in the UK divided by value-added produced in the UK, taken from plant level investment data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. Firm size and firm age are measured at the firm-industry pair level in 1996 (age is equal to 1 if the firm enters that industry pair after 1996). Average firm size and age are calculated from the firm-industry pair data and average over the years 1996-2001. In regression with interactions, all main effects are evaluated at sample means.

Table 7: Outside option

dependent variable: vi_{ijk}	(1)	(2)	(3)	(4)
	Technology measure = R&D intensity		Technology measure = Investment intensity	
Share of costs (jk)	0.154 (0.027)	0.328 (0.042)	0.161 (0.024)	0.359 (0.039)
technology, producing (j)	0.029 (0.005)	0.010 (0.016)	-0.002 (0.007)	-028 (0.027)
x Share of costs	0.863 (0.338)	1.571 (0.628)	1.352 (0.440)	0.358 (1.330)
technology, supplying (k)	-0.004 (0.003)	-0.005 (0.007)	-0.031 (0.005)	-0.073 (0.014)
x Share of costs	-0.762 (0.287)	-1.313 (0.595)	-1.282 (0.475)	0.400 (1.129)
ln Firm size (ij)	0.0052 (0.0002)	0.0024 (0.0004)	0.0052 (0.0002)	0.0023 (0.0004)
ln Firm age (ij)	0.0009 (0.0001)	0.0004 (0.0005)	0.0009 (0.0001)	0.0003 (0.0005)
ln Average firm size, producing (j)	-0.0012 (0.0006)	0.016 (0.002)	-0.0001 (0.0006)	0.003 (0.002)
ln Average firm size, supplying (k)	-0.0022 (0.0004)	-0.0018 (0.0010)	0.0029 (0.0005)	0.0008 (0.0011)
ln Average firm age, producing (j)	0.010 (0.002)	-0.004 (0.009)	0.009 (0.002)	-0.004 (0.009)
ln Average firm age, supplying (k)	0.017 (0.002)	0.021 (0.005)	0.017 (0.002)	0.023 (0.005)
ln number of firms, producing (j)	-0.0018 (0.0003)	-0.0004 (0.0009)	-0.0019 (0.0003)	-0.0004 (-0.0009)
ln number of firms, supplying (k)	0.0055 (0.0003)	0.019 (0.002)	0.0053 (0.0003)	0.019 (0.002)
Observations	2,973,008	891,942	2,973,008	891,942
Fixed effects	no	yes	no	yes

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. Standard errors in parentheses are clustered at the level of 3,840 industry pairs. R&D and investment intensity are R&D and investment in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. Firm size and firm age are measured at the firm-industry pair level in 1996 (age is equal to 1 if the firm enters that industry pair after 1996). Average firm size and age are averages over the years 1996-2001. In regression with interactions, all main effects are evaluated at sample mean.

Table A.1: Summary of input-output table statistics

Supplying industry	% sales for intermediate consumption: all industries	% sales for intermediate consumption: manufacturing	Largest purchasing industry	% of total supplying industry sales going to this purchaser	% of total purchasing industry purchases coming from this supplier
8 Meat processing	0.441	0.250	29 Leather goods	0.019	0.502
9 Fish and fruit processing	0.461	0.195	9 Fish and fruit processing	0.085	0.141
10 Oils and fats	0.694	0.534	10 Oils and fats processing	0.161	0.208
11 Dairy products	0.403	0.183	11 Dairy products	0.109	0.143
12 Grain milling and starch	0.666	0.580	14 Bread biscuits etc	0.192	0.191
13 Animal feed	0.751	0.020	8 Meat processing	0.017	0.009
14 Bread biscuits etc	0.433	0.009	14 Bread biscuits etc	0.001	0.002
15 Sugar	0.817	0.639	16 Confectionery	0.217	0.153
16 Confectionery	0.380	0.160	16 Confectionery	0.097	0.176
17 Other food products	0.332	0.137	17 Other food products	0.043	0.075
18 Alcoholic beverages	0.100	0.070	18 Alcoholic beverages	0.065	0.127
19 Soft drinks and mineral waters	0.192	0.003	19 Soft drinks & mineral waters	0.003	0.005
20 Tobacco products	0.001	0.001	20 Tobacco products	0.001	0.001
21 Textile fibres	0.681	0.646	27 Knitted goods	0.189	0.269
22 Textile weaving	0.274	0.261	28 Wearing apparel & fur products	0.158	0.092
23 Textile finishing	0.976	0.415	23 Textile finishing	0.053	0.129
24 Made-up textiles	0.193	0.052	24 Made-up textiles	0.010	0.026
25 Carpets and rugs	0.356	0.115	25 Carpets and rugs	0.004	0.009
26 Other textiles	0.513	0.427	28 Wearing apparel & fur products	0.203	0.109
27 Knitted goods	0.030	0.019	28 Wearing apparel & fur products	0.013	0.013
28 Wearing apparel and fur products	0.099	0.018	28 Wearing apparel & fur products	0.017	0.048
29 Leather goods	0.312	0.078	30 Footwear	0.042	0.076
30 Footwear	0.219	0.065	30 Footwear	0.064	0.177
31 Wood and wood products	0.894	0.519	31 Wood and wood products	0.203	0.395
32 Pulp paper and paperboard	0.672	0.559	33 Paper and paperboard products	0.243	0.285
33 Paper and paperboard products	0.885	0.520	33 Paper and paperboard products	0.055	0.133
34 Printing and publishing	0.613	0.211	34 Printing and publishing	0.132	0.334
35 Coke ovens refined petroleum & nuclear	0.450	0.108	38 Organic chemicals	0.013	0.093
36 Industrial gases and dyes	0.535	0.459	19 Soft drinks & mineral waters	0.089	0.120

