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,‡
Department of Economics, University College London and IFS

Fabrizio Zilibotti§
Institute for International Economic Studies Stockholm

November 17, 2003

Abstract

This paper investigates the determinants of vertical integration and confronts some of the predictions of the leading approach to the internal organization of the firm with data from the UK manufacturing sector. Consistent with the theory, it shows that an upsteam and downstream activity pair are more likely to be vertically integrated, when the downstream (the producer) is more technology intensive and the upstream (the supplier) is less technology intensive. Also consistent with the theory, the magnitude of both effects are substantially amplified when the upstream inputs are an important fraction of the total costs of the downstream producer. These results are generally robust and hold with a variety of 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: holdup, incomplete contracts, internal organization of the firm, investment, R&D, technology, vertical integration.

Preliminary, Comments Welcome

*e-mail: daron@mit.edu.
†e-mail: paghion@arrow.fas.harvard.edu.
‡e-mail: rgriffith@ifs.org.uk.
§e-mail: fabrizio.zilibotti@iies.su.se.

1. Introduction

Many commentators believe that recent technological developments and globalization are transforming the internal organization of the firm. First, it is argued, new technologies, especially information technology, are creating a shift from the old integrated firms towards more delayered organizations and outsourcing.[1] Second, globalization often creates a tendency to transfer certain labor-intensive parts of the production process to countries with lower wages.[2] Third, it is 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.[3] Summarizing many of these views, the Business Week wrote: "Globalization and the arrival of the information economy have rapidly demolished all the old precepts. The management of global companies, which must innovate simultaneously and speed information through horizontal globe-spanning networks, has become a daunting challenge. Old, rigid hierarchies are out ....".[4]

Despite the importance of these issues in public debate and a large theoretical literature on vertical integration,[5] the economics profession knows surprisingly little about the empirical determinants of vertical integration in general and about the relationship between technological change and vertical integration in particular. This paper provides detailed empirical evidence on the determinants of vertical integration and confronts a number of theoretical predictions with data.

The most influential theory of vertical integration, developed by Williamson (1975, 1985), Klein, Crawford and Alchian (1978), Grossman and Hart (1986) and Hart and Moore (1990), builds on incomplete contracts, relationship-specific investments and "holdup". In the absence of complete contracts, specific investments may not be rewarded appropriately; or put differently, one of the parties in the relationship can hold up the other one after its investments. Vertical integration, or other organizational

[1] Breshanan et al. (1999) find that IT use is associated with more decentralized decision-making within firms. Helper (1991) document the increase in outsourcing in the U.S. automobile industry.

[2] See, for instance, Antras (2003b), Feenstra (1998), Feenstra and Hanson (1996), Grossman and Helpman (2003).

[3] See, for instance, Milgrom and Roberts (1990), Athey and Schmutzler (1995) and Marin and Verdier (2002 and 2003).

[4] Business Week "The 21st Century Corporation", coverstory August 21-28, 2000.

[5] See, among others, Klein, Crawford and Alchian (1978), Williamson (1975, 1985), Grossman and Hart (1986), Hart and Moore (1990), Bolton and Winston (1992), Aghion and Tirole (1994 and 1997) and Legros and Newman (2003), and the survey in Holmstrom and Tirole (1989). See also McLaren (2000), Grossman and Helpman (2002, 2003) and Antras (2003a,b) for models of vertical integration in the context of industry or trade equilibria, and Acemoglu, Aghion and Zilibotti (2003) on the relationship between technical change and vertical integration.

forms, are then ways of mitigating this holdup problem and encouraging greater investments ex ante. The important insight by Grossman and Hart (1986), henceforth GH, is that ownership provides residual rights of control to the revenues of physical and human assets. As a consequence, whether a particular supplier-producer relationship is vertically integrated or not has an important effect on ex ante investments. Perhaps the most important implication of this theoretical approach is that, in order to encourage investment, ownership (and thus residual rights of control) should be allocated to the party with more important investment opportunities. To illustrate this point more concretely, consider a relationship between a supplier (upstream firm) and a (downstream) producer. Also 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. Vertical integration gives greater power to the producer, encouraging its investment, while at the same time reducing the ex post bargaining power and the investment incentives of the supplier. Non-integration, on the other hand, gives greater investment incentives to the supplier. Whether the producer's or the supplier's investments are more important determines whether vertical integration or non-integration generates greater efficiency.

We first develop this insight using a simple theoretical framework building on GH and derive a 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. Consequently:

1. The importance of (specific) investments by the producer and supplier have opposite effects on the likelihood of vertical integration.

2. If the relevant margin is backward integration, vertical integration is less likely when the supplier is more technology intensive, and more likely when the producer is more technology intensive. The opposite results apply when the relevant margin is forward integration.

3. Backward integration is more likely (and forward integration less likely) when the supplier accounts for a larger fraction of the input costs of the producer, because, in this case, there is greater scope for holdup by the supplier, and backward integration protects the producer against this holdup.

4. Vertical integration is more responsive to technology intensity 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 a detailed microdata set on all British manufacturing plants, the UK Census of Production (ARD) between 1992 and 2001. The theoretical predictions above emphasize the importance of looking at the technology intensity of both supplying and producing industries. For this purpose, 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. Clearly, these measures do not distinguish between backward or forward integration, since we do not observe who has residual rights of control.

We proxy technology and investment intensity with a number of different measures: these are R&D to value added ratio, investment to value added ratio, the rate of labor productivity growth and the rate of total factor productivity growth. All of these variables are defined at the industry level and calculated over a sample predating our vertical integration measures (between 1992 and 95, while the vertical integration measures refer to 1996-2001). One advantage of these measures is that they are explicitly about technology, so they are particularly informative about the relationship between technology and the internal organization of the firm. A possible disadvantage is that they include general as well as relationship-specific investments. It is nonetheless highly likely that industries that are technology intensive also have more specific investments than less technology-intensive industries.

We find that:

- Greater technology intensity of the producing (downstream) industry is associated with greater vertical integration, while greater technology intensity of the supplying (upstream) industry is associated with less vertical integration.

- A higher share of cost of the supplying industry in the total cost of the producing industry (for short, "share of cost") is associated with greater vertical integration.

- The effect of the technology intensity of both producing and supplying industries on vertical integration is greater when the share of cost is greater.

These results are generally robust, and they hold with all four measures of technology intensity and with or without a variety of industry and firm-level controls. The results that the technology intensity of producing and supplying industries have opposite effects on the likelihood of vertical integration and that these effects are magnified when the share of costs is larger are consistent with the predictions of the theory. In addition, the rest of the results, in particular, the direction of these effects, are consistent with the theory provided that the marginal form of vertical integration in the data is *backward integration.* Existing evidence from case studies suggests that backward integration is indeed more common in the manufacturing sector than forward integration.[6]

It has to be emphasized at this point 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. However, in theory, and most likely in practice, vertical integration also affects investment. Moreover, other factors omitted in the regression could influence both vertical integration and technology intensity. As an imperfect attempt to deal with these endogeneity problems, we also report results where the technology intensity of each industry is instrumented with the technology intensity of the same industry in the U.S.. This instrumentation strategy generally yields results similar to the ordinary least squares strategy.

Finally, we also investigate the empirical relationship between competition and vertical integration. Although the results here are somewhat less robust than the ones discussed above, they are broadly consistent with the simple theory we outline in the next section, and interestingly, not easily reconcilable with the popular claims that greater competition is leading to the demise of large vertically-integrated firms.

In addition to the theoretical studies mentioned above, this paper is related to a small empirical literature. However, as illustrated by the references in footnote 6, most empirical studies of the internal organization of the firm focus on a single industry; see, for example, the famous study by Joskow (1987). More recently, Baker and Hubbard (2000, 2002) study the trucking industry, Lerner and Merges (1998) look at the biotech sector, and Woodruff (2002) studies integration in the Mexican footwear industry. The only cross-industry evidence relevant to our investigation that we are aware of comes from Antras (2003a) who looks at the share of intra-firm imports over total imports for

[6] See, for example, the studies surveyed in Perry (1989), as well as Monteverde and Teece (1982) on the automobile industry, Stuckey (1983) on aluminum refineries, Masten (1984) on aerospace, and Baker and Hubbard (2000, 2002) on the trucking industry. The only exception is the oil industry, which features both backward and forward integration (see Teece, 1976). In contrast to manufacturing, forward vertical integration seems to be common in retail (see, e.g., Perry, 1989, or Woodruff, 2002).

23 U.S. industries, and relates this to capital intensity. Our data, which are at the firm level, are much more detailed and enable us to investigate more specific predictions on the relationship between technology and vertical integration.

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 results and robustness checks. Section 5 concludes.

2. Theory and Empirical Predictions

2.1. The Environment

The basic model is an extension of GH. We consider a one period relationship between a producer and a supplier, which are both risk-neutral. Both parties can undertake specific investments to increase the productivity of the relationship. Decision rights over these investments cannot be transferred between the two parties, for example, because the investments require specific tacit knowledge or human capital. This implies that the producer cannot undertake 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 respectively the producer and the supplier. 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 are 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 $T_P^z + T_S^z = 0$.[7] There are no credit imperfections, so $T_i(z)$ can be negative.

2. The supplier decides whether or not to accept the offer. If the offer is not accepted,

[7] The feature that the producer makes the organizational form offer is without loss of any generality.

the game ends with payoffs $\{O_P^{NI}, O_S^{NI}\}$ defined below.[8] Otherwise, the producer and the supplier simultaneously choose their investments, e_P and $\hat{e}_S$.

3. The supplier and the producer bargain over the division of the revenue, according to the Nash bargaining solution given the organizational form z.[9] Output is realized and shared.

The production technology of the relationship is:

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

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 increased further by the producer's and the supplier's investments, e_P and e_S. The parameters p and s designate the relative importance of the 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.[10] Note that ϕ also determines the importance of the supplier's investment, $\hat{e}_S$. This production function has also normalized the level of output without 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 that for the results that we want to emphasize, such complementarities are not essential.