Supplying industry	% sales for intermediate consumption		Largest purchasing industry	% of total supplying industry sales going to this purchaser	% of total purchasing industry purchases coming from this supplier
	all industries	manufacturing			
37 Inorganic chemicals	0.676	0.603	37 Inorganic chemicals	0.043	0.111
38 Organic chemicals	0.056	0.052	41 Pesticides	0.005	0.051
39 Fertilisers	0.839	0.154	39 Fertilisers	0.146	0.273
40 Plastics & Synthetic resins etc	0.599	0.557	48 Plastic products	0.249	0.209
41 Pesticides	0.508	0.006	41 Pesticides	0.001	0.005
42 Paints varnishes printing ink etc	0.692	0.477	42 Paints varnishes printing ink etc	0.054	0.107
43 Pharmaceuticals	0.380	0.125	43 Pharmaceuticals	0.094	0.206
44 Soap and toilet preparations	0.229	0.105	44 Soap and toilet preparations	0.085	0.146
45 Other Chemical products	0.185	0.116	45 Other Chemical products	0.033	0.085
46 Man-made fibres	0.273	0.259	21 Textile fibres	0.058	0.100
47 Rubber products	0.552	0.243	47 Rubber products	0.041	0.102
48 Plastic products	0.759	0.405	48 Plastic products	0.081	0.179
49 Glass and glass products	0.775	0.549	49 Glass and glass products	0.133	0.276
50 Ceramic goods	0.402	0.125	50 Ceramic goods	0.046	0.113
51 Structural clay products	0.723	0.003	51 Structural clay products	0.001	0.004
52 Cement lime and plaster	0.882	0.336	53 Articles of concrete stone etc	0.286	0.136
53 Articles of concrete stone etc	0.851	0.024	53 Articles of concrete stone etc	0.022	0.044
54 Iron and steel	0.596	0.561	54 Iron and steel	0.138	0.253
55 Non-ferrous metals	0.658	0.611	55 Non-ferrous metals	0.218	0.450
56 Metal castings	0.973	0.790	62 Mechanical power equipment	0.175	0.107
57 Structural metal products	0.472	0.115	57 Structural metal products	0.025	0.050
58 Metal boilers and radiators	0.348	0.158	58 Metal boilers & radiators	0.084	0.174
59 Metal forging pressing etc	0.964	0.864	80 Aircraft and spacecraft	0.077	0.212
60 Cutlery tools etc	0.545	0.453	60 Cutlery tools etc	0.024	0.063
61 Other metal products	0.690	0.563	71 Insulated wire and cable	0.014	0.114
62 Mechanical power equipment	0.285	0.210	62 Mechanical power equipment	0.055	0.125
63 General purpose machinery	0.202	0.110	63 General purpose machinery	0.020	0.042
64 Agricultural machinery	0.132	0.009	64 Agricultural machinery	0.006	0.015
65 Machine tools	0.181	0.133	65 Machine tools	0.011	0.023
66 Special purpose machinery	0.293	0.246	66 Special purpose machinery	0.046	0.102
67 Weapons and ammunition	0.554	0.170	67 Weapons and ammunition	0.154	0.336