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

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

Notice that the investment costs of the supplier are multiplied by ϕ. This ensures that costs are proportional to the scale of operation and that the socially optimal levels of

[8] The assumption that the parties receive their non-integration outside options is not important for the results, and any other combination of reservation payoffs that sum to less than the maximum attainable output of the relationship would give identical results.

[9] Following the other papers in this literature, we are using the Nash bargaining solution. See Binmore, Rubinstein and Wolinsky (1985) for a potential justification for Nash bargaining and a discussion of alternative bargaining rules. Our qualitative results do not depend on Nash bargaining.

[10] 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 by the supplier in question. Even though, 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.

both e_P and $\hat{e}_S$ are independent of ϕ.[11]

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.

With *backward vertical integration* (VIB), the producer owns all the assets, and in the event of ex post breakup, the supplier simply walks away from the firm, and receives nothing. 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 $\lambda \in [0,1)$.[12] 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 that relies on those inputs. The supplier can sell her input in the market, but with 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 looses if she sells the input outside of the specific relationship.[13] The parameter θ is determined by, among other things, the degree of competition in the market, and perhaps the relative number

[11]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 "technologically" 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. The empirical results we present below suggest an unambiguous effect of the share of costs of the supplier in the producer's inputs, and are therefore more consistent with the formulation in the text.

[12]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 a breakup.

[13]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 does not affect our main results.

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 that the supplier loses in case of disagreement.

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

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

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 his or her own outside option, and decreasing in that of the other party. This feature creates a link between outside options and investment incentives, and via this channel, between organizational forms and investment incentives.

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

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

We can now characterize the unique equilibrium of the game specified in the previous subsection. Unless specified otherwise, we refer to an equilibrium by the on-the-equilibrium-path actions and revenues, $\left(\hat{z}, \hat{T}_P, \hat{T}_S, \hat{e}_P, \hat{e}_S, \hat{y}_P, \hat{y}_S\right)$.

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

$$\mathcal{S}^z = U_S^z(y_S^z(\hat{e}_P^z, \hat{e}_S^z), \hat{e}_S^z) + U_P^z(y_P^z(\hat{e}_P^z, \hat{e}_S^z), \hat{e}_P^z),$$

where $\hat{e}_i(z)$ denotes party i's optimal investment under the 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:

$$\mathcal{S}^z = F(\hat{e}_P^z, \hat{e}_S^z) - \Gamma_P(\hat{e}_P^z) - \Gamma_S(\hat{e}_S^z). \tag{2.5}$$

Since both parties have access to perfect credit markets, the subgame perfect equilibrium will always pick the organizational form that maximizes the surplus, $\mathcal{S}$.[14] In other words, $\mathcal{S}^{\hat{z}} \geq \mathcal{S}^{z}$ for all $z \in \{VIB, NI, VIF\}$.[15]

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

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 $\{\hat{e}_S^z, \hat{e}_P^z\}$ such that:

$$\hat{e}_P^z \in \arg\max_{e_P}\{y_P^z(e_P, \hat{e}_S^z) - \Gamma_P(e_P)\} \text{ and } \hat{e}_S^z \in \arg\max_{e_S}\{y_S^z(\hat{e}_P^z, e_S) - \Gamma_S(e_S)\},$$

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

$$\hat{e}_P^{VIB} = p \text{ and } \hat{e}_S^{VIB} = \frac{\lambda}{2}s, \tag{2.6}$$

$$\hat{e}_P^{NI} = \left(1 - \frac{\phi}{2}\right)p \text{ and } \hat{e}_S^{NI} = \frac{1+\theta}{2}s, \tag{2.7}$$

$$\hat{e}_P^{VIF} = \frac{\lambda'}{2}p \text{ and } \hat{e}_S^{VIF} = s. \tag{2.8}$$

These expressions highlight the effect of the different ownership structures on investment incentives. The investment of the producer is highest under backward vertical integration (i.e., $\hat{e}_P^{VIB} > \hat{e}_P^{NI} > \hat{e}_P^{VIF}$), while the investment of the supplier is highest under forward vertical integration (i.e., $\hat{e}_S^{VIF} > \hat{e}_S^{NI} > \hat{e}_S^{VIB}$). Furthermore, most relevant for our empirical analysis, backward vertical integration leads to greater producer investment and lower supplier investment relative to non-integration. This is a fundamental

[14] With credit constraints, the party that is less constrained may end up being 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, 1994, or Legros and Newman, 2003).

[15] Suppose not, and that the equilibrium involves $\hat{z}$, but $\mathcal{S}^{\hat{z}} < \mathcal{S}^{z'}$. 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 his payoff. Namely, he can offer

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

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

result in this class of models: (backward) vertical integration reduces the outside option of the supplier, 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, because 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 this 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 $\hat{e}_S^z$ and $\hat{e}_P^z$ in (2.5), we obtain the total surplus under the three ownership structures $\mathcal{S}^{VIB}$, $\mathcal{S}^{NI}$, and $\mathcal{S}^{VIF}$, and the comparison of the surpluses gives the following proposition:

Proposition 1. *There exist three thresholds $\underline{r}(\phi,\theta,\lambda')$, $\overline{r}(\phi,\theta,\lambda)$, and $\widehat{r}(\phi,\lambda,\lambda')$ such that the unique subgame perfect equilibrium ownership structure, $\hat{z}$, is given as follows:*

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

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

- *If $\underline{r}(\phi,\theta,\lambda') \geq \overline{r}(\phi,\theta,\lambda)$, then $\hat{z} = VIB$ for $p/s > \widehat{r}(\phi,\lambda,\lambda')$, and $\hat{z} = VIF$ for $p/s < \widehat{r}(\phi,\lambda,\lambda')$. Moreover,*

$$\frac{\partial\widehat{r}(\phi,\lambda,\lambda')}{\partial\phi} > 0 \text{ and } \frac{\partial\widehat{r}(\phi,\lambda,\lambda')}{\partial\theta} = 0.$$

This proposition is proved in Appendix A, which also gives the expressions for $\mathcal{S}^{VIB}$, $\mathcal{S}^{NI}$, and $\mathcal{S}^{VIF}$, and for the thresholds $\underline{r}(\phi,\theta,\lambda')$, $\overline{r}(\phi,\theta,\lambda)$, and $\widehat{r}(\phi,\lambda,\lambda')$. The proposition and the proof also make it clear that these thresholds are defined such that when $p/s = \underline{r}(\phi,\theta,\lambda')$, $\mathcal{S}^{NI} = \mathcal{S}^{VIF}$; when $p/s = \overline{r}(\phi,\theta,\lambda)$, $\mathcal{S}^{VIB} = \mathcal{S}^{NI}$; and when $p/s = \widehat{r}(\phi,\lambda,\lambda')$, $\mathcal{S}^{VIB} = \mathcal{S}^{VIF}$.

This proposition gives the most important comparative static results that will be tested 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 will be the equilibrium organizational form; for intermediate values, non-integration may emerge; and when this ratio is small, forward integration will result 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 the supplier's. In contrast, when s is large, backward integration becomes less likely, because now the supplier's investment is more important, and backward integration, by reducing the outside option of the supplier, discourages her investment. The opposite comparative static results apply for forward integration.

Second, as long as we are in the first case with $\underline{r}(\phi, \theta, \lambda') < \overline{r}(\phi, \theta, \lambda)$, non-integration is a possibility, which is clearly the empirically relevant case. In this case, an increase in ϕ makes backward integration relative to non-integration more likely and non-integration relative to forward integration also more likely. 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 means that there are important interaction effects: the effect of p/s on vertical integration is amplified by ϕ. To see the interaction effects in the comparison between non-integration and backward integration, which will be our empirical focus below, let us denote the difference in surplus between these two organizational forms by $\Delta^B \mathcal{S} \equiv \mathcal{S}^{VIB} - \mathcal{S}^{NI}$. Then we have that:

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

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 more producers, after a breakup the supplier is more likely to find a suitable match in the secondary market.

In summary, the most important empirical prediction of this framework is that the technology intensity of the producer and 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; in particular, 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, which appears to be the more common form of vertical integration in the manufacturing sector (see footnote 6), the results suggest that greater technology intensity of the producers should be associated with greater vertical integration, greater technology intensity of the 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 sub-firm level. 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 the supplying industry $k = 1, 2, .., K$:

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

This measure provides a direct answer to the question of whether the firm can supply some of its own inputs k necessary for the production of product j. It does not, however, use any information on how much of its required inputs the firm does (or can) supply from its own plants.

We also construct an alternative measure which uses 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 get 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:[16]

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

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

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

In practice, we will see that there is little difference between the two measures, because when a firm owns a plant in a supplying industry, this is typically more or less sufficient to cover all of its inputs requirements from that industry.

Our main source of data is the annual UK Census of Production (called the ARD or ABI) for the years 1996-2001.[17] This is collected by the UK Office of National Statistics (ONS) and it is a legal obligation for firms to reply. These data provide us with information on input costs and output for all production plants located in the UK at the 4-digit industry level and on the ownership structure of these plants.[18] 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 is available at the 2/3-digit industry level from the Input-Output Domestic Use Table for 1995. The Input-Output table contains information on domestic input flows between 77 manufacturing industries, giving 5,929 pairs of producing-supplying industries, of which 3,840 have positive flows. Appendix Table A3 lists all 77 (supplying) industries together with their largest purchaser.

[16] The actual extent of "vertical integration" may be less than this if a firm produces intermediate goods that could be used in the production process of another plant in the firm, but then sells part of this intermediate good production to the market, and buys the intermediate intput from other suppliers in the market.

Davies and Morris (1995) use an index that is similar in spirit, though is constructed with quite different (and more aggregate) data. Similarly, Fan and Lang (2000) measure corporate relatedness using a measure that is similar in spirit.

[17] In constructing the dichotomous measure, vi_{ijk}, we consider a firm producing product j as vertically integrated in the industry pair jk if it owns at least one plant in industry k in any one of the years 1996-2001. The continuous measure, $\overline{vi}_{ijk}$, instead calculates the average amount of the required inputs that the firm can supply from its own plants over the entire sample period.

[18] Data on employment is available for all plants. Data on other inputs and output is available for all plants with over 100 employees, data from smaller plants are collected from a random stratified sample and values for non-sampled plants are imputed.

Because of the level of aggregation of the Input-Output table, one difficulty arises when we look at industries that use inputs from 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. Overall, the vertical integration measures are constructed using data on 46,392 manufacturing firms over the period 1996-2001.[19]

3.2. Technology intensity and the share of costs

We construct a number of measures of technology intensity at the industry level. Our main measure is R&D intensity, i.e., R&D expenditure divided by total value added, which most closely corresponds to the notion of investment in new technologies. This measure, like all of our technology intensity measures, is calculated on a sample predating the vertical integration sample, in this case 1994-1995 (we do not have R&D information before this date). The total value added in the denominator includes both firms that perform R&D and those that do not.

R&D intensity is our prefered measure because it is a direct measure of investment in new technologies. However, one concern is that the distribution of R&D across industries is rather skewed. Another concern might be that R&D could be spuriously correlated with vertical integration, e.g., because it is better reported in larger firms and larger firms are more likely to be vertically integrated (though, in many specifications, we also control for firm size). We consider alternative indicators of industry technology including physical investment intensity, growth in labour productivity and growth in total factor productivity.

Our first alternative measure is investment intensity, which we expect to be correlated with technology intensity and the extent of relationship-specific investments. This measure is calculated also from the ARD as the average ratio of investment to value added for each industry over the years 1992-1995.

Because some types of technology investments may not be reflected in R&D or physical investment spending, we also use the growth in labour productivity and growth in total factor productivity as additional measures. These are are also constructed from the ARD and are averages for the years 1992-1995. Growth in labour productivity is

[19] We exclude single plant firms with fewer than 20 employees.

measured as:

$$g_{jt}^{LP} = \ln\left(\frac{VA_t}{VA_{t-1}}\right) - \ln\left(\frac{E_t}{E_{t-1}}\right) \tag{3.4}$$

where j indexes industry, VA is value-added and E is numbers employed, and growth in total factor productivity is measured using a superlative index (Caves et al. 1982a,b):

$$g_{jt}^{TFP} = \ln\left(\frac{VA_t}{VA_{t-1}}\right) - \widetilde{s}_{t-1,t}^{l} \ln\left(\frac{E_t}{E_{t-1}}\right) - \left(1 - \widetilde{s}_{t-1,t}^{l}\right) \ln\left(\frac{K_t}{K_{t-1}}\right) \tag{3.5}$$

where $\widetilde{s}_{t-1,t}^{l} = \left(s_t^l + s_{t-1}^l\right)/2$ and s^l is the wage bill over value-added and K is capital stock.

Finally, the share of costs between each industry jk, sc_{jk}, pair 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 high or low R&D intensity, and whether the share of costs (of the producer made up by the inputs from the supplying industry) are high or low. There are 4,352,810 firm-industry-pair observations in 5,929 industry-pairs, of which 3,840 have positive flows. Since pairs without positive flows cannot be vertically integrated, we focus on the 3,840 pairs would positive flows throughout the analysis.

The first row gives the mean and standard deviation of $\overline{vi}_{ijk}$. The mean of 0.008 indicates that only a small fraction of total input flows, less than 1 percent, take place within vertically-integrated structures. But the standard deviation of 0.087 shows that there is substantial variation across firms and at the subfirm level. In fact, there is also substantial variation within industry pairs. To illustrate this, we have also calculated (not shown in the table) the average within-industry-pair standard deviation of $\overline{vi}_{ijk}$, which is 0.086. 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.

We also calculate that the mean of $\overline{vi}_{ijk}$ conditional on $\overline{vi}_{ijk} > 0$ is 0.92 (not reported). This indicates that if a firm can produce some of input k in-house, it can typically produce *all* of that input (k) necessary for production.[20] This motivates our focus on the

[20]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 production-share-weighted average of $\overline{vi}_{ijk}$, i.e., $\sum_k w_{jk}\overline{vi}_{ijk}$, conditional on $\overline{vi}_{ijk'} > 0$ for some k' is 0.053, so firms that are vertically integrated in any one input produce, on average, only about 5% of their total inputs in-house.

simpler variable dummy vi_{ijk}, which designates whether the firm owns a plant producing input k required for the production of good 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 show the differences in the extent of vertical integration when we separate firm-industry pairs into low and high producer technology intensity and supplier technology intensity using the R&D and the investment measures. 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 and the investment intensity of the producing industry are high. For example, the mean of vi_{ijk} for observations where producer R&D intensity is above average is 0.013, while for those where producer R&D intensity is below average, it is 0.008. They also show that there is less vertical integration when the investment intensity of the supplying industry is high. Interestingly, however, they do not show a difference between vertical integration when we cut the sample by whether the R&D intensity of the supplying industry is high or low. The regression analysis below will show a relatively robust effect of supplier R&D and other measures of technology intensity on vertical integration, but the high-low cut in Table 1 does not show this result because of nonlinearities in this relationship (see Table 7).

The table also indicates that there are differences in firm size (but interestingly not in firm age) correlated with producer and supplier technology intensity, motivating the importance of controlling for firm-level covariates in our regression analysis. In addition, the table gives descriptive statistics for the various technology measures. Another noteworthy feature here is that, as also documented in Appendix Table 2, the correlation between the R&D intensity and the other technology intensity measures is quite low. Appendix Table 2 shows that this is also true for investment intensity, but growth in labor productivity and TFP are somewhat more highly correlated. The relatively weak correlation between these measures means that each of these measures is a highly imperfect proxy for the overall technology intensity of the sector, and consequently, there is likely to be considerable downward 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 and investment intensity, so it is useful to look at the relationship between each of them and vertical integration separately.

4. Results

4.1. Main results

Table 2 reports the main results using the R&D intensity measure. 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 share of costs, RD_j^P is R&D intensity in producing industry j, RD_k^S is R&D intensity in supplying industry k, X_{ijk} is a vector including the constant term and firm and producing or supplying 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.[21] We report linear probability models in the text, but in Appendix Table 4 show the main specficiations estimated using a probit models, with very similar results to the linear probability model.

The first two columns of Table 2 look at 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 quite 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

[21] This assumes that the decisions of firms in sector j of whether or not to vertically integrate with supplying industries k and k' are uncorrelated (once we condition on firm and industry-level covariates), and that the decisions of a firm to vertically integrate its operations in industry pair jk and industry pair $j'k'$ are also uncorrelated. We have adopted a number of more conservative strategies, such as clustering only on producing industry j or supply industry k, with little effect on the statistical significance of the main results. In the next version of the paper, we plan to calculate standard errors that allow for the overlapping correlation structure in the variance-covariance matrix.

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 integration. Since we find the same pattern in all of our specifications, from now on we take the relevant comparison to be the one between backward vertical integration and non-integration. Recall that this is also consistent with the evidence from the case studies, which indicates that the most common form of vertical integration in the manufacturing sector is backward integration (see the references in footnote 6).

Note also that the magnitude of the effect of R&D intensity in the producing and supplying industries are quite different, with the effect of the producing industry more than four times the size of the effect of R&D intensity in the supplying industry. We do not have a good explanation for this pattern. However, we will see that with other measures of technology intensity, the effect of R&D intensity in the supplying industry is sometimes larger than the effect of R&D intensity in the producing industry. Therefore, our conjecture is that the differences in the relative magnitudes of these coefficients are due to the fact that each of the intensity measures is capturing a different part of the variation in the sample, combined with potentially heterogeneous effects of technology intensity on vertical integration.[22]

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 (4.1) to:

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

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 γ_S should have the same sign as β_P, so that the effects of R&D intensity in producing and supplying industries are amplified when there is a greater share of costs. Throughout, when we include interaction terms, we report the main effects evaluated at the 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 (-0.009 with a standard error of 0.353).

Columns 5 and 6 add a number of firm and industry-level characteristics, namely firm size and age, and average firm size and average firm age in producing and supplying

[22]See, for example, Angrist and Krueger (2000) on the weighting function of the OLS with heterogeneous coefficients.

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 0.029 in column 6). The effects of these 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.

The results in Table 2 show an interesting pattern of opposite effects from technology intensity in producing and supplying industries. They also show that these effects are magnified when the share of costs that the supplying industry accounts in the total costs of the producing industry is large. These results are consistent with the predictions of the incomplete contracts theory discussed above, and the direction of the effects are also consistent with the theory as long as the relevant margin in the data is backward integration. The rest of the paper shows that these results also hold with alternative measures of technology intensity and are robust to a variety of alternative controls and estimation strategies.

4.2. Results with alternative measures of technology intensity

Table 3 reports estimates from equations (4.1) and (4.2), with alternative measures of technology intensity replacing the RD_j^P and RD_k^S variables. For each of the three alternative measures, investment intensity, growth in labor productivity, and growth in TFP, we report the equivalent of specifications 3 and 6 from Table 2.[23]

The results are generally consistent with those in Table 2. Technology intensity in the producing industry is positively associated with vertical integration, while technology intensity in the supplying industry is negatively associated with integration. For example, column 1, which corresponds to column 3 of Table 2 with investment to value added ratio as the measure of technology, the coefficient on producer technology intensity is 0.030 (standard error of 0.006), while the coefficient on supplier technology intensity is

[23] In the regressions with growth in labor productivity (or growth in TFP), we also include a dummy that takes the value 1 when an industry exhibits negative growth in labor productivity (or TFP) over the entire sample where this is calculated (1992-95). This is motivated by the conjecture that these industries are experiencing a recession or a structural problem that should not have a direct effect on their integration decision. The results are similar when these industries are excluded or when the dummy is omitted.

-0.046 (standard errorof 0.004). The only exceptions to this pattern are for producer technology intensity with investment and growth in TFP and the full set of additional covariates (columns 4 and 6): in this case, although supplier technology intensity continues to be highly significant, producer technology intensity is no longer statistically significant (though with the investment measure, there is a significant effect of producer technology at higher shares of cost).

There is also a positive effect of share of costs in all six columns, and the magnitude of this effect is very similar to those in Table 2. In addition, the interaction terms between the share of costs and technology intensity in the producing and supplying industries also show the same pattern as in Table 2, and are typically significant; the only exceptions are for the interaction terms between the share of costs and technology intensity in the producing industry in columns 2 and 4 (i.e., for labor productivity and TFP growth), whereby the estimates are positive, but statistically insignificant.