Supplying industry	% sales for intermediate consumption		Largest purchasing industry	% of total supplying industry sales going to this purchaser	% of total purchasing industry purchases coming from this supplier
	all industries	manufacturing			
68 Domestic appliances nec	0.187	0.019	68 Domestic appliances nec	0.007	0.014
69 Office machinery & computers	0.060	0.049	69 Office machinery & computers	0.035	0.094
70 Electric motors and generators etc	0.390	0.301	70 Electric motors and generators etc	0.072	0.165
71 Insulated wire and cable	0.550	0.304	71 Insulated wire and cable	0.031	0.073
72 Electrical equipment nec	0.491	0.388	74 Transmitters for TV radio and phone	0.112	0.274
73 Electronic components	0.126	0.115	69 Office machinery & computers	0.048	0.052
74 Transmitters for TV radio and phone	0.204	0.020	74 Transmitters for TV radio and phone	0.014	0.032
75 Receivers for TV and radio	0.134	0.091	75 Receivers for TV and radio	0.055	0.126
76 Medical and precision instruments	0.360	0.129	76 Medical and precision instruments	0.036	0.090
77 Motor vehicles	0.200	0.112	64 Agricultural machinery	0.005	0.200
78 Shipbuilding and repair	0.536	0.061	78 Shipbuilding and repair	0.061	0.142
79 Other transport equipment	0.275	0.154	79 Other transport equipment	0.152	0.305
80 Aircraft and spacecraft	0.107	0.011	80 Aircraft and spacecraft	0.011	0.029
81 Furniture	0.233	0.077	81 Furniture	0.053	0.114
82 Jewellery and related products	0.020	0.017	82 Jewellery & related products	0.016	0.044
83 Sports goods and toys	0.014	0.001	12 Grain milling and starch	0.001	0.001
84 Miscellaneous manufacturing nec & recycl	0.543	0.406	56 Metal castings	0.034	0.117

Source: United Kingdom Office of National Statistics, 1995 Input Output Tables

Table A.2: Robustness checks

dependent variable: vi_{ijk}	(1) exclude bottom quartiles of firms by size	(2) exclude top quartiles of firms by size	(3) dependent variable: $\overline{vi}_{ijk}$
Share of costs (jk)	0.201	0.103	0.114
	(0.030)	(0.021)	(0.016)
R&D intensity, producing (j)	0.030	0.014	0.022
	(0.006)	(0.003)	(0.004)
x Share of costs	1.146	0.477	0.554
	(0.389)	(0.254)	(0.224)
R&D intensity, supplying (k)	-0.0074	-0.005	-0.004
	(0.0038)	(0.002)	(0.002)
x Share of costs	-0.953	-0.474	-0.413
	(0.352)	(0.211)	(0.199)
Observations	2,249,095	2,234,459	2,973,008
Covariates	yes	yes	yes

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.

Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. Standard errors in parentheses are clustered at the industry pair level. R&D intensity is R&D carried out in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. In regression with interactions, all main effects evaluated at sample means. Covariates in all specifications include: firm size and age, producing and supplying industry average size and average age.

Table A.3: Nonlinearities

	(1)	(2)	(3)	(4)
dependent variable: vi_{ijk}				
Share of cost:				
2nd quartile			0.0027	0.0024
			(0.0004)	(0.0004)
3rd quartile			0.0092	0.0081
			(0.0007)	(0.0006)
4th quartile			0.0186	0.0176
			(0.0011)	(0.0011)
Producing industry R&D intensity:				
2^{nd} quartile	0.0036		0.0034	0.0015
	(0.0010)		(0.0009)	(0.0008)
3^{rd} quartiles	0.0050		0.0040	0.0028
	(0.0010)		(0.0009)	(0.0008)
4^{th} quartile	0.0057		0.0049	0.0029
	(0.0010)		(0.0008)	(0.0009)
Supplying industry R&D intensity:				
2^{nd} quartile		-0.0080	-0.0062	-0.0049
		(0.0014)	(0.0011)	(0.0010)
3^{rd} quartiles		-0.0039	-0.0019	0.0002
		(0.0014)	(0.0012)	(0.0011)
4^{th} quartile		-0.0043	-0.0040	-0.0008
		(0.0014)	(0.0012)	(0.0012)
Covariates	no	no	no	yes

Source: Authors' calculations using the United Kingdom Office of National Statistics data. All statistical results remain Crown Copyright.
Notes: The dependent variable is a dummy for whether a firm is vertically integrated in that industry pair at any time over the years 1996-2001. There are 2,973,008 observations at the firm-industry pair level. Standard errors in parentheses are clustered at the level indicated. R&D intensity is R&D carried out in the UK divided by value-added produced in the UK, taken from plant level data, aggregated to the 2/3-digit industry level and average over the years 1994-1995. Share of costs is from the 1995 input-output table and is at the industry pair level. Covariates included where indicated are: producing firm size, age, mean firm size and mean firm age in producing and supplying industries. The reference group is zero share of costs and bottom quartiles of R&D intensity in producing and supplying industries.