The effects of the other covariates in all columns are also similar to those in Table 2. Finally, notice that with investment intensity and growth in TFP as the technology measures, the magnitude of the effect of technology intensity in the supplying industry is greater than that for the producing industry, which contrasts with the pattern for R&D intensity in Table 2 and for growth in labor productivity in Table 3. Our interpretation is once again that different technology measures are capturing a different part of the variation in the sample and that there are potentially heterogeneous effects of technology intensity on vertical integration.

4.3. An instrumental variable strategy

Tables 2 and 3 show statistically significant associations between vertical integration and the technology intensity in the producing and supplying industries. These associations do not necessarily correspond to the causal effects of the technology intensity variables on vertical integration decisions, however. First, as highlighted by the theory above, vertical integration also affects investment, so there is room for reverse causality. Second, and potentially more important, there may be other variables that are omitted from the regressions, which causally affect both technology intensity and vertical integration.

One way to deal with these problems is to use an instrumental variable strategy, with instruments that affect technology intensity, but have no other effect on vertical integration (i.e., they should be orthogonal to the error term, ε_{ijk}, in equations (4.1) and (4.2)). We use measures of technology intensity in the same industry in the U.S. as instruments. These instruments are useful in avoiding the potential reverse causality problems and in removing the effect of UK-specific omitted variables (though this procedure would not remove the effects of omitted industry-specific variables that are common across the US and the UK).

The first-stage equations for the model in (4.1) are:

$$\begin{aligned} RD_j^P &= \alpha^{US,P} sc_{jk} + \beta_P^{US,P} RD_j^{P,US} + \beta_S^{US,P} RD_k^{S,US} + \left(X_{ijk}^{US}\right)' \eta^{US,P} + u_{ijk}^S \\ RD_k^S &= \alpha^{US.S} sc_{jk} + \beta_P^{US,S} RD_j^{P,US} + \beta_S^{US,S} RD_k^{S,US} + \left(X_{ijk}^{US}\right)' \eta^{US,S} + u_{ijk}^S, \end{aligned}$$

where $RD_j^{P,US}$ is technology intensity in the producing industry j in the U.S. and $RD_k^{S,US}$ is technology intensity in the supplying industry k in the U.S..[24]

For the model in (4.2), there are four first-stage equations, since the interaction terms between technology intensity and the share of costs also have to be instrumented, and we create two additional instruments by interacting US technology intensity measures with the UK share of costs.

While we have corresponding measures of investment intensity, growth in labor productivity, and growth in TFP in the U.S., we do not have R&D data in the U.S. at the same level of aggregation. For this reason, in the regressions with R&D intensity, we use all three of the alternative measures of technology intensity as instruments. The IV results for the specifications in columns 3 and 6 of Table 2 and all columns of Table 3 are reported in Table 4. The specifications in the even-numbered columns include the additional firm and industry-level covariates, but these covariates are not reported to save space. The bottom panels of the table reports some of the main first-stage coefficients and the p-value of the F-statistics for the first-stage regressions (except for the R&D regressions where, since there are more instruments, we only report the p-values). All of the first-stage relationships are highly significant and show an intuitive pattern: producer technology intensity in the U.S. has a strong effect on producer technology intensity in the UK, and no effect on the supplier technology intensity in the UK, while this pattern is reversed for supplier technology intensity in the U.S..

The results are generally consistent with those in Tables 2 and 3. The weakest results are for R&D intensity, which is not surprising given the absence of an appropriate instrument for R&D. Recall that R&D intensity is only weakly correlated with the other measures of technology intensity even when we use only UK data (see Appendix, Table A.2). Consequently, the first stages for the R&D intensity variables are relatively weak when we use other measures of technology intensity from the U.S. as instruments.

[24] As in the OLS estimates, the producing and supplying industry R&D intensities are assigned to all corresponding firm-industry pairs, so that both the first and the second stages are estimated at the firm-industry pair level.

Nevertheless, even in this case, the main effects of R&D intensity in producing and supplying industries are of the right sign and significant without the additional covariates, but with the additional covariates, only the R&D intensity in the supplying industry is significant.[25]

The results using investment intensity, growth in labor productivity, and growth in TFP are stronger and more precisely estimated than those for R&D intensity. In all cases, the main effects have the same pattern as the OLS estimates, and are highly significant without covariates, and generally significant with the covariates; the only main effect that is not significant at the 5 percent level is for technology intensity in the supplying industry with growth in TFP, which is, instead, significant at the 10 percent level. Interestingly, the effects of producer technology intensity with either investment intensity or growth in TFP, which were insignificant in the OLS when all the covariates were included, are highly significant in the IV regressions; this is consistent with the presence of downward attenuation bias in the OLS regressions because of measurement error.

However, the interaction terms are typically imprecise and insignificant. This is not surprising, given the difficulty of simultaneously instrumenting for four regressors, some of which are highly correlated. It is remarkable, however, that even with the interaction terms instrumented, the main effects of technology intensity in the producing and supplying industries are reasonably precisely estimated and similar to the OLS results.

Another noteworthy feature is that the 2SLS estimates of the main effects of the technology intensity variables are typically larger than the corresponding OLS values in Table 2 and 3. This might be because the 2SLS procedure reduces the attenuation bias resulting from classical measurement error or from the fact that our measures are only imperfect proxies of the importance of relationship-specific technology investments. Another possibility 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).

Overall, the IV estimates using the U.S. values yield results that are similar to the

[25] If we use the U.S. values of each of the three other measures separately as instruments for R&D intensity in the producing and supplying industries, the results are similar to those reported in columns 1 and 2 with growth in labor productivity and investment intensity, but substantially weaker with growth in TFP.

OLS estimates, and make us more confident that the OLS estimates are informative about the effect of technology intensity on vertical integration decisions.

4.4. Economic significance

How large are the magnitudes implied by the estimates in Tables 2-4? Table 5 answer this question. It shows the mean values of the various technology variables, their standard deviations, the coefficient estimates for on the technology variables in the OLS and IV regressions, and also the coefficients on the relevant interaction terms. Using these values, we calculate the effect of the one standard deviation increase in the technology intensity variable on vertical integration. In interpreting this exercise, it is important to recall that the mean of the vertical integration variable is 0.009, while its standard deviation is 0.091.

The table shows that the magnitude of the main effect is generally small relative to the standard deviation of vertical integration. For example, using the OLS estimate, one standard deviation increase in producer R&D intensity raises the probability of vertical integration by one sixteenth of a percentage point (i.e., 0.0016), which is a very small effect compared both to the mean and the standard deviation of the vertical integration measure. This effect is evaluated at the mean share of producer costs, which is 1%. The presence of only a small effect for the average relationship, which constitutes a very small fraction (i.e., 1%) of the input costs of the producer, is not surprising. More relevant, the table also shows that the impact is much larger when we consider industry pairs where the share of producer costs is larger. If we evaluate the impact of an increase in R&D intensity in an industry pair where the relationship is more important, for example where the share of producer costs is 3% (which is the 90th percentile of the share of cost distribution), the economic impact is considerably larger: one standard deviation increase in producer R&D intensity increases the probability of vertical integration by 2.3 percentage points (i.e., 0.0233). The equivalent effect from increasing supplier industry R&D intensity is to decrease the probability of vertical integration by around 2 percentage points (0.0196).

The same calculations using the IV estimates are shown in the bottom part of Table 5. The effects evaluated at the mean of the share of costs are generally somewhat larger than the OLS, but still small. The effects at the 90th percentile are more imprecise (given the imprecise 2SLS estimates of the interaction terms). Nevertheless, with the exception of R&D intensity which were the least precise results in Table 4, the effects are considerably larger than the effects at the mean.

This table therefore shows that on average the magnitude of the effect of technology intensity on vertical integration is small. This is not surprising given the fact that most industry pairs do not trade much, and consequently, are less likely to be vertically integrated, and changes in their technological characteristics are less likely to have an effect on vertical integration. However, the table also shows that the quantitative effect of technological characteristics is sizable for industry pairs where trade is substantial.

4.5. Robustness

Table 6 shows the robustness of the results in Tables 2 and 3 to a variety of checks. In all cases, we only report the specification corresponding to column 6 in Table 2 (though we do not report the covariates to save space). The top panel excludes pairs where the supplying and producing industries are the same, and shows that this has little effect on the coefficients of interest. For example, the main effect of R&D intensity in the producing industry is now 0.035 (standard error of 0.006), while that for R&D intensity in the supplying industry is -0.005 (standard error of 0.002).

The second panel excludes the bottom quartiles of firms by size, again with little effect on the results. For example, the main effect of R&D intensity in the producing industry is 0.031 (standard error of 0.005), and the main effect of R&D intensity in the supplying industry is -0.011 (standard error of 0.004). The third panel excludes the top quartile of firms by size, again with very similar results. These results show that the patterns documented in Tables 2-4 are not driven by the comparison of the largest or smallest firms to the rest.

Table 7 looks further into potential nonlinearities, and instead of the continuous measure of R&D intensity, it includes dummies for a producing (or supplying) industry being at the second, third or fourth quartile of the corresponding distribution (with the first quartile as the omitted group). The share of costs is also parameterized with dummies for medium and high (with low 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 the smallest effect relative to the omitted first quartile (and is in fact positive in column 3). This non-monotonic pattern is the reason why the difference in vertical integration by the R&D intensity of the suppliers is not visible in the descriptive statistics in Table 1. In regressions using other technology measures (not reported), we find that the coefficients broadly follow the patterns described above; generally, consistent with the theory, there appears to be a monotonic effect of technology intensity; but in some cases, for example, for the investment intensity of the supplying industry and for growth in TFP of both producing and supplying industries, there is a non-monotonic pattern. We do not have a good explanation for these non-monotonic patterns.

A more demanding test of the relationship between technology in intensity and vertical integration is to investigate whether a particular producing industry is more vertically integrated with supplying industries that are less technology intensive relative to its mean propensity to be vertically integrated. Similarly, for a supplying industry, whether it is more vertically integrated with producing industries that are more technology intensive relative to its mean. We cannot investigate both of these questions simultaneously, since equations (4.1) and (4.2) would not be identified with a full set of producing and supplying and industry dummies. However, we can put one set of dummies at a time. In other words, we can estimate

$$vi_{ijk} = \alpha sc_{jk} + \beta_P RD_j^P + \delta_k^S + X'_{ijk}\eta + \varepsilon_{ijk}, \tag{4.3}$$

and

$$vi_{ijk} = \alpha sc_{jk} + \delta_j^P + \beta_S RD_k^S + X'_{ijk}\eta + \varepsilon_{ijk} \tag{4.4}$$

where δ_j^P denotes a full set of producing industry dummies and δ_j^S denotes a full set of supplying industry dummies. Estimates from these equations and from similar ones with interaction terms are reported in Table 8 (the top panel corresponds to equation (4.3) and the bottom panel to (4.4)). The estimates of the main effects are generally similar to those in Tables 2 and 3, and show a positive effect of producer technology intensity and negative effect of supplier technology intensity (with the exceptions of the producer intensity with growth in TFP and investment). For example, the effect of R&D intensity in the producing industry is 0.024 (standard error of 0.004) both with and without the additional covariates, and the effect of R&D intensity the supplying industry is -0.005 (standard error of 0.001) without the covariates and -0.008 (standard error of 0.002) with the covariates.

Finally, Table 9 reports results using the continuous measure $\overline{vi}_{ijk}$ rather than the dummy variable vi_{ijk}. Not surprisingly given the close correspondence between the two measures, the results are very similar to those in Tables 2 and 3.

4.6. The effect of competition and the number of firms

Finally, Table 10 investigates the effect of competition and the number of firms in producing and supplying industries on vertical integration decisions. Table 10 shows the effect of the (log) number of firms in producing and supplying industries, which have statistically significant and robust effects. 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 lower vertical integration.

The coefficient on the number of firms in the supplying industry is about three 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 a good explanation, these results are consistent with the theory. There 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 outside option of the supplier, while more firms in the supplying industry should reduce it, which is the pattern we find data.[26]

Interestingly, if we think of an increase in overall competition as a proportional increase in the number of producing and supplying firms, our estimates suggest that there should be an increase in vertical integration (since the coefficient on the number of supplying firms is larger). Although this result is not our main focus, it sheds some doubt on the popular claims that greater competition is necessarily leading to less integrated firms.

5. 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 confronts some of the predictions of the leading approach, the incomplete contracts theory of the firm, with data from the entire population of UK manufacturing plants, and provides a number of robust and striking results on the relationship between technology intensity and vertical integration.

Our results show that vertical integration in a pair of industries (products) is less likely when the supplying industry is more technology intensive and the producing industry is less technology intensive. Moreover, both of these effects are substantially larger when inputs from the supplying industry form a large fraction of the total costs

[26] We also experimented with Hirfindahl indices calculated separately for producing and supplying industries. Although the Hirfindahl indices were sometimes significant, the results were not robust, and in particular, in specifications with all the covariates, there was no significant effect. 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.

of the producing industry. This pattern of opposite effects of technology intensity of producing and supplying industries is consistent with the incomplete contracts theory of the firm. In addition, the direction of these effects, i.e., that vertical integration is more likely when the producing industry is more technology intensive etc., and the other patterns we report are consistent with the theory provided that the relevant margin in the data is the choice between backward vertical integration and non-integration.

We show that these results are robust to the choice of technology intensity measure and to a battery of robustness checks. We also report similar results instrumenting UK technology intensity measures with U.S. measures. We also 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).

Although these results are not a direct test of the incomplete contracts approach and may also be consistent with other theories, they are striking and robust enough that we believe that future theories should strive to be consistent with these general patterns.

References

Acemoglu, D, Aghion, P, and F. Zilibotti (2003) "Vertical Integration and Distance to Frontier", *Journal of the European Economic Association, Papers and Proceedings.*

Aghion, P and J. Tirole (1994) "On the Management of Innovation", *Quarterly Journal of Economics,* 109, 1185-1207.

Aghion, P and J. Tirole (1997) "Formal and Real Authority in Organizations", *Journal of Political Economy,* 105, 1-29.

Angrist, Joshua D. and Guido W. Imbens (1995), "Two-Stage Least Squares Estimates of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association* 90(430), 431-42.

Angrist, J and A. Krueger (1999) "Empirical Strategies in labor Economics", Chapter 23 in Ashenfelter O and D. Card, eds, *The Handbook of Labor Economics,* Vol III, North Holland

Antras, P (2003a) "Firms, Contract and Trade Structure", forthcoming in the *Quarterly Journal of Economics.*

Antras, P (2003b) "Incomplete Contracts and the Product Cycle", NBER Working Paper No 9945.

Athey, S. and A. Schmutzler (1995) "Product and Process Flexibility in an Innovative Environment" *The RAND Journal of Economics*, 26, 557-574.

Baker, G, and T. Hubbard (2000) "Contractibility and Asset Ownership: On-Board Computers and Governance in US Trucking", NBER Working Paper 7634.

Baker, G, and T. Hubbard (2002) "Make Versus Buy in Trucking: Asset Ownership, Job Design and Information", mimeo Harvard.

Barnes, M. and R. Martin (2002) "Business Data Linking: An Introduction", *Economic Trends*, 581, April, pp.34-41.

Binmore, K, Rubinstein, A, and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modeling", *Rand Journal of Economics,* 17, 176-188.

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

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

Caves, D, Christensen, L, and Diewert, E (1982a) 'The economic theory of index numbers and the measurement of input, output and productivity', *Econometrica*, 50, 6, 1393-1414

Caves, D, Christensen, L, and Diewert, E (1982b) 'Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers', *Economic Journal*, 92, 73-86

Davies, S and Morris, C (1995) "A new index of vertical integration: some estimates for

UK manufacturing" *International Journal of Industrial Organization* 13, 151-177.

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

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

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

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

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

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

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

Holmstrom, B and J. Tirole (1989) "The Theory of the Firm," Chapter 2 in *The Handbook of Industrial Economics,* Vol I, Schmalensee, R and R. Willig, eds, Elsevier.

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

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

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

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

Klein, B, Crawford, R and A. Alchian (1978) "Vertical Integration, appropriable Rents, and the Competitive Contracting process," *Journal of Law and Economics,* 21, 297-326.

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

McLaren, J (2000) "Globalization and Vertical Structure," *American Economic Review,* 90, 1239-1254.

Marin D. and T. Verdier (2002) "Globalization and the Empowerment of Human Capital", Mimeo, University of Munich and Delta.

Marin D. and T. Verdier (2003) "Globalization and the New Enterprise", *Journal of the European Economic Association, Papers and Proceedings.*

Masten, S (1984): "The Organization of Production: Evidence from the Aerospace Industry", *Journal of Law and Economics,* 27, 403-417.

Milgrom, P. and J. Roberts (1990) "The Economics of Modern Manufacturing: Technology, Strategy, and Organization" *American Economic Review*, 80, 511-528.

Monteverde, K and D. Teece (1982) "Supplier Switching Costs and Vertical Integration in the Automobile Industry", *Bell Journal of Economics*, 13, 206-213.

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

Teece, D (1976) *Vertical Integration and Vertical Divestiture in the US Oil Industry*, Institute for Energy Studies, Stanford University

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

Williamson, O (1975) *The Economic Institutions of Capitalism,* Free Press, New York.

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

6. Appendix A: Proof of Proposition 1

First, substituting from (2.6), (2.7), and (2.8) into (2.5), we obtain the levels of social surplus under the three organizational forms as:

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

Next, let $\Delta^B\mathcal{S} \equiv \mathcal{S}^{VIB}-\mathcal{S}^{NI}$ and $\Delta^F\mathcal{S} \equiv \mathcal{S}^{VIF}-\mathcal{S}^{NI}$. Then $\Delta^B\mathcal{S}$ can be written as:

$$\Delta^B\mathcal{S} = p^2\Lambda_P^B - s^2\phi\Lambda_S^B,$$

where

$$\Lambda_P^B \equiv \frac{1}{2}-\left(1-\frac{1-\phi/2}{2}\right)\left(1-\frac{\phi}{2}\right)=\frac{1}{8}\phi^2>0$$

and

$$\Lambda_S^B \equiv \frac{1+\theta}{2}\left(1-\frac{1+\theta}{4}\right)-\frac{\lambda}{2}\left(1-\frac{\lambda}{4}\right)=\chi\left(\frac{1+\theta}{2}\right)-\chi\left(\frac{\lambda}{2}\right)>0$$

where

$$\chi(t)\equiv t\left(1-\frac{t}{2}\right),$$

which is increasing in t for $t<1$, so that $\Lambda_S^B>0$ (recall that $\lambda<1$).

Similarly,

$$\Delta^F\mathcal{S} = p^2\Lambda_P^F - s^2\phi\Lambda_S^F,$$

where

$$\Lambda_P^F \equiv \frac{\lambda'}{2}\left(1-\frac{\lambda'}{4}\right)-\left(1-\frac{1-\phi/2}{2}\right)\left(1-\frac{\phi}{2}\right)=\chi\left(\frac{\lambda'}{2}\right)-\frac{1}{2}+\frac{1}{8}\phi^2>0,$$

(since $\phi<1$ and $\chi(t)>0$), and

$$\Lambda_S^F \equiv \frac{1}{2}-\frac{1+\theta}{2}\left(1-\frac{1+\theta}{4}\right)=\frac{1}{2}-\chi\left(\frac{1+\theta}{2}\right)>0,$$

where the latter inequality results from that fact that $\theta<1$ and that $\chi(t)<1/2$ as long as $t<1$.

Then, the producer and the supplier will be indifferent between VIB and NI if and only if $\Delta^B\mathcal{S}=0$, or equivalently, if and only if

$$\frac{p}{s}=\sqrt{\frac{8(\chi(\frac{1+\theta}{2})-\chi(\frac{\lambda}{2}))}{\phi}}\equiv\overline{r}(\phi,\theta,\lambda)>0.$$

When $p/s > \overline{r}(\phi, \theta, \lambda)$, we have $\Delta^B \mathcal{S} > 0$ and backward integration will be preferred to non-integration, and when $p/s < \overline{r}(\phi, \theta, \lambda)$, we have $\Delta^B \mathcal{S} < 0$ and non-integration will be preferred. Moreover, differentiation establishes that

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

Similarly, the two parties will be indifferent between NI and VIF, if and only if $\Delta^F \mathcal{S} = 0$, or equivalently, if and only if

$$\frac{p}{s} = \sqrt{\frac{\phi(\chi(1) - \chi(\frac{1+\theta}{2}))}{\chi\left(\frac{\lambda'}{2}\right) + \frac{1}{2} - \frac{1}{8}\phi^2}} \equiv \underline{r}(\phi, \theta, \lambda') > 0.$$

When $p/s > \underline{r}(\phi, \theta, \lambda')$, $\Delta^F \mathcal{S} > 0$ and forward integration will be preferred to non-integration, and when $p/s < \underline{r}(\phi, \theta, \lambda')$, $\Delta^B \mathcal{S} < 0$ and non-integration will be preferred. We also have:

$$\frac{\partial \underline{r}(\phi, \theta, \lambda')}{\partial \phi} > 0 \text{ and } \frac{\partial \underline{r}(\phi, \theta, \lambda')}{\partial \theta} < 0.$$

The above analysis establishes that if

$$\underline{r}(\phi, \theta, \lambda') < \overline{r}(\phi, \theta, \lambda),$$

then the equilibrium organizational form is simply given by

$$\hat{z} = \left\{ \begin{array}{cc} VIB & \text{if } \frac{p}{s} > \overline{r}(\phi, \theta, \lambda) \\ NI & \text{if } \frac{p}{s} \in (\underline{r}(\phi, \theta, \lambda'), \overline{r}(\phi, \theta, \lambda)) \\ VIF & \text{if } \frac{p}{s} < \underline{r}(\phi, \theta, \lambda') \end{array} \right\};$$

with indifference at the threshold levels of $\underline{r}(\phi, \theta, \lambda')$ and $\overline{r}(\phi, \theta, \lambda)$. If, on the other hand,

$$\underline{r}(\phi, \theta, \lambda') \geq \overline{r}(\phi, \theta, \lambda),$$

then the above analysis needs to be modified since NI is always dominated by either VIF and VIB. The producer and the supplier will choose VIB when $\mathcal{S}^{VIB} > \mathcal{S}^{VIF}$ or equivalently, when $\Delta^B \mathcal{S} = \mathcal{S}^{VIB} - \mathcal{S}^{NI} > \Delta^F \mathcal{S} = \mathcal{S}^{VIF} - \mathcal{S}^{NI}$, or when

$$\frac{p}{s} > \sqrt{\frac{\phi(\chi(1) - \chi(\frac{\lambda}{2}))}{\frac{1}{2} - \chi\left(\frac{\lambda'}{2}\right)}} \equiv \widehat{r}(\phi, \lambda, \lambda').$$

When the opposite inequality applies, they will choose VIF. Moreover, we have

$$\frac{\partial \widehat{r}(\phi, \lambda, \lambda')}{\partial \phi} > 0 \text{ and } \frac{\partial \widehat{r}(\phi, \lambda, \lambda')}{\partial \theta} = 0,$$

which completes the proof of the proposition.

7. 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 it is a legal obligation for firms 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 appendix for details.

7.1. 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 is also available on all plants (called local units) within each reporting unit, their location and number of employees is also reported. There are over 150,000 reporting units with non-zero employment in the ARD each year 1996-2000. Detailed data is collected from a random stratified sample.[27] Data on value-added and costs for non-sampled reporting units are imputed.

Table 1 shows the total number of reporting units and the number of reporting units in multi and single plant firms by year. 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.

Firms in the ARD are classified by 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 within an IO industry these are aggregated so that there is only one observation per firm in each IO industry. We take an average of all firm level variables over the period 1996-2001. 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. Age is truncated at 21. 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 shows means for these variables.

[27] 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. See Griffith (1999) and Barnes and Martin (2002).

7.2. The Input-Output table

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 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. Appendix Table A.3 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 average 3.7 percent) and the share of the purchaser's total costs this input represents (which ranges from zero to 37 percent and averages 2.7 percent).

7.3. Measuring vertical integration

To calculate both the discrete and the continuous measures of vertical integration, vi_{ijk} and $\overline{vi}_{ijk}$ as given in (3.1) and (3.3), we need information on the production activity of firms at a disaggregated level (to measure p_{ik} and c_{ij}) and information on the cost structure of production (to measure w_{jk}). Our source of data on production activity is at the plant level and is the data that underlies the annual UK Census of Production (called the ARD or ABI). This is collected by the UK Office of National Statistics (ONS) and it is a legal obligation for firms to reply. These data provide us with information on input costs and output for all production plants located in the UK at the 4-digit industry level and on the ownership structure of these plants.[28] These data do not, however, tell us directly whether a plant purchases inputs from a related plant in the same firm. Data on demand for intermediate inputs is available at the 2/3-digit industry level from the Input-Output table.

We implement our measure of vertical integration using data on 46,392 manufacturing firms over the period 1996-2001.[29] This is combined with information from the Input-Output Domestic Use Table for 1995. The Input-Output table contains information on domestic trade flows between 77 manufacturing industries, giving 5,929 pairs of producing-supplying industries, of which 3,840 have positive trade flows. In addition, the Input-Output table indicates the share of total inputs that are imported. In 2,766 industry pairs at least one firm is vertically integrated to some extent. There are 4,352,810 firm-industry-pair observations in 5,929 industry-pairs.

7.4. 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 industry of the R&D reporting unit for the years 1994-1995.

[28] Data on employment is available for all plants. Data on other inputs and output is available for all plants with over 100 employees, data from smaller plants are collected from a random stratified sample and values for non-sampled plants are imputed.

[29] We exclude single plants with fewer than 20 employees.

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. Growth in labour productivity and growth in total factor productivity are measured at the industry level and average over the years 1992-1995 using the ARD data. The correlation between the four measures is given in Appendix Table A.2.

Table 1: Descriptive statistics

Variable	Mean (s.d.)	Producer R&D		Supplier R&D		Producer investment		Supplier investment	
		low	high	low	high	low	high	low	high
Mean $\overline{vi}_{ijk}$	0.008	0.007	0.010	0.008	0.009	0.008	0.009	0.011	0.006
	(0.087)	(0.078)	(0.096)	(0.084)	(0.089)	(0.084)	(0.092)	0.101)	(0.073)
Mean of vi_{ijk}	0.009	0.009	0.013	0.010	0.011	0.010	0.012	0.015	0.007
	(0.091)	(0.093)	(0.114)	(0.101)	(0.104)	(0.100)	(0.108)	(0.120)	(0.085)
Firm age	10	10	10	10	10	10	10	10	10
	(7)	(7)	(7)	(7)	(7)	(7)	(7)	(7)	(7)
Firm employment	111	99	125	109	112	95	143	109	112
	(455)	(346)	(559)	(444)	(465)	(456)	(453)	(447)	(463)
Share of producer costs on this intermediate input	0.010	0.010	0.010	0.012	0.009	0.011	0.10	0.014	0.007
	(0.034)	(0.038)	(0.029)	(0.040)	(0.028)	(0.036)	(0.030)	(0.041)	(0.026)
Producing industry									
R&D over value-added	0.027	0.004	0.055	0.026	0.028	0.024	0.033	0.027	0.027
	(0.055)	(0.002)	(0.072)	(0.055)	(0.055)	(0.047)	(0.067)	(0.055)	(0.055)
Investment over value-added	0.101	0.095	0.109	0.102	0.101	0.079	0.147	0.101	0.102
	(0.041)	(0.031)	(0.049)	(0.040)	(0.041)	(0.017)	(0.038)	(0.041)	(0.041)
Growth in labour productivity	0.058	0.048	0.069	0.057	0.058	0.052	0.069	0.058	0.057
	(0.038)	(0.028)	(0.044)	(0.037)	(0.038)	(0.030)	(0.048)	(0.037)	(0.038)
Growth in TFP	0.042	0.036	0.050	0.042	0.043	0.042	0.044	0.043	0.042
	(0.039)	(0.030)	(0.046)	(0.038)	(0.039)	(0.030)	(0.052)	(0.038)	(0.039)
Mean number of firms in industry	5757	8267	2763	5755	5759	7476	2231	5571	5914
	(6585)	(7978)	(1635)	(6525)	(6636)	(7348)	(1748)	(6331)	(6789)
Supplying industry									
R&D over value-added	0.046	0.044	0.050	0.005	0.082	0.047	0.045	0.031	0.059
	(0.107)	(0.103)	(0.113)	(0.003)	(0.137)	0.109)	(0.104)	(0.078)	(0.126)
Investment over value-added	0.122	0.123	0.122	0.106	0.136	0.122	0.123	0.079	0.159
	(0.057)	(0.057)	(0.057)	(0.038)	(0.067)	0.057)	(0.057)	(0.016)	(0.054)
Growth in labour productivity	0.068	0.067	0.069	0.061	0.075	0.069	0.066	0.064	0.072
	(0.058)	(0.059)	(0.058)	(0.039)	(0.070)	(0.059)	(0.056)	(0.044)	(0.068)
Growth in TFP	0.045	0.044	0.046	0.040	0.049	0.046	0.043	0.050	0.041
	(0.059)	(0.059)	(0.059)	(0.039)	(0.071)	0.059)	(0.058)	(0.045)	(0.068)
Mean number of firms in industry	2316	2320	2309	3347	1433	2296	2355	3730	1120
	(3730)	(3727)	(3733)	(5065)	(1471)	(3715)	(3759)	(4960)	(1324)

Source: Authors' calculations using ONS data. Means reported for 2,973,008 *observations. All statistical results remain Crown Copyright*

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 (0.027)	0.184 (0.028)
R&D intensity, producing (j)	0.038 (0.006)		0.040 (0.005)	0.044 (0.005)	0.038 (0.005)	0.029 (0.005)
x Share of costs				1.112 (0.402)	1.099 (0.398)	1.085 (0.383)
R&D intensity, supplying (k)		-0.010 (0.002)	-0.007 (0.001)	-0.013 (0.003)	-0.013 (0.003)	-0.012 (0.003)
x Share of costs				-0.909 (0.353)	-0.906 (0.351)	-0.916 (0.338)
Firm size (ij)					0.015 (0.001)	0.015 (0.001)
Firm age (ij)					0.042 (0.002)	0.039 (0.002)
Average firm size, producing (j)						0.026 (0.008)
Average firm size, supplying (k)						-0.017 (0.002)
Average firm age, producing (j)						0.003 (0.001)
Average firm age, supplying (k)						0.0001 (0.0005)

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. All variables are normalized by their sample mean.

Table 3: Alternative measures of technology

	(1)	(2)	(3)	(4)	(5)	(6)
dependent variable: vi_{ijk}						
technology variable:	Investment Intensity		growth labour productivity [a]		growth TFP [a]	
Share of costs (jk)	0.203 (0.028)	0.190 (0.024)	0.206 (0.028)	0.219 (0.026)	0.203 (0.028)	0.223 (0.029)
technology intensity, producing (j)	0.030 (0.006)	0.003 (0.008)	0.065 (0.008)	0.036 (0.009)	0.042 (0.008)	0.012 (0.008)
x Share of costs		1.416 (0.457)		0.472 (0.320)		0.155 (0.352)
technology intensity, supplying (k)	-0.046 (0.004)	-0.049 (0.005)	-0.035 (0.004)	-0.018 (0.005)	-0.032 (0.005)	-0.018 (0.004)
x Share of costs		-1.578 (0.493)		-1.353 (0.289)		-1.088 (0.259)
Firm size (ij)		0.015 (0.001)		0.015 (0.001)		0.015 (0.001)
Firm age (ij)		0.039 (0.002)		0.039 (0.002)		0.039 (0.002)
Average firm size, producing (j)		0.042 (0.009)		0.025 (0.008)		0.043 (0.008)
Average firm size, supplying (k)		-0.013 (0.002)		-0.015 (0.002)		-0.016 (0.002)
Average firm age, producing (j)		0.003 (0.001)		0.003 (0.001)		0.003 (0.001)
Average firm age, supplying (k)		0.0007 (0.0005)		0.0002 (0.0004)		0.0003 (0.0004)

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, investment averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. All variables are normalized by their sample mean. [a] Dummies indicating observations where growth in labour productivity or growth in TFP are negative are included.
.

Table 4: Instrumental variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dependent variable: vi_{ijk}								
technology variable:	R&D		Investment		growth labour productivity [a]		growth TFP [a]	
Share of costs (jk)	0.199	0.227	0.203	0.183	0.210	0.225	0.204	0.243
	(0.029)	(0.054)	(0.028)	(0.024)	(0.027)	(0.026)	(0.030)	(0.030)
technology intensity, producing (j)	0.044	0.012	0.083	0.054	0.132	0.100	0.138	0.103
	(0.009)	(0.013)	(0.010)	(0.013)	(0.017)	(0.022)	(0.025)	(0.030)
x Share of costs		0.524		-0.062		0.170		0.535
		(1.187)		(0.971)		(0.853)		(0.918)
technology intensity, supplying (k)	-0.031	-0.020	-0.051	-0.048	-0.115	-0.105	-0.036	-0.035
	(0.004)	(0.011)	(0006)	(0.007)	(0.017)	(0.015)	(0.034)	(0.020)
x Share of costs		-0.808		-2.210		-1.228		-1.923
		(1.396)		(1.043)		(0.483)		(0.615)
Firm and industry controls		yes		yes		yes		yes
First stage, producing industry technology								
US producing industry technology	-	-	1.070	0.998	0.685	0.542	0.438	0.370
			(0.025)	(0.024)	(0.015)	(0.016)	(0.019)	(0.019)
US supplying industry technology	-	-	-0.005	-0.011	0.013	0.010	0.004	0.001
			(0.016)	(0.016)	(0.018)	(0.016)	(0.016)	(0.014)
F-stat P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R^2	0.288	0.342	0.469	0.520	0.300	0.385	0.304	0.362
First stage, supplying industry technology								
US producing industry technology	-	-	-0.003	0.011	-0.0003	-0.024	0.007	-0.021
			(0.031)	(0.033)	(0.0365)	(0.029)	(0.041)	(0.036)
US supplying industry technology	-	-	1.006	0.967	0.524	0.487	0.260	0.395
			(0.020)	(0.022)	(0.033)	(0.028)	(0.039)	(0.031)
F-stat (P-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R^2	0.386	0.391	0.538	0.544	0.326	0.524	0.352	0.476

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, investment averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. All variables are normalized by their sample mean. [a] Dummies indicating observations where growth in labour productivity or growth in TFP are negative are included.

Table 5: Economic significance

Measure of technology	Technology variable: Mean	Technology variable: standard deviation	Coefficient on: technology [a]	Coefficient on: interaction of technology with share of costs [a]	Impact on vi of increasing technology by one s.d.: at mean share of costs (1% of costs)	Impact on vi of increasing technology by one s.d.: at 90 pctile (3% of costs)
OLS estimates						
Producing industry						
R&D	0.027	0.055	0.029	1.085	0.0016	0.0233
Investment	0.101	0.041	0.003	1.416	0.0001	0.0284
Growth labour productivity	0.058	0.038	0.036	0.472	0.0014	0.0108
Growth TFP	0.042	0.039	0.012	0.155	0.0005	0.0036
Supplier industry						
R&D	0.046	0.107	-0.012	-0.916	-0.0013	-0.0196
Investment	0.122	0.057	-0.049	-1.578	-0.0028	-0.0344
Growth labour productivity	0.068	0.058	-0.018	-1.353	-0.0010	-0.0281
Growth TFP	0.045	0.059	-0.018	-1.088	-0.0011	-0.0228
IV estimates						
Producing industry						
R&D	0.027	0.055	0.012	0.524	0.0007	0.0111
Investment	0.101	0.041	0.054	-0.062	0.0022	0.0010
Growth labour productivity	0.058	0.038	0.100	0.170	0.0038	0.0072
Growth TFP	0.042	0.039	0.103	0.535	0.0040	0.0147
Supplier industry						
R&D	0.046	0.107	-0.020	-0.808	-0.0021	-0.0183
Investment	0.122	0.057	-0.048	-2.210	-0.0027	-0.0469
Growth labour productivity	0.068	0.058	-0.105	-1.228	-0.0061	-0.0307
Growth TFP	0.045	0.059	-0.035	-1.923	-0.0021	-0.0405

[a] From column 6 Table 2 and columns 2,4,6 Table 3.

Table 6: Robustness – subsamples

	(1)	(2)	(3)	(4)
dependent variable: vi_{ijk}				
Technology indicator	R&D intensity	investment	GLP [a]	GTFP [a]
exclude industry pairs where producing=supplying, 2,916478 observations				
Share of costs (jk)	0.219	0.213	0.260	0.255
	(0.025)	(0.026)	(0.031)	(0.035)
technology, producing (j)	0.035	0.005	0.038	0.014
	(0.006)	(0.008)	(0.009)	(0.009)
x Share of costs	2.152	1.834	1.031	0.736
	(0.694)	(0.502)	(0.625)	(0.610)
technology, supplying (k)	-0.005	-0.051	-0.020	-0.019
	(0.002)	(0.005)	(0.005)	(0.004)
x Share of costs	-0.023	-1.888	-1.518	-1.194
	(0.226)	(0.475)	(0.371)	(0.326)
exclude bottom quartiles of firms by size, 2,249,095 observations				
Share of costs (jk)	0.203	0.209	0.241	0.244
	(0.028)	(0.023)	(0.025)	(0.028)
technology, producing (j)	0.031	-0.003	0.031	0.009
	(0.005)	(0.008)	(0.009)	(0.008)
x Share of costs	1.156	1.598	0.473	0.198
	(0.395)	(0.455)	(0.325)	(0.365)
technology, supplying (k)	-0.011	-0.051	-0.016	-0.016
	(0.004)	(0.005)	(0.005)	(0.004)
x Share of costs	-0.982	-1.780	-1.400	-1.117
	(0.364)	(0.471)	(0.270)	(0.246)
exclude top quartiles of firms by size, 2,234,459 observations				
Share of costs (jk)	0.106	0.108	0.129	0.131
	(0.022)	(0.021)	(0.023)	(0.026)
technology, producing (j)	0.015	0.012	0.036	0.019
	(0.003)	(0.006)	(0.006)	(0.006)
x Share of costs	0.499	0.636	0.155	-0.019
	(0.266)	(0.388)	(0.249)	(0.272)
technology, supplying (k)	-0.009	-0.032	-0.014	-0.013
	(0.002)	(0.004)	(0.003)	(0.003)
x Share of costs	-0.517	-1.043	-0.929	-0.713
	(0.228)	(0.419)	(0.260)	(0.236)

Source: Authors' calculations using ONS data.
Notes. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. All variables are normalized by their sample mean. [a] *Dummies indicating observations where growth in labour productivity or growth in TFP are negative are included. All regressions include producing firm size, age, mean firm size and mean firm age in producing and supplying industries.*

Table 7: Robustness – nonlinearities

dependent variable: vi_{ijk}	(1)	(2)	(3)	(4)	(5)	(6)
Share of cost:						
Medium	0.0031			0.0039	0.0038	0.0031
	(0.0004)			(0.0005)	(0.0005)	(0.0005)
High	0.0171			0.0167	0.0167	0.0161
	(0.0009)			(0.0009)	(0.0009)	(0.0010)
Producing industry R&D intensity:						
2^{nd} quartile		0.0034		0.0028	0.0019	0.0009
		(0.0009)		(0.0009)	(0.0009)	(0.0008)
3^{rd} quartiles		0.0053		0.0040	0.0036	0.0029
		(0.0010)		(0.0009)	(0.0009)	(0.0008)
4^{th} quartile		0.0054		0.0046	0.0038	0.0021
		(0.0010)		(0.0009)	(0.0009)	(0.0009)
Supplying industry R&D intensity:						
2^{nd} quartile			-0.0076	-0.0059	-0.0059	-0.0056
			(0.0013)	(0.0011)	(0.0011)	(0.0011)
3^{rd} quartiles			0.0032	-0.0021	-0.0021	-0.0012
			(0.0014)	(0.0012)	(0.0012)	(0.0012)
4^{th} quartile			-0.0038	-0.0038	-0.0038	-0.0026
			(0.0014)	(0.0012)	(0.0012)	(0.0012)
Firm size (ij)					0.0148	0.0146
					(0.0009)	(0.0009)
Firm age (ij)					0.0426	0.0400
					(0.0020)	(0.0021)
Average firm size, producing (j)						0.0199
						(0.0081)
Average firm size, supplying (k)						-0.0144
						(0.0029)
Average firm age, producing (j)						0.0028
						(0.0006)
Average firm age, supplying (k)						-0.0006
						(0.0006)

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. Reference group is zero share of costs and bottom quartiles of R&D intensity in producing and supplying industries.

Table 8: Robustness – including producer or supplier industry dummies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dependent variable: vi_{ijk}								
Technology indicator	R&D		Investment		growth labour productivity		growth TFP	
Producing industry effect, conditioning on supplying industry dummies								
Share of costs (jk)	0.167	0.169	0.166	0.168	0.167	0.167	0.166	0.167
	(0.033)	(0.032)	(0.033)	(0.034)	(0.033)	(0.032)	(0.033)	(0.032)
technology, producing (j)	0.039	0.024	0.030	-0.003	0.063	0.035	0.041	0.011
	(0.010)	(0.006)	(0.012)	(0.014)	(0.018)	(0.018)	(0.018)	(0.018)
x Share of costs		0.112		0.257		-0.226		-0.480
		(0.242)		(0.291)		(0.390)		(0.339)
Firm and industry controls		yes		yes		yes		yes
Supplying industry effects, conditioning on producing industry dummies								
Share of costs (jk)	0.205	0.192	0.204	0.195	0.207	0.219	0.205	0.225
	(0.039)	(0.033)	(0.039)	(0.034)	(0.037)	(0.032)	(0.037)	(0.036)
technology, supplying (k)	-0.007	-0.008	-0.047	-0.044	-0.036	-0.019	-0.034	-0.019
	(0.004)	(0.005)	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.012)
x Share of costs		-0.331		-0.731		-1.346		-1.142
		(0.286)		(0.576)		(0.361)		(0.316)
Firm and industry controls		yes		yes		yes		yes

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on producing industry in the top half of the table and on supplying industry in the bottom half. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. [a] Dummies indicating observations where GTFP, GLP is negative are included. All regressions include producing firm size, age, mean firm size and mean firm age in producing and supplying industries.

Table 9: Robustness – dependent variable

	(1)	(2)	(3)	(4)
dependent variable: $\overline{vi}_{ijk}$				
Technology indicator	R&D	Investment	growth labour productivity [a]	growth TFP [a]
Share of costs (jk)	0.116 (0.016)	0.118 (0.014)	0.134 (0.015)	0.137 (0.016)
technology, producing (j)	0.021 (0.004)	-0.0001 (0.0062)	0.029 (0.007)	0.011 (0.007)
x Share of costs	0.567 (0.230)	0.672 (0.302)	0.257 (0.211)	0.084 (0.226)
technology, supplying (k)	-0.007 (0.002)	-0.037 (0.004)	-0.015 (0.004)	-0.016 (0.003)
x Share of costs	-0.447 (0.208)	-0.769 (0.317)	-0.794 (0.164)	-0.666 (0.140)
Firm size (ij)	0.010 (0.001)	0.010 (0.001)	0.010 (0.001)	0.010 (0.001)
Firm age (ij)	0.029 (0.002)	0.029 (0.002)	0.029 (0.002)	0.028 (0.002)
Average firm size, producing (j)	0.022 (0.007)	0.034 (0.007)	0.021 (0.007)	0.034 (0.006)
Average firm size, supplying (k)	-0.014 (0.002)	-0.010 (0.001)	-0.011 (0.002)	-0.012 (0.002)
Average firm age, producing (j)	0.0026 (0.0005)	0.0024 (0.0005)	0.0022 (0.0005)	0.0024 (0.0005)
Average firm age, supplying (k)	0.0003 (0.0004)	0.0007 (0.0004)	0.0003 (0.0003)	0.0004 (0.0003)

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. Dummies indicating observations where GTFP is negative are included. All variables are normalized by their sample mean. [a] Dummies indicating observations where growth in labour productivity or growth in TFP are negative are included.

Table 10: Outside option

dependent variable: vi_{ijk}	(1) R&D	(2) Investment	(3) growth labour productivity [a]	(4) growth TFP [a]
Share of costs (jk)	0.154 (0.026)	0.161 (0.023)	0.189 (0.028)	0.183 (0.026)
technology, producing (j)	0.025 (0.005)	0.0003 (0.0073)	0.010 (0.008)	0.029 (0.008)
x Share of costs	0.846 (0.334)	1.323 (0.427)	0.023 (0.312)	0.396 (0.291)
technology, supplying (k)	-0.002 (0.003)	-0.022 (0.004)	-0.018 (0.004)	-0.003 (0.004)
x Share of costs	-0.728 (0.280)	-1.247 (0.463)	-0.917 (0.277)	-1.170 (0.303)
Firm size (ij)	0.014 (0.001)	0.015 (0.001)	0.015 (0.001)	0.015 (0.001)
Firm age (ij)	0.040 (0.002)	0.039 (0.002)	0.039 (0.002)	0.039 (0.002)
Average firm size, producing (j)	-0.003 (0.008)	0.009 (0.009)	0.012 (0.008)	0.0005 (0.0081)
Average firm size, supplying (k)	-0.013 (0.002)	0.014 (0.002)	0.014 (0.002)	0.012 (0.002)
Average firm age, producing (j)	0.0022 (0.0006)	0.0019 (0.0006)	0.0018 (0.0006)	0.0020 (0.0006)
Average firm age, supplying (k)	0.0038 (0.0005)	0.0038 (0.0005)	0.0038 (0.0005)	0.0038 (0.0005)
Log number of firms, producing (j)	-0.0018 (0.0003)	-0.0019 (0.0003)	-0.0020 (0.0003)	-0.0016 (0.0003)
Log number of firms, supplying (k)	0.0054 (0.0003)	0.0051 (0.0003)	0.0051 (0.0003)	0.0052 (0.0003)

Source: Authors' calculations using ONS data. All statistical results remain Crown Copyright Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. All variables are normalized by their sample mean. LHS variable and firm size averaged 1996-2001, R&D averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. [a] Dummies indicating observations where GTFP, GLP is negative are included.

Table A.1: Number of single, small and single and multi-plant firms, by year

	Number of firm-industry observations	Number of single plant firm observations	Number of single plant firms with fewer than 20 employees	Number of foreign firm-industry observations	Number of multi-plant firm-industry observations	Number of firm-industry observations used
1996	155342	139749	125444	1990	15593	28146
1997	160432	146136	131090	1899	14296	27653
1998	158654	144295	130319	1093	14359	27273
1999	159771	146472	132801	1136	13299	25860
2000	156171	143446	130521	2659	12725	23644
2001	153402	140428	126275	3458	12974	24551

Source: Authors' calculations using ONS data. All statistical results remain Crown Copyright
Notes: Total of 52,918 firms across all years and industries; sample excludes reporting units that report zero employment, industries classifications are those used in 1995 Input-Output Tables..

Table A2: Correlation between technology measures

Technology indicator	R&D intensity	investment intensity	GLP
R&D intensity	-	-	-
investment intensity	0.213	-	-
GLP	0.091	0.303	-
GTFP	-0.024	0.081	0.896

Table A.3: 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: ONS, 1995 Input Output Tables

Table A.4: Probit estimation, marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dependent variable: vi_{ijk}								
technology variable:	R&D		Investment		growth labour productivity		growth TFP	
Share of costs (jk)	0.090 (0.007)	0.073 (0.006)	0.203 (0.028)	0.190 (0.024)	0.089 (0.007)	0.078 (0.006)	0.086 (0.007)	0.077 (0.006)
technology intensity, producing (j)	0.027 (0.002)	0.015 (0.002)	0.030 (0.006)	0.003 (0.008)	0.055 (0.006)	0.028 (0.006)	0.037 (0.006)	0.009 (0.006)
x Share of costs		0.210 (0.091)		1.416 (0.457)		0.059 (0.102)		0.000 (0.119)
technology intensity, supplying (k)	-0.009 (0.002)	-0.005 (0.002)	-0.046 (0.004)	-0.049 (0.005)	-0.035 (0.005)	-0.011 (0.005)	-0.030 (0.005)	-0.011 (0.004)
x Share of costs		-0.183 (0.111)		-1.578 (0.493)		-0.276 (0.072)		-0.206 (0.062)
Firm size (ij)		0.0026 (0.0001)		0.015 (0.001)		0.0027 (0.0001)		0.0027 (0.0002)
Firm age (ij)		0.038 (0.002)		0.039 (0.002)		0.038 (0.002)		0.037 (0.002)
Average firm size, producing (j)		0.024 (0.004)		0.042 (0.009)		0.022 (0.004)		0.032 (0.004)
Average firm size, supplying (k)		-0.039 (0.005)		-0.013 (0.002)		-0.034 (0.004)		-0.034 (0.004)
Average firm age, producing (j)		0.0027 (0.0005)		0.003 (0.001)		0.0024 (0.0005)		0.0026 (0.0005)
Average firm age, supplying (k)		0.0006 (0.0004)		0.0007 (0.0005)		0.0007 (0.0004)		0.0008 (0.0004)

Source: Authors' calculations using ONS data.
Notes: Regression include 2,973,008 observations on 3,840 industry pairs. Standard errors in () are clustered on industry pairs. LHS variable and firm size averaged 1996-2001, investment averaged 1994-1995, age is in 1996 (or 1 if it enters after 1996), share of costs is from 1995 input-output table. All variables are normalized by their sample mean